11-9-2012

Developing a Process to Create and Validate an Instrument Assessing Student Attainment of Competencies at an Intercultural University in Mexico

Martha Fernanda Pineda

Florida International University, mapineda@fiu.edu

DOI: 10.25148/etd.FI12120701

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

Recommended Citation

Pineda, Martha Fernanda, "Developing a Process to Create and Validate an Instrument Assessing Student Attainment of Competencies at an Intercultural University in Mexico" (2012). *FIU Electronic Theses and Dissertations*. 793.
<https://digitalcommons.fiu.edu/etd/793>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

DEVELOPING A PROCESS TO CREATE AND VALIDATE AN INSTRUMENT
ASSESSING STUDENT ATTAINMENT OF COMPETENCIES AT AN
INTERCULTURAL UNIVERSITY IN MEXICO

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF EDUCATION

in

CURRICULUM AND INSTRUCTION

by

Martha Fernanda Pineda Castilleja

2012

To: Dean Delia C. Garcia
College of Education

This dissertation, written by Martha Fernanda Pineda Castilleja, and entitled Developing a Process to Create and Validate an Instrument Measuring Student Attainment of Competencies at an Intercultural University in Mexico, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Benjamin Baez

Mohammed Farouk

Joan Wynne

Hilary Landorf, Major Professor

Date of Defense: November 9, 2012

The dissertation of Martha Fernanda Pineda Castilleja is approved.

Dean Delia C. Garcia
College of Education

Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2012

© Copyright 2012 by Martha Fernanda Pineda Castilleja

All rights reserved.

DEDICATION

To Gary, with love

ACKNOWLEDGMENTS

Having now reached the finish line, I want to take this opportunity to acknowledge those whose support, in innumerable ways, made it possible. Their love, patience, knowledge, wisdom, honest and critical feedback, time, prayers, and compassion helped me finish this race called “Doctoral Degree.” This is not an exhaustive list of names, so please know that if you were part of my struggle and my joy with this project, you are on my heart’s list!

My first BIG thank you is to God- Thank you for giving me life to learn and share. Your mercy endures forever.

Students from the Intercultural University- Thank you for welcoming, teaching, and inspiring me. Mexico is a better country because of young leaders like you.

Educators from the Intercultural University: Thank you for letting me learn from you and teaching me so much. I respect and commend you for your dedication to the IU’s mission and vision.

To my Committee members, Drs. Landorf, Wynne, Baez, and Farouk- Thank you for your dedication, support, feedback, and openness. I respect each of you profoundly. To Dr. Hilary Landorf, my chair, mentor, and great friend, thank you for your guidance, patience, and support. You have been a great teacher and have seen me through every step of the way in graduate school. My gratitude to you is immense. Dr. Wynne, thank you for helping me keep my assessment-prone vein in line with my social-justice heart. You embody theory, practice, and advocacy like nobody else. I am *your* fan! Dr. Baez, your critical eye and sharp questions made my dissertation a better product. Your cautious approach to assessment helped me to keep my research in check. Thank you!

Dr. Farouk, despite the distance, you were there for me. You introduced me to my early readings of Giroux and Aronowitz in our Curriculum classes. I am hooked. Thank you!

To the College of Education staff: Caprila Almeida, Katie Schmidbauer, Aleida Alfonso, Maria Porras, Susy Torres, Maria Tester - you so much for years of help navigating systems, and all this always done with a smile.

Dr. Linda Bliss, thank you for spending numerous hours reading my drafts, helping me with APA, and reminding me one way or another that researchers do not detach themselves from their products, and that all words carry power and meanings. Your feedback not only improved my paper- it helped my writing and my thinking.

To all my friends and colleagues: Thank you, fellow doctoral students, for your feedback during and after our dissertation boot camp sessions. Having your perspectives was very valuable and helpful. To Stephanie Doscher and Jackie Peña, thank you for being big sisters to me. I am very grateful for the times you sat with me during different stages of this process, helped me, and pushed my thinking to the next level. Claudia Grigorescu and Joan Oviawe: Grad school was so much better with you around, my sisters. I love you. Geraldine Cochran, Tony Delgado, Masha Plakhotnik, and Kathy Perez, thank you so much for your help with APA formatting. To all my friends outside graduate school, especially Claudia Peñuela, Denise Perry, and Lis Palma: thank you for your encouragement, prayers, and understanding. I turned down many invitations to go out so I could finish my dissertation. I can hang out now!

To Drs. Isadore & Carole Newman- Dear Isadore, you came into my life at the right moment: to rescue me from my proposal ditch! And you became my mentor and my friend in the process. I am very grateful for your guidance, kindness, and immense

patience. You kept me focused, and on top of my work's consistency, transparency, and quality. Thanks for taking me under your wing. Dear Carole, thank you for reading my drafts and giving me very useful, insightful feedback all the time. I always received words of wisdom and encouragement from you. I love you both very much. Thank you!

To my parents: Mami and papi, I made it! Thank you for believing in me, loving me, and for cheering me on throughout the process. You believed in a “Dr. Fernanda Pineda” before I learned to read and write. You instilled in me love for learning and a commitment to social justice and solidarity. I hope that my professional and personal life reflect this in as many ways as possible.

Mami: I could not have come this far without your emotional, spiritual, financial, and intellectual support, and all the sacrifices this entailed. From day one of my life you have been there for me with your love and wise guidance. Thank you for your years of sacrifice as I completed my degrees. My education was your priority for as long as I can remember, and you were by my side from beginning to end. You are an inspiration of courage, discipline, integrity, compassion, and hard work, and the human being I admire the most in the whole world. Thank you for teaching me with your example, *má. Te amo.*

To Gary: I could not have done this without you by my side, holding my hand. The day I said, “I need to lock myself up in a room to be able to finish my dissertation” you took on a role that nobody else did. You helped me find big chunks of uninterrupted work, but also found fun things for us to do as a reward whenever we had had a productive day. You even came up with creative alternatives to our dates, like bringing headphones for me to watch football with you in a loud sports restaurant while I worked

on my paper. Those moments were crucial to keep my sanity. You stayed up numerous nights reading and editing my work, and always with a loving attitude. You allowed me to bounce off ideas about content validity, reliability, rubrics, performance-based assessments, research, academia, and everything under the moon! You were there for me to vent, cry, and share my *Eureka!* moments. Both my parents were in the ER and/or hospitalized twice during my dissertation writing process. When my mom was in the hospital you took my place and stayed with her so I could go home to work on my dissertation. When my dad was in the hospital in Mexico and could not come to the community where the IU is located, you took me there giving my family peace of mind. When he was hospitalized in Miami you spent hours playing chess and playing music for him. You helped me rehearse my dissertation defense presentation almost ten times with love and patience, but yet, with a critical stance. When I felt like giving up, you cheered me on with all your warmth and patience, reminding me this was the last stretch, and that awesome opportunities around the world for us were ahead. You are God's gift to me: my love, best friend, partner, coach, logistics pro, fellow progressive thinker, editor, owner of the shoulder I cry on, and more. Your selfless love, solidarity, patience, and unwavering support were crucial for me to finish this project. Thank you with all my heart, Gary. I love you immensely. I dedicate this dissertation to you.

ABSTRACT OF THE DISSERTATION

DEVELOPING A PROCESS TO CREATE AND VALIDATE AN INSTRUMENT

ASSESSING STUDENT ATTAINMENT OF COMPETENCIES AT AN

INTERCULTURAL UNIVERSITY IN MEXICO

by

Martha Fernanda Pineda Castilleja

Florida International University, 2012

Professor Hilary Landorf, Major Professor

This study took place at one of the intercultural universities (IUs) of Mexico that serve primarily indigenous students. The IUs are pioneers in higher education despite their numerous challenges (Bertely, 1998; Dietz, 2008; Pineda & Landorf, 2010; Schmelkes, 2009). To overcome educational inequalities among their students (Ahuja, Berumen, Casillas, Crispín, Delgado et al., 2004; Schmelkes, 2009), the IUs have embraced performance-based assessment (PBA; Casillas & Santini, 2006). PBA allows a shared model of power and control related to learning and evaluation (Anderson, 1998). While conducting a review on PBA strategies of the IUs, the researcher did not find a PBA instrument with valid and reliable estimates.

The purpose of this study was to develop a process to create a PBA instrument, an analytic general rubric, with acceptable validity and reliability estimates to assess students' attainment of competencies in one of the IU's majors, Intercultural Development Management. The Human Capabilities Approach (HCA) was the theoretical framework and a sequential mixed method (Creswell, 2003; Teddlie & Tashakkori, 2009) was the research design. IU participants created a rubric during two

focus groups, and seven Spanish-speaking professors in Mexico and the US piloted using students' research projects.

The evidence that demonstrates the attainment of competencies at the IU is a complex set of actual, potential and/or desired performances or achievements, also conceptualized as “functional capabilities” (FCs; Walker, 2008), that can be used to develop a rubric. Results indicate that the rubric's validity and reliability estimates reached acceptable estimates of 80% agreement, surpassing minimum requirements (Newman, Newman, & Newman, 2011).

Implications for practice involve the use of PBA within a formative assessment framework, and dynamic inclusion of constituencies. Recommendations for further research include introducing this study's instrument-development process to other IUs, conducting parallel mixed design studies exploring the intersection between HCA and assessment, and conducting a case study exploring assessment in intercultural settings.

Education articulated through the HCA empowers students (Unterhalter & Brighouse, 2007; Walker, 2008). This study aimed to contribute to the quality of student learning assessment at the IUs by providing a participatory process to develop a PBA instrument.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Background to the Study.....	1
Problem Statement.....	8
Purpose Statement.....	12
Research Questions.....	13
Conceptual Framework.....	13
Assumptions.....	15
Delimitations.....	16
Significance of the Study.....	16
Definition of Terms.....	16
Summary.....	19
II. LITERATURE REVIEW	19
The Intercultural Universities in Mexico.....	21
Research on Assessment of Competencies in the Intercultural Universities in Mexico.....	23
Creating and Estimating the Validity and Reliability of Rubrics in Higher Education	25
Human Capabilities Approach and Education.....	40
Summary.....	46
III. METHODS	48
Purpose of the Study.....	48
Research Questions.....	50
Research Design.....	50
Sampling	53
Data Collection	55
Data Analysis	72
Data Management	73
Integrity Measures	74
Role of the Researcher.....	78
Summary.....	80
IV. RESULTS	81
Data Analysis	81
Students' Evidence of Attainment of Competencies	82
Estimates of Content Validity	87
Focus Group Participants' Feedback	92
Estimates of Reliability	99
Scorers' Feedback.....	109

	The Researcher's Analysis of the Instrument-Development	
	Steps	114
	Summary	120
V.	DISCUSSION AND RECOMMENDATIONS.....	123
	Summary of the Study	123
	Answer to Research Questions	128
	Analysis of Results	131
	Limitations of the Study.....	136
	Implications for Practice	139
	Recommendations for Future Research and Practice	145
	Conclusions	148
	REFERENCES	150
	APPENDICES.....	160
	VITA.....	226

LIST OF TABLES

TABLE		PAGE
1.	Comparison of the Examples of Practices or Studies Regarding Collaborative Development of Rubrics in Higher Education	36
2.	<i>Documentos Recepcionales</i> Scored and Return Date	101
3.	Total Scores in each Scoring Rounds Limitations of the Study	102
4.	Scores on each Competency and the Inter-rater Agreement Calculations	105
5.	Intra-rater Reliability Estimates	108

LIST OF ACRONYMS

ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
CGEIB	<i>Coordinación General de Educación Intercultural Bilingüe</i> (General Office of Intercultural Bilingual Education)
CVR	Content Validity Ratio
CVI	Content Validity Index
DR	<i>Documento Recepcional</i> (Final Research Project or Thesis)
FC	Functional Capabilities
FIU	Florida International University
FODA	<i>Fortalezas, Debilidades, Oportunidades and Amenazas (SWOT)</i>
HCA	Human Capabilities Approach
INEGI	<i>Instituto Nacional de Estadística, Geografía e Informática</i> (National Institute of Statistics, Geography, and Information Technology)
IU	Intercultural Universities
IMD	Intercultural Management for Development Major (<i>Licenciatura en Gestión Intercultural para el Desarrollo or LGID</i>)
LGID	Licenciatura en Gestión Intercultural para el Desarrollo
SEP	Ministry of Public Education (<i>Secretaría de Educación Pública</i>)
SWOT	Strengths, Weaknesses, Opportunities, and Threats
ToS	Table of Specifications
UM	University of Miami

CHAPTER I

INTRODUCTION

The purpose of this study was to develop a process to create a performance-based assessment instrument that has acceptable standards of validity and reliability estimates. The instrument was designed to assess the attainment of competencies outlined in the Intercultural Management for Development (IMD) undergraduate program at one of Mexico's 10 intercultural universities (IUs). The IUs serve primarily students of indigenous origin. The instrument was created and its content validity estimates were obtained with the participation of IU senior students, one recent graduate, and the researcher. Inter- and intra-rater score reliability estimates were obtained with the participation of seven Spanish-speaking professors in higher education.

This chapter begins with the background to the study, problem and purpose statement, research questions, followed by the conceptual framework. Next, definition of terms, the significance and the organization of the study are discussed.

Background to the Study

Assessment of learning, usually in the form of standardized testing, has gained unprecedented attention in this era of accountability (Benveniste, 2002; Pineda, 2010; Spring, 2008). As a result, assessment of learning is at the forefront of educational policy-making locally, nationally (Linn, Baker, & Dunbar, 1991), and more recently, internationally (Benveniste, 2002; Kamens & McNeely, 2010; Pineda, 2010; Spring, 2008). However, there is growing international evidence that standardized tests are limited in terms of measuring complex performance skills. The research of Darling-

Hammond (1994) in the U.S. and Treviño (2006) and Casillas and Santini (2006) in Mexico illustrate this point.

Notions of political power are associated with testing and accountability (Apple, 2000; Gabbard, 2000). This is particularly salient when dealing with standardized testing and populations such as the indigenous people of Mexico (Schmelkes, 2009; Treviño, 2006). In her study on the assessment practices at one of the IU's, Caro (2010) unveiled the limitations of the use of standardized testing for evaluating indigenous students. She is not alone in making these arguments. Numerous scholars suggest that educational models for indigenous students need to move away from the often times prescriptive, individualistic, subtractive, and even oppressive nature of mainstream education (Casillas & Santini, 2006; Guerra, 2005; Schmelkes, 2009; Treviño, 2006). Scholars also denounce the disparity in achievement between indigenous and mainstream students, usually exacerbated by standardized testing (Casillas & Santini, 2006; Treviño, 2006).

When educators seek an accurate measurement of indigenous students learning, they must acknowledge the inherent limitations of overvaluing standardized tests. An assumption in standardized testing is that the knowledge tested is standard and representative in content (Frederiksen & Collins, 1989, as cited in Messick, 1994), and relevant to all students. This is a useful feature in particular contexts but problematic in more ethnically diverse ones (Pineda, 2010; Treviño, 2006). Acknowledging academic achievement heterogeneity within all ethnic groups, Schmelkes (2009) analyzed recent official indicators regarding indigenous students in higher education and testing. She reports that even though indigenous people make up approximately 10% of the Mexican population, only 1-3% of the students enrolled in higher education are indigenous.

Comparing scores of indigenous students on the Mexican higher education admission test called *Examen Nacional de Ingreso a la Educación Superior*, Schmelkes (2009) found indigenous students scored significantly lower than the national mean across states, and even lower when compared to the private school mean. Another study conducted in Mexico by Parker, Rubalcava, and Teruel (2002) showed that indigenous students consistently scored lower than their non-indigenous classmates, despite the heterogeneity of academic achievement among indigenous students.

Performance-Based Assessment

Standardized tests do not capture the complexity of learning and learners' proactivity (Darling-Hammond, 1994; Messick, 1994), which are pivotal aspects in educational models that aim to assess indigenous student learning from a different paradigm (Casillas & Santini, 2006). Standardized testing may not be the only or ideal way to allow indigenous students to demonstrate the attainment of their competencies (Casillas & Santini, 2006; Treviño, 2006). An alternative is a more flexible, constructivist, authentic, and participatory approach to assess student learning in the form of performance-based assessment (Anderson, 1998; Moskal, 2000). Performance-based assessment is "a process for collecting information through systematic observation in order to make decisions about an individual" (Mertens, 2005, p. 369), in forms different from traditional pen-and-pencil, test formats (Linn et al., 1991). Performance-based assessment is also known as "authentic assessment" (Darling-Hammond, 1994), because students demonstrate that they possess the required skills to apply their knowledge in problem solving in real-world situations (Casillas & Santini, 2006; Darling-Hammond, 1994).

The most commonly used instruments for performance-based assessments are rubrics (Jonsson & Svingby, 2007; Moskal, 2000; Reddy & Andrade, 2009). Rubrics are scoring tools or documents that lay out specific expectations regarding authentic tasks or assignments (Jonsson & Svingby, 2007; Stevens & Levi, 2005), also subject to content validity and score reliability (Onwuegbuzie & Daniel, 2002, 2004; Thompson & Vacha-Haase, 2000; Vacha-Haase, Kogan, & Thompson, 2000; Witta & Daniel, 1998) standards as other assessment instruments (Messick, 1994). Rubrics are “tools to help educators [and learners] establish criteria needed to make decisions to fairly evaluate and assess student work” (Quinlan, 2006, p. 13).

Performance-Based Assessment and Indigenous Students

The use of performance-based assessment is pertinent among indigenous students, or to any minority student group performing markedly lower than counterparts, for several reasons. Performance-based assessment allows for consistency of scoring, the possibility to facilitate accurate judgment of complex competencies, and the promotion of active learning (Jonsson & Svingby, 2007; Messick, 1994; Stevens & Levi, 2005). It also promises a more authentic and direct appraisal of students’ competencies, which eventually reflects positively on educators’ practices (Messick, 1994).

Aside from pedagogical and assessment rationales, there are also educational equity reasons for the use of performance-based assessment. For example, performance-based assessment embraces a shared model of power and control related to learning and evaluation using more democratic processes (Anderson, 1998). With performance-based assessments, students have the possibility of being active stakeholders in their learning

assessment and of exercising self-assessment, both individually and as a group (Anderson, 1998; Stevens & Levi, 2005).

Performance-based assessment also allows students and educators to move away from a model of educational accountability based on high-stakes testing (Anderson, 1998; Mathison & Ross, 2008), documented to have pernicious effects on those students who live in poverty (Treviño, 2006) and/or may not have had the same schooling opportunities (Pineda, 2010; Schmelkes, 2009; Treviño, 2006). Not surprisingly, these student populations living in poverty are usually of indigenous origin (Treviño, 2006). Moreover, in Mexico several educators have also called for performance-based assessment, especially when working with populations like indigenous youth (Casillas & Santini, 2006; Treviño, 2006). In particular, educators at the Intercultural Universities (IUs) have been among the proponents of this approach.

The Intercultural Universities

Since 2003, the education system through the Ministry of Public Education (*Secretaría de Educación Pública*, SEP) has instituted policies that address a multicultural Mexico in higher education via the creation of IUs throughout the nation. The IUs have been created to serve indigenous students from different ethnic groups in the country, but also *Mestizo* students that live in regions with numerous indigenous communities, or that embrace the IU vision. The IUs are considered to be one way to respond to the historical and recent demands of indigenous people (Schmelkes, 2009) for access and representation in Mexican higher education. Despite numerous problems and contradictions in policy and practice (Bertely, 1998; Dietz, 2008; Pineda & Landorf, 2010; Schmelkes, 2009), these IUs are pioneers of higher education in Mexico.

The IUs seek to acknowledge the diverse voices of Mexico, and to create spaces that allow diverse social actors to bring indigenous and rural communities into the dynamics of the modern world in an inclusive and respectful manner (Casillas & Santini, 2006). Educating critical and professional individuals committed to the development of their people and their regions (Schmelkes, 2009; Casillas & Santini, 2006) are also part of the mission of the IUs. Presently there are 10 IUs throughout the country. Seven are directly operated by the General Office of Intercultural Bilingual Education (*Coordinación General de Educación Intercultural Bilingüe*, CGEIB), and three are associated with the network of intercultural universities (Schmelkes, 2009). The IUs are open to indigenous and non-indigenous students, with an intended enrollment quota of at least 70% indigenous youth and at least 20% *Mestizo* or non-indigenous, mainstream students (Schmelkes, 2009).

All IUs offer associate's degrees (*profesional asociado*) and bachelor's degrees (*licenciatura*), with some offering graduate degrees as well (Casillas & Santini, 2006; Schmelkes, 2009). While the institutions across the IU network do not share a standardized curriculum (*plan de estudio*), most offer at least the following five majors: Language and Culture (*Lengua y Cultura*), Sustainable Development (*Desarrollo Sustentable*), Intercultural Communication (*Comunicación Intercultural*), Intercultural Management for Development (*Licenciatura en Gestión Intercultural para el Desarrollo*), and Alternative/Sustainable Tourism or Ecotourism (*Turismo Alternativo o Sustentable*; Casillas & Santini, 2006; Schmelkes, 2009).

There is not a standardized curriculum at the IU because, “in congruence with the recognition of diversity, [the IUs] do not propose a fixed approach to their educational

activities. Each university... [finds] its own way of responding to the characteristics, needs, and potential of the region in which it is located” (Schmelkes, 2009, p. 9).

However, they do share a freshman course. In their freshman year (*ciclo de formación básica*), all students take classes that emphasize language and communication skills, as well as many other skills necessary for college and for achieving the competencies of the IUs (Casillas & Santini, 2006).

This study was conducted at an IU¹ that was established less than half a decade ago, and serves different indigenous groups and *mestizo* students from throughout the state. The majors offered are Sustainable Development, Intercultural Management for Development, and Language and Culture. There were approximately 700 students enrolled as of the summer of 2009. The most popular major is Intercultural Management for Development. The IMD program has the highest enrollment at the IU where this study took place. The IMD program has five pre-established student competencies, indicating that students should be able to (a) articulate regional/indigenous knowledge and initiatives, (b) facilitate advocacy for, and management of resources and information, (c) strengthen and design intercultural initiatives, (d) highlight regional/indigenous knowledge and initiatives, and (e) generate diagnostic and proposal-oriented knowledge.

Problem Statement

The establishment of universities with the pedagogical and sociopolitical goals like the IUs has no precedent in Mexico. Their existence and operations are invariably challenging (Fábregas, 2008; Schmelkes, 2009). Some of the challenges they face include funding, curriculum development, current and historical living conditions of the

¹ To comply with research protocol of protection of subjects, the institution’s name and specific location are confidential

students, political issues (Schmelkes, 2008; 2009), resistance to the revitalization/recognition of indigenous knowledge (Pineda, 2007), and racism (Fábregas, 2008). Accessing government funding and ensuring degree accreditation (Schmelkes, 2009) is also challenging. The IUs have financial constraints due to limited resources (Dietz, 2008; Fábregas, 2008; Guerra, 2005; Schmelkes, 2009).

Despite financial constraints, the IUs strive to engage in quality production of knowledge (Schmelkes, 2009) and overcome unequal educational experiences of the indigenous students (Ahuja et al., 2004; Schmelkes, 2009). To address the resulting educational needs, the IUs have embraced performance-based assessment. The IUs' policy documents state that the student assessment process should be one in which "the student will build his/her knowledge [through] evaluating [both] his/her previous experience and... his/her potential of *knowing how-to-do*" (Casillas & Santini, 2006, p. 207). That is, it should be a performance-based process.

Assessment centered on performance, or performance-based assessment, according to the IUs' policy documents, "implies that students show that they possess the required skills to apply their knowledge of problem solving in diverse situations" (Casillas & Santini, 2006, p. 208). The criteria for assessment should be analyzed in a group setting, and be a product of consensus, and the IUs' student learning assessment "should be continuous, and consider generic and specific competencies developed throughout... the process" (Casillas & Santini, 2006, p. 208).

While conducting a review of the IUs performance-based assessment strategies, no instrument with valid and reliable estimates to assess students' attainment of competencies was found, and no research project to develop such an instrument was

being conducted (CGEIB, 2011) at the writing of this research. Moreover, research is scant on the IUs assessment of learning, including self-assessment, performance-based assessment and pertinent instruments such as rubrics, common tools of performance-based assessment. While the governmental institution that orchestrates the creation and operations of the IUs, the, CGEIB supports numerous research projects, no sponsored research was found on assessment and performance-based assessment instruments at the IUs. Excluding Caro's (2010) master's thesis on the IU's process of student assessment, there were no research projects that addressed learning assessment strategies and/or students' self-assessment strategies at the IUs at the time this study was conducted.

Caro's (2010) educational assessment research at one of the IU's pointed out several future possibilities of research endeavors. In brief, she held interviews with indigenous and *Mestizo* students and coded the data thematically. In addition to highlighting the inadequacy of using standardized testing with populations like IU students, Caro found that students considered the performance-based assessment processes at that university to lack organization and consistency among the students' advisors.

Some of Caro's participants expressed lacking methodological tools and academic support to complete their *documento de reflexión del aprendizaje* (DRs, or Learning Reflection document) as the research projects are called at that IU, despite the 12 mandatory sessions of oral defense practice throughout their university years. Caro (2010) also found that according to many of the interviewed students, communication skills and confidence when defending DRs seemed to carry more weight in obtaining a good grade than evidence of content knowledge. Several of her participants of the *ch'ol*

and *mayo* ethnic groups considered that advisors should be more “strict” while assessing them, and should not give grades as gifts to students. Furthermore, many of her participants considered the defense of their DRs as a source of excitement, genuine learning, and pride. It must be highlighted that this IU does not utilize a performance-based assessment tool such as a rubric to organize and conduct their student assessment, but they do utilize student action-research projects or *documentos recepcionales* (DRs) as their final assessment tool.

The Documentos Recepcionales

From their freshman year IU students spend at least 40% of their time doing field work, and the compilation of their work becomes their final research project (DR; Dietz, personal communication, 2010). These projects are open-ended, action-based, participatory research projects (*investigación vinculada*). Students choose a community group to work with, identify a problem to be addressed, and creatively design solutions to this problem, and have flexibility in choosing their topic and the format of their DR, such as papers, handbooks, or videos.

Students are taught to guide their analyses by the acronym FODA: Strengths, Weaknesses, Opportunities, and Threats (*Fortalezas, Debilidades, Oportunidades and Amenazas*). For example, if a group of traditional healers have numerous plants with different healing attributes and usages and would like to centralize this information, the students create a catalogue and a photographic archive. The strength of this project could be that the group of healers has an archive of their knowledge for posterity and to share with their apprentices. A weakness of this project could be that the procedures might be too complex to be recorded, so the students can only provide a limited record of

those procedures. Opportunities that could emerge from this project for example are that the healers could come together to discuss their needs and desired outcomes, that their knowledge is recorded, and that new healers experience this process early in their development. Threats could be that the act of recording their procedures and cataloguing their plants could be considered inappropriate, should the practices be secrets only known to the healers.

Once the students determine the feasibility of a project, they design a solution or improvement to the need or problem. Students have freedom to choose the area of need related to economic issues or health issues, for example, and present their DR in the desired format, mostly in written format such as a catalogue, a manual, a sales brochure, or production plan. During their senior year, students complete and defend their DR before a committee. Their DR is the main vehicle through which seniors' attainment of their major's competencies is evaluated, and the only graduation prerequisite.

Purpose Statement

Based on an extensive literature review on performance-based assessment (PBA), the SEP IU policy book, findings of Caro's (2010) study, and personal communication with SEP and IU educators, the need for a PBA instrument, or a rubric, was established. The purpose of this study was to develop a process to create a performance-based assessment instrument with acceptable standards of validity and reliability estimates, for measuring students' attainment of competencies at one of the IU's undergraduate programs, the Intercultural Management for Development (IMD). Based on a search of the established PBA procedures at the IU where this study took place, it was determined

that the instrument would be piloted using *documentos recepcionales*, unless this study's participants expressed otherwise.

Research Questions

The research questions guiding this study were the following:

- (a) According to IU seniors, what evidence demonstrates the attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?
- (b) Can acceptable standards of validity and reliability estimates be established for an instrument developed to measure students' attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?

Conceptual Framework

The conceptual framework guiding this study is the human capabilities approach (Nussbaum, 1997; Sen, 1995, 1999; Walker, 2008). The writings of philosopher Sen and practitioners such as Walker provide theoretical foundations of the human capabilities approach.

The Human Capabilities Approach

The human capabilities approach, originally developed by Sen, winner of the 1998 Nobel Prize for Economics, is a recent way to interpret, evaluate, and advocate for human development. In brief, Sen's capabilities approach allows for contextually evaluating well-being from a broader perspective than the utilitarian economic approach of equating income with well-being. This approach highlights human capabilities,

understood as potential functionings, which pertains to what people are able to do or achieve, that is, their freedoms.

The approach is “a concentration on [people’s] freedom to achieve in general, and the[ir] capabilities to function in particular” (Sen, 1995, p. 266). Freedom relates to “the range of options a person has in deciding what kind of life to lead” (Drèze & Sen, 1995, p. 10). Functionings are achievements, whereas capabilities are abilities to achieve (Sen, 1999). Capabilities comprise what a person is able to do or be, such as “the ability to be well nourished, to avoid escapable morbidity or mortality, to read, write and communicate, to take part in the life of the community, to appear in public without shame” (Sen, 1990, p. 126, as cited in Saito, 2003). Sen (1992) highlights the difference between actual achievement and the freedom to achieve: “Achievement is concerned with what we manage to accomplish, and freedom with the real opportunity that we have to accomplish what we value” (p. 31, italics in original).

The decision-making process regarding which capabilities comprise the kind of life an individual or a community values is facilitated through democratic dialogue (Sen, 1999). Democratic dialogue is the engagement of public discourse to negotiate and reach consensus on values, goals, and priorities (Sen, 1999). Every individual and community should be able to develop his/her/their own capabilities in a participative process, and education is fertile soil for the inclusion of democratic dialogue (Landorf & Pineda, 2010).

In education, the HCA has been used as a framework for assessing student learning and overall experiences in higher education from the students’ perspective, (Walker, 2008). In the HCA, assessment is not restrictive or standardized, but based on

students' input and negotiation to establish what they value in their education, as Walker (2008) argues. The HCA has enormous potential to address complex issues of equality, quality, and measurement (Unterhalter & Brighouse, 2007). This approach calls for educators to move away from narrow definitions of educational achievement, providing a more comprehensive analysis of complex settings. The HCA allows for the analysis of "aspects of education deemed valuable and hence issues about... complex class, gender, race, and ethnic inequalities" (p. 73).

The IUs are institutions designed to move away from the historic failures of assimilationist education for indigenous groups and a utilitarian approach to education (Casillas & Santini, 2006). This can be done several ways, such as exploring different ways of being and learning, and implementing the acquired and existing types of knowledge or *saberes*, comprised of indigenous knowledge and academic knowledge (Casillas & Santini, 2006). A concern of many indigenous leaders is the lack of opportunities in their communities that force many young people to leave, as opposed to leaving by choice. Many students from these indigenous communities enroll in urban universities and often study subjects unconnected to the development of their communities of origin (Ahuja et al., 2004). The HCA, with its overriding concern for the quality of people's life and keen criticism of social arrangements that limit people's freedoms (Sen, 1992, 1999; Walker, 2008), responds to the mission of the IUs.

Assumptions

An assumption of this study was the author's full access to participants and the truthfulness of their responses during the focus group(s). It was also assumed that the participants were representative of the population, and that they would be able to work as

a group and communicate their experiences at the IU, related to the attainment of the competencies of their major.

Delimitations

The scope of this study was delimited to one institution in the network of 10 IUs in Mexico, as of January of 2011, and only seniors in and one graduate of the IMD major participated.

Significance of the Study

This study sought to contribute to the quality of the IUs' assessment of student learning and students' final research projects by providing a process for developing a performance-based assessment instrument with acceptable standards of validity and reliability estimates. Because of their commitment to battle against centuries-long inequalities, the IUs promote their mission and vision based on the quality of their students' education, and the impact that students' work could make on development of their communities (Fábregas, 2008). It is hoped that this instrument-development process can be used across the IU network.

Definition of Terms

Capability

Defined as “[T]he ability to achieve” (Sen, 1987, p. 36), a capability is what a person is able to do or be, or value doing or being (Sen, 1999). Examples of capabilities are “the ability to be well nourished, to avoid escapable morbidity or mortality, to read, write and communicate, to take part in the life of the community, to appear in public without shame” (Sen, 1990, p. 126).

Competencies

“[T]he collection of knowledge, skills, attitudes, values, and responsibilities that the student develops in her/his forming trajectory to successfully complete a particular professional activity with a high sense of commitment and professional ethic” (Casillas & Santini, 2006, p. 185).

Democratic dialogue

It is public discourse and the assemblage of community consensus on values, goals and priorities (Sen, 1999).

Freedom

“[T]he range of options a person has in deciding what kind of life to lead” (Drèze & Sen, 1995, p. 10).

Functionings

Functionings are actual achievements, and are related to the way and the conditions in which people live their daily lives (Sen, 1995). Functionings reflect what a person may value doing or being (Sen, 1999).

Human Capabilities Approach

The Human Capabilities Approach is a philosophical approach to human development and well-being. It focuses on the “freedom to achieve in general and the capabilities to function in particular” (Sen, 1995, p. 266).

Indigenous

In this study, “indigenous” is the term related to ethnic groups (*pueblos o grupos indígenas*). It is a concept of colonial origins that defines a population “that shares a cultural tradition of Prehispanic roots... and retains among its most salient features

speaking an Amerindian language or the assuming of an identity attached to this tradition” (Zolla & Zolla, 2004, para. 1).

Indigenous knowledge

Knowledge that is local and community-rooted. It is usually taught by elders or otherwise “wise people,” members of indigenous communities, to transfer their usually problem-solving based empirical knowledge (Mendoza, 2009).

***Mestizo* culture**

The culture of people who are ethnically of mixed descent, predominantly of European/Spanish, and indigenous people. The language used is usually Spanish (Mendoza, 2009).

Performance-based assessment

The kind of assessment that requires students to perform to demonstrate that they have attained the required skills, and that they can apply their knowledge and problem-solving skills to different situations (Casillas & Santini, 2006).

Rubrics

Authentic assessment scoring tools that serve as vehicles to articulate specific expectations regarding authentic tasks or assignments (Jonsson & Svingby, 2007; Stevens & Levi, 2005).

Summary

This dissertation contains five chapters. This first chapter serves as the introduction of the study, whose purpose was developing a process to create an instrument to measure the attainment of competencies outlined in the program of study of the Intercultural Management for Development major at one of Mexico’s intercultural

universities. This instrument was created with the participation of senior students and a graduate and the researcher, and piloted by a group of Spanish-speaking professors in higher education. The conceptual framework for this study was the Human Capabilities Approach.

Chapter 2 provides a review of the literature and Chapter 3 describes the sequential mixed methods design along with the procedures for data collection, analysis, and management. Integrity measures are also discussed. Chapter 4 presents the results of this study, and Chapter 5 contains findings, implications, conclusions and recommendations for further research.

CHAPTER II

LITERATURE REVIEW

The chapter is divided into four sections: (a) an overview of Mexico's intercultural universities, (b) research on assessment of competencies at the intercultural universities of Mexico, (c) research on creating rubrics collaboratively in higher education, and (d) the human capabilities approach to education. A description of the search procedures including sources, languages, relevant key words, and timeframes precedes the discussion on the literature review findings. The chapter concludes with a summary.

The first section of this chapter provides further background information on the IUs. The second section of this review is a summary and discussion on the sole piece of literature on the subject pertinent to this dissertation study, a relevant master's thesis conducted at an IU (Caro, 2010).

The third section examines research on creating rubrics collaboratively in higher education. Based on several studies (e.g., Parker et al., 2002; Schmelkes, 2009; Treviño, 2006) and the IUs' policy documents (Casillas & Santini, 2006), an argument for performance-based assessment practices was presented in Chapter 1. It was also argued that performance-based assessment alone does not guarantee equity or best educational practices for indigenous students, such as those enrolled in Mexico's IUs. Poor instrument design, a lack of active collaboration among stakeholders, and an inconsistency between content and pedagogy can all be detrimental (Darling-Hammond, 1994). Hence, the third section of this chapter examines the body of literature for examining the best practices for collaborative instrument-making in higher education,

and possible gaps in the research. Because rubrics are the most common tool for engaging in performance-based assessment (Jonsson & Svingby, 2007; Moskal, 2000; Reddy & Andrade, 2009) a search for relevant studies on rubrics was also conducted. The fourth section of this review is devoted to the Human capabilities approach to education. A review of key writings of this approach is included.

This literature review was conducted predominantly online by accessing databases such as JSTOR, ERIC, Science Direct, and WilsonWeb. Additional electronic and/or printed journals and books were accessed through the Florida International University and University of Miami library systems or from journals' websites directly. Websites of Mexican institutions such as the Ministry of Education, Office for the Development of Indigenous Communities, and the General Office of Intercultural Bilingual Education, were also accessed. Other seminal works were obtained through the researcher's contacts: researchers at the IUs, an expert methodologist in the United States, and a fellow doctoral candidate. The research reviewed for this study was in English and Spanish, languages in which the researcher is fluent.

The inclusive search terms used for the review on creating rubrics were "rubric(s)," "rubrics in higher education," "collaboration and rubrics," "developing instruments and collaboration," "performance-based assessment rubric(s)," "performance-based assessment instrument(s)," and "alternative assessment." For empirical research articles, only higher education studies were included. For the intercultural universities section, the search terms were "intercultural universities and Mexico." In Spanish, these search terms were "*universidades interculturales y México*." The search for literature on the section on the Human capabilities approach was

predominantly based on Sen's work. In addition, search terms such as "Human Capabilities Approach and education," "Human Capabilities Approach and higher education," and "Human Capabilities Approach and intercultural/indigenous education" were used.

No timeframe was set for the search of material for the second, third, and fourth sections, which were devoted to the IUs and the HCA. For the first section, empirical research articles and applications from 2000-present were given priority to limit the possibilities of obsolete practices. Only one study and one model published prior to 2000 are included.

The Intercultural Universities in Mexico

Historically, indigenous populations have experienced subtractive public education (Bertely, 1998; Mendoza, 2009; Treviño, 2006). This means that attending public school became synonymous with the experience of forgetting their indigenous language, and having it be replaced by Spanish (Mendoza, 2009). Often times, many curricula did not offer contextualized learning and appropriate cultural contents for these groups (Mendoza, 2009).

For decades, these had been the educational experiences of indigenous students (Bertely, 1998; Pineda & Landorf, 2010). There were some incipient political and educational movements in the 1960's and 1970's that allowed the indigenous populations to gain terrain in the owning of their education, paving the way for the intercultural education initiatives as we know them today (Mendoza, 2009; Pineda & Landorf, 2010).

Particularly since the decade of the 90's, Mexico's public education system has transformed. A key event was the 1996 uprising of indigenous and peasant groups in

the Southern state of Chiapas, also known as the Zapatista Movement, organized by the *Ejército Zapatista de Liberación Nacional* or *Zapatista National Army for National Liberation*. “¡Nunca más un México sin nosotros!” (“Never again a Mexico without us!”), was the Zapatista declaration, crystallizing the centuries-long struggle of indigenous groups to be recognized, respected and treated with justice (Pineda & Landorf, 2010). Since the 1990’s, the Mexican government implemented different intercultural bilingual programs from the elementary to high school education level. However, it was not until the early 2000’s that intercultural universities were established. Mexico is not alone in its initiatives establishing intercultural universities in regions with high concentrations of indigenous populations. This has been a widely seen phenomenon throughout Latin America (Mato, 2008).

The Mexican intercultural universities are public institutions of higher education, located in regions with high concentration of indigenous population such as Chiapas, Guerrero, Oaxaca, Tabasco, and Veracruz (Casillas & Santini, 2006; Mendoza, 2009; Schmelkes, 2008; 2009). A network of 10 IUs has been established in Mexico. The vision of the IUs includes many tenets. These include the educating professionals and intellectuals committed to the development of their communities and regions (Casillas & Santini, 2006). These individuals should be able to own their academic formation by first strengthening the knowledge and appreciation of the students’ roots first (Casillas & Santini, 2006).

The IUs consider research as fundamental, and they emphasize research on language, culture, and regional development. Among their goals is to reevaluate, revitalize, and consolidate languages, knowledge and cultural expressions of Mexico’s

indigenous communities (Casillas & Santini, 2006; Dietz, 2008; Schmelkes, 2008; 2009). Moreover, knowledge formation is also a priority at the IUs. Particularly at the IU where this study took place, much importance is given to the connections between Western or conventional, and indigenous knowledge (Dietz, 2003, as cited in Mendoza, 2009), as well as the inclusion of processes where inter- and cross-cultural dialogues take place among “wise people” or the elders of the community, students, researchers, and the community at large (Mendoza, 2009).

The project of the IUs is groundbreaking in Mexican higher education; however, the endeavor is not free from challenges. Documentation on the IUs proactive involvement in the inclusion of other minorities in the country such as the Afromexicans was not found. The inclusion of other ethnic minorities with similar socioeconomic circumstances as the indigenous minorities is necessary if the Mexican Ministry of Education strives to eventually have intercultural education for all (Ahuja et al., 2004; Schmelkes, 2009). These institutions are young and are addressing millenary debts to minorities in the country. Further research on how to strengthen these initiatives should take place.

Research on Assessment of Competencies in the Intercultural Universities in Mexico

The search for performance-based assessment and self-assessment projects regarding the IUs yielded no results. Only a thesis accessed directly through the IU student advisor was retrieved. Caro’s (2010) thesis was conducted at an institution that is part of the IUs’ network. This was a qualitative study that consisted of a series of interviews with *ch’ol*, *huichol*, *yoreme*, *mayo*, and *mestizo* students regarding their perception of the graduation process in place at their IU.

The graduation process at this university is performance-based assessment. There are no standardized exams, although there are term exams. No description was provided on these exams. The IUs students create a final research project called the “document of reflection of learning” (*documento de reflexión del aprendizaje*) which can be completed individually or in groups. The completion of this document is based on students’ knowledge, and on what he/she is challenged to research and understand more deeply. In other words, it aims to allow students to demonstrate their knowledge and abilities. Once all the final research projects are completed, a committee reviews them, and the student(s) defend the project before the committee. It was not clarified if the defense is done individually or in groups. The students’ final research project is competencies-based and it is revised twice throughout the student’s program, along with 12 sessions of preparation. If a student does not pass the oral defense the first time around, he/she receives oral feedback from the committee, suggesting things to work on and is given another chance until satisfactory results are achieved.

Caro (2010) analyzed the interviews and coded the data into general themes and then refined the data analysis to find themes that were particular to one or more indigenous groups. Some of her findings were that the five groups of students found the process to be lengthy, and that their good communication skills and confidence when defending their thesis seemed to help more than their actual content knowledge. Only the *ch’ol* and *huichol* students reported that the graduation process was tailored towards understanding the learning process; in knowing the “how” as much as the “what.”

Other findings and discussions in Caro’s (2010) study suggested that the process is problematic for some of the students because of the lack of uniform criteria for

assessment, with the processes left up to each individual bachelor program. A problem is that the tutors, who prepare students for completing their final research projects, do not have a clear procedure to follow. Students complained about the flaws in procedures, injustices, and a lack of objectivity in the system.

Creating and Estimating the Validity and Reliability of Rubrics Collaboratively in Higher Education

This section is divided in three subsections. The first section focuses on two meta-reviews on the development, validation, and use of rubrics. The second section provides an in-depth examination of some examples of studies and applications. The third subsection presents two models of instrument development that can be adapted to developing rubrics.

Research on and the use of rubrics in higher education has exponentially grown in the last two decades in the US and around the world. Particularly useful for this study was the work of Jonsson and Svingby's (2007) and Reddy and Andrade's (2009). These two studies provided meta-reviews of literature on empirical research regarding rubric use. Jonsson and Svingby's (2007) review included a variety of educational levels, whereas Reddy and Andrade (2009) focused their review strictly on higher education. Both reviews suggested the invariability of the benefits of rubrics as performance-based assessments. Among the benefits are the promotion of learning and achievement among students, and the improvement of procedures to give and receive timely and thorough feedback. Rubrics also have potential benefit also for teaching and even program evaluation (e.g., Dunbar, Brooks, & Kubicka-Miller, 2006; Knight, 2006; Oakleaf, 2006). And, according to both meta-reviews, rubrics increase trustworthiness and consistency in

scoring. This quality highlights rubrics' value when used in performance-based assessment.

Both reviews highlighted the widely-used practice of including students' views on rubrics, but seldom as active developers of rubrics. The Reddy and Andrade (2009) review included a more detailed discussion on students' perceptions on the use of rubrics. Reddy and Andrade reported several studies (e.g., Andrade & Du, 2005; Bolton, 2006; Powell, 2001; Schneider, 2006) indicating that "students value rubrics because they clarify the targets for their work, allow them to regulate their progress and make grades transparent and fair" (p. 438).

Neither Reddy and Andrade (2009) nor Jonsson and Svingby (2007) reported finding a process where the students participate not only in providing quality description of work for the different grading levels of the rubric, but in providing the evidence, in portfolio format or any format, to use a rubric. All the studies that these two set of authors reported on have predetermined tasks to be completed and students' input is focused on describing attainment levels. This is a gap in the research that this study addressed by placing students' input at the center of the rubric-development process. The students in this study provided content for the rubric by discussing, negotiating, and describing what the competencies at the IUs mean, what each attainment level should include, and the evidence that would demonstrate their attainment.

Another highlight of the Jonsson and Svingby (2007) and Reddy and Andrade (2009) reviews was the diverse approaches and findings regarding rubrics' validity and score reliability estimates. The two reviews pointed out how pivotal clear language, expectations and criteria, and a comprehensive framework for validity and reliability

estimates are. Both reviews found that inter-rater score reliability estimates in the form of consensus agreements are among the most popular, and are crucial for increasing trustworthiness and consistency. Pearson's correlations, Cronbach's *alpha*, and Spearman's *rho* were the most commonly used. Jonsson and Svingby (2007) argued however, that among the major threats to the reliability of rubrics is the lack of consistency of individual scorers. In practice, when the rubric is implemented, it is usually one professor/teacher scoring students' work. Out of 75 studies that Jonsson and Svingby (2007) reviewed, more than half reported inter-rater score reliability in some form, but only seven reported intra-rater score reliability estimates.

Jonsson and Svingby (2007) reported that out of the 75 studies, only 25 reported on validity estimates. The most common validity estimate practices that both reviews found were content, criterion, construct validity, and using experts in the process. Also called "logical validity," content validity estimates how representative are the items on an instrument of the content matter the instrument is designed to measure (Newman et al., 2006). Criterion validity allows the researcher to estimate validity by examining the relationship between the instrument, and the scores of another well-established criterion variable (Johnson & Christensen, 2004). Criterion validity is concurrent and predictive validity estimates combined (Newman et al., 2006). This means that criterion validity estimates are based on the simultaneous relationship between the test score and a criterion score, or "concurrent validity," or an anachronous relationship between the test score and a criterion score obtained later, or "predictive validity" (Johnson & Christensen, 2004). Construct validity is a conglomeration of all types of validity estimates, important to use when interpreting tests scores as measures of a construct or

attribute (Newman et al., 2006). If the IUs' community members consider a criterion variable to be adequate in the future, construct validity estimates could be calculated in the future.

While an instrument can have high reliability estimates without validity (Newman et al., 2006), reliability is a necessary yet insufficient aspect (Reddy & Andrade, 2009) of validity, but when an instrument is valid, it is also reliable (Newman et al., 2006).

However, only four of the 20 studies that Reddy and Andrade (2009) reviewed reported validity estimates. Hence, more time and attention should be invested on rubrics' validity and clearly, not ignoring score reliability (Newman et al., 2006; Reddy & Andrade, 2009). This dissertation implemented two validity procedures to increase the methodological rigor in the IUs rubric: content validity estimates with experts, and the calculation of Lawshe's Content Validity Ratio (CVR). Of the studies reviewed herein, none reported the use of the CVR; no study was found discouraging its use. Lawshe's CVR is a useful tool for assuring that each item on an instrument proposed for inclusion is judged relevant to the content, by a group of expert judges (Guion, 1977).

Studies and Applications of Rubrics in Higher Education

There were several studies and applications regarding the use of rubrics in higher education accessed in this review, and selected examples throughout the decade are presented. These examples served the purpose of informing of current practices in instrument development and validation and point to gaps in the research (see Table 1). None of these examples were included in the Jonsson and Svingby (2007) and Reddy and Andrade (2009) reviews.

The first example is Anderson's (1998). She developed the "4x4 method" for developing rubrics. She did not report on details of sampling, content, and/or validity and reliability which is a limitation, but reported on the model she developed and her experience with it. The 4x4 method in summary is a process inclusive of students where they, in small groups, identify four characteristics of good quality in the tasks assigned. In an iterative process, the class reaches consensus about these four criteria. Then, the students write four descriptors with a corresponding score. The scores were one to four, low to high. These descriptors are submitted for consensus as well. The end product is a rubric upon which students and instructor agreed. Anderson reported that hers was a good model for including students thoroughly and there were no high financial costs implied, but that it was time consuming. The instructor should be ready to provide plenty of examples and support. Anderson's 4x4 method has inspired other researchers like Stevens and Levi (2005).

Stevens and Levi (2005) developed rubrics in their Freshman Inquiry and Teacher Training workshops, separately. Their model encompassed developing rubrics in four stages: reflecting, listing, grouping and labeling for performance descriptions and application. The instructor carried out the reflection stage. This reflection stage entailed reflecting on students' expectations, the rationale for assignments, and his/her past experiences. Reflection *per se* is a highlight of Stevens and Levi's (2005) model, which was not found in other studies. Although instructor's "reflection" does not guarantee the automatic pursue of improvement backed up by research or strong theoretical frameworks, it could be used as a measure of integrity.

The listing stage can include professor and students, though it is principally done by the professor, where details on the assignments and learning objectives are outlined. The final stage is application, where the rubric takes final form. Stevens and Levi's (2005) work was pragmatic and creative in terms of pedagogical strategies, not cost-demanding, and students seemed to be genuine "stakeholders" (p. 49); however, the authors did not discuss any procedures for reliability and validity, which is a limitation.

Moreover, Stevens and Levi (2005) suggest involving colleagues such as the Writing Center staff, in the development of rubrics. Stevens and Levi consider their inclusion as a benefit, for it might improve communication among educators and tutors, and help the researchers delegate work in the process of developing the rubric. One limitation of Stevens and Levi (2009)'s study was that, unlike a study conducted by Allen and Knight (2009), they did not seek field experts' input on their instruments.

Allen and Knight (2009) conducted an empirical study in which they developed a method for collaboratively developing and validating a rubric, observably more elaborated than the other studies. Their resulting rubric was intended to be academically-sound but also employer relevant. In summary, Allen and Knight (2009) had an eight-step process. Step #1 was to develop learning objectives for their Public Relations course. The creation of these learning objectives was based on Public Relations literature, and once created, the learning objectives were subject to revisions from professionals in the field. Step #2 was to identify sample of work to link learning through the review professional sources, reports, and conference proceedings.

The third step in the Allen and Knight (2009) process was to develop the evaluation rubric with primary and secondary research data, professional input from four

experts in the field, and repeated discussions with faculty about samples of students' work. Allen and Knight also conducted numerous interviews with professionals and faculty members at that point. This part of the process was one of the strengths of this study because it entailed iteration in the input discussions.

The fourth step in the Allen and Knight (2009) study was titled "longitudinally test student learning as measured by the rubric." During this fourth step, the professional baseline was established, and refined through academic and professional input. The fifth step was to identify problems whenever there was sub-optimal performance. This step also entailed using longitudinal data to assess the rubric and related pedagogical strategies, and to evaluate uniformity when multiple sections of the Public Relations course were taught by different faculty members.

The sixth step of the Allen and Knight (2009) study was to "improve construct validity of the rubric." The researchers used the *Delphi* technique with professionals and then with faculty to distill knowledge and reach consensus on weights of each category of the rubric. The seventh step dealt with assessing the stability of the rubric's standards to determine the ability of the rubric to differentiate between submissions. Separately, faculty and professionals used the rubric to grade a weak and a strong students' assignments ($N=61$ assignments), and used a two-way ANOVA to test if significant difference existed. They also used F test for equal variances. The last step was to analyze data to improve reliability and further validate the rubric. The *Delphi* technique and the F test for equal variances were used again. Inter-rater reliability and rater repeatability and reproducibility processes were carried out.

Limitations of the Allen and Knight (2009) study were that they do not have student input and that there could have been financial implications which they did not report on. These financial implications could include the consulting fees of field experts. Also, there needs to be much willingness from several stakeholders and in some cases this is not available. The authors pointed out that “research is notably sparse on the collaborative process of developing and validating a rubric that integrates data collected from academic and professionals” (p. 2). No other study in this literature review included field professionals as thoroughly as Allen and Knight’s did. Their rigorous and systematic approach to reliability and validity using multiple processes were a highlight.

The last example of a rubric-development application included in this literature review was the Mansilla, Duraisingh, Wolfe, and Haynes (2009)’s study. Their empirically-tested instrument was developed through an iterative process that included a theoretical framework development, validation, and further revisions to their instrument, titled “Targeted Assessment Rubric for Interdisciplinary Writing.” This rubric was designed for the field of interdisciplinary writing among undergraduate college students.

Mansilla et al. (2009) used a framework developed and polished by the first author and other colleagues, since 1994. Since then, the framework had been subject to experts’ scrutiny using a variety of validation standards. This framework is called “Targeted Assessment Framework.” Mansilla et al.’s resulting rubric had a strong theoretical foundation, and its format was similar to all of the studies discussed above. This rubric differed from other rubrics included in this review in the labels used to describe the levels of students’ interdisciplinary understanding. Mansilla et al.’s used the following labels to indicate levels of attainment: *naïve*, novice, apprentice, and master.

These labels might not be the most adequate for the IU's rubric because they could be interpreted as being based on learners' deficiencies and not on efforts.

The labels that were used for the instrument developed in this dissertation were "beginning," "developing," "accomplished," and "exemplary." These labels were chosen because they are tactful, yet clear (Stevens & Levi, 2005), and indicated a process in learning and performance. These labels avoid the concept of failing, lacking ability, or other punitive labels, as the label "poor" would suggest, for example. Also, the IU rubric had four levels of attainment to avoid having scorers remain in a neutral position towards the students' work, as a five-level numeric scale would.

For the reliability estimates of the rubric, Mansilla, Duraisingh, Wolfe, and Haynes (2009) had a systematic analysis of students' essays, which served as sample work ($N=84$ essays). The essays became increasingly longer in average of pages: four for freshmen, 13 for sophomores, and 83 to seniors. For the reliability test, the first two authors scored 40 essays and then discrepancies were settled by consensus. No description is provided on the consensus strategies used.

The two second authors of the Mansilla et al.'s study scored the same 40 essays again for calibration of the rubric. After each round of scoring, the rubric was revised to reflect the agreements reached. Then, the first two authors scored independently the remaining 44 essays. Then, the inter-rater score reliability estimate was assessed by the number of scores or "judgments" on which all raters agreed divided by the total number of scores observed.

Their inter-rater score reliability of the Targeted Assessment Rubric for Interdisciplinary Writing was 83.5%. It was not explained why only 40 of 84 essays

were scored twice by different scorers, or how were these chosen. Having consistency in the procedure or providing an explanation of the decisions made throughout the Mansilla et al.'s study would have increased the study's trustworthiness regarding their inter-rater score reliability estimates.

The validity estimate of the Mansilla et al. (2009) rubric was done with a series of 4x1 ANOVAs. Each ANOVA corresponded to each dimension of the rubric. The authors hypothesized that seniors would score significantly higher than the freshmen and sophomores on the four categories of the rubric, and they did. To test for lengths of the essays and addressing the issue of number of pages as a possible confounding variable, they conducted *post hoc* tests. The *post hoc* tests were four ANCOVA tests, using the number of pages in the essays as the covariate and the student category as the independent variable.

The results of the Mansilla et al. (2009) study were that for the four tests, student category significantly predicted scores and essay length did not. The developmental differences between freshmen and seniors were a function of educational experience and not the length of their essays. The authors concluded that their Targeted Assessment Rubric for Interdisciplinary Writing had good estimates of inter-rater reliability to score essays with a wide array of topics, and also good estimates of validity, because it captured developmental differences among students regardless of essay length. The Mansilla et al.'s study strengths were the thoroughly-validated framework used by the researchers to build their rubric's categories, the systematic validation testing, the large number of essays scored, and the implied collaboration among authors.

The possible weaknesses of the Mansilla et al.'s study were the lack of explanations of procedure for their reliability estimates, the lack of input from students and of experts other than the authors. The procedure to develop the rubric at the IUs did not have a large sample of students' work as the Mansilla et al.'s study, but it attempted to include students and experts, and to provide detailed narratives of the validity and reliability estimate procedures.

In conclusion, the review of these rubric-development studies suggests that including students and experts, other than the authors, was a common practice, although their inclusion had varying degrees. It is also suggested that assuring consistency in the rigor and reporting on the reliability and validity estimates is important. This review shows that the inter-rater score reliability estimates are commonly carried out, but not intra-rater score reliability. Intra-rater score reliability addresses changes in the scoring results of a given rater over time (Moskal & Leydens, 2000; Newman et al., 2006).

Furthermore, no study was found where higher education students were included in the making of the instrument to the maximum level possible. Also, no studies were found where Lawshe's CVR was used to estimate content validity, and no argument was found discouraging its use.

Table 1

Comparison of the Examples of Practices or Studies Regarding Collaborative Development of Rubrics in Higher Education

Study/Model	Inclusion of Students?	Reliability Estimates	Validity Estimates	Statistical Tests	Strengths	Weaknesses
Anderson (1998)	Marginal - only in defining quality levels and scoring levels	Not discussed	Not discussed	None	Inclusive and simple, and probably cost-effective	Time consuming if students do not have previous experiences, and narrow in scope (no experts included); no rigorous procedures at any time
Stevens & Levi (2005)	Marginal- only in defining quality levels and scoring levels	Not discussed	Not discussed	None	Inclusive of students, simple, creative, reflective, flexible, and probably cost-effective	Narrow in scope (no experts included); no rigorous procedures at any time

Study/Model	Inclusion of Students?	Reliability Estimates	Validity Estimates	Statistical Tests	Strengths	Weaknesses
Allen & Knight (2009)	Marginal to none	Inter-rater reliability, rater repeatability and reproducibility	Construct Validity (<i>Delphi</i> technique), Expert Judges (faculty and outside professionals)	Two-way ANOVA, <i>F</i> Test	Thorough iterative processes for content validity (qualitative and quantitative); inclusion of professionals in the field	One might need to have access to willing professionals, and there might be financial costs involved
Mansilla, Duraisingh, Wolfe, & Haynes (2009)	None	Inter-rater reliability	Scores' reliability across students' categories (freshmen, sophomores and seniors)	ANOVA, ANCOVA, post-hoc tests for length of students' essays	Thorough testing of validity; strong theoretical framework to develop rubric; large sample of students' essay	The authors did not include other experts than themselves; they did not have any student input; the reliability procedure lacked a thorough description

Two Approaches to Instrument Development

This section was devoted to two approaches for instrument development: the Newman and McNeil's (1998) and Onwuegbuzie, Bustamante, and Nelson's (2010). The Newman and McNeil (1998) approach was not as current as the second one, but it is written in a question and answer format that is very comprehensive. The authors presented general guidelines to follow sequential steps in instrument development. This simple yet thorough style allowed for this approach to be easily adapted for rubrics.

Although it is not stated that it is for mixed methodology, the nature of the Newman and McNeil (1998) approach lent itself to a research question-led process and for the use of multiple sources of data. For example, the first stages highlighted the crucial role that "stakeholders" hold in planning for and creating the instrument, which was a qualitative-type of decision. Consecutive stages included quantitative instrument psychometric considerations.

Although the instrument-development process presented in this dissertation was fully developed by the author/researcher, the Newman and McNeil approach was explored and included only for its adequacy to keep integrity measures, or a check-off list to make sure that the minimum components of instrument development research were present. The researcher answered each question and reported the answers as a checklist when planning, creating, and piloting her instrument (Appendix C). The stages and questions of the Newman and McNeil checklist were:

Stage 1: Specify the intent of the instrument

Question 1: What are the basic questions I want my instrument to answer?

Question 2: How do I plan to use the information?

Question 3: Who do I plan to collaborate with to obtain those answers?

Stage 2: Assess the Available Resources

Question 4: What assistance do I need in creating this instrument?

Question 5: What types of analyses are required?

Question 6: Are computer programs available?

Stage 3: Define the population

Question 7: What is the population to which I am interested in generalizing?

Question 8: What are the characteristics of the population to which I am interested in generalizing?

Question 9: How does the accessible population differ from the population which I want to generalize?

Question 10: How would volunteering affect my results?

Stage 4: Review of the Pertinent Literature

Question 11: Where can I look for information?

Question 12: What are the key terms for my literature search, for the section of instrument making?

Question 13: What are the best sources of information?

Stage 5: Determine the data collection techniques

Question 14: What instrument procedure should I use?

Question 15: Should I use an existing rubric or develop my own?

Question 16: Who should administer the rubric?

Stage 6: Develop the instrument

Question 17: What are the item format considerations?

Question 18: What are the psychometric considerations?

Question 19: What are the piloting considerations?

Question 20: What are the training considerations?

Stage 7: Determine the sampling procedure

Question 21: How should I sample?

Stage 8: Analyze results and prepare a report

Question 22: What analyses should I perform?

Question 23: How should I write my report?

The strengths of the Newman and McNeil approach were its clarity, consistency, and organization, and that the format itself could be used as a measure of integrity in the form of a self-debriefing list. The date of publication could be a disadvantage. However, an advantage is that the questions did not lead to a particular set of practices; rather, they allowed for the researcher to survey different sources and make educated decisions regarding instrument making.

The second approach considered for this review is the Onwuegbuzie, Bustamante and Nelson's (2010). In contrast to the first approach, this one was specifically designed for mixed methodology. The Instrument Development and Construct Validation (IDCV) model was developed by Onwuegbuzie, Bustamante, and Nelson (2010) to optimize the creation and validation of quantitative instruments using mixed methods research.

The IDCV is a research-based meta-framework, "designed to help instrument developers undergo a rigorous and comprehensive process during instrument development/construct validation" (p. 60). The use of the IDCV comprises different frameworks and models: (a) Onwuegbuzie and Combs' (2010) crossover analyses, (b)

Green, Caracelli, and Graham's (1989) rationale for combining qualitative and quantitative data, (c) Onwuegbuzie, Daniel, and Collins' (2009) meta-validation framework, and (d) Onwuegbuzie, Leech, and Collins' (2008) framework for debriefing the researcher. The IDCV included ten elaborated stages and numerous procedures for validity and reliability estimates. The stages are:

1. Conceptualize the construct of interest
2. Identify and describe behaviors that underlie the construct
3. Develop initial instrument
4. Pilot-test initial instrument
5. Design and field-test revised instrument
6. Validate revised instrument: Quantitative analysis phase
7. Validate revised instrument: Qualitative analysis phase
8. Validate revised instrument: Mixed analysis phase: Qualitative-dominant crossover analyses
9. Validate revised instrument: Mixed analysis phase: Quantitative-dominant crossover analyses
10. Evaluate the instrument development/construct evaluation process and product

The IDCV approach also suggested a thorough review of the literature and consultation with local experts and key informants. The authors used the IDCV to develop the School-Wide Cultural Competence Observation Checklist, which served as a heuristic example of the use of the IDCV.

The strengths of the IDCV approach were the thorough and sophisticated use of meta-frameworks and the iterative nature characteristic of mixed methods, in addition to

being published more recently. However, the IDCV might not be the most appropriate model for this study because it might require advanced instrument-development skills or a diverse team of researchers. The researcher might consider using the IDCV in further research endeavors. The Newman and McNeil checklist was used as a check-off list and as a transparency measure for this study, instead.

The Human Capabilities Approach and Education

The human capabilities approach (HCA) was introduced in Chapter 1. The human capabilities approach is a fairly recent approach to human well-being, but it has gained attention worldwide in the last two decades. Nevertheless, the research on performance-based assessment and the HCA was scant at the time of this review. Flores's and Walker's work were the two studies found. This section is devoted to a more detailed description of the HCA, an HCA-based literature review (Flores, 2007), and a discussion about an educational-assessment study that uses the HCA as a framework (Walker, 2008).

The human capabilities approach is, in Sen's words, "a concentration on freedom to achieve in general and the capabilities to function in particular, and the core concepts of this approach are 'functionings' and 'capabilities'" (Sen, 1995, p. 266). Freedom relates to "the range of options a person has in deciding what kind of life to lead" (Drèze & Sen, 1995, p. 10), functionings are achievements, "whereas a capability is the ability to achieve" (Sen, 1987, p. 36).

The concepts of capabilities, freedoms, and functionings, are related. In Sen's (1992) words: "Capability is primarily a reflection of the freedom to achieve valuable functionings" (p. 49). Well-being, according to Sen should be evaluated based on

capabilities, which pertain to what people are able to do or achieve, going beyond traditional well-being approaches (Sen, 1990; 1999). As suggested by Sen (1999), every individual and community should be able to develop his/her/their own capabilities engaging in democratic dialogue, in a participative process.

Education offers fertile soil for adopting and further developing the human capabilities approach, essentially in the inclusion of key elements such as democratic dialogue (Landorf and Pineda; 2010). Sen (1999) regards democratic dialogue as important on three levels: directly, instrumentally, and constructively. Educators have the responsibility to model and facilitate democratic dialogue in order for students to learn its direct importance: the fact that freedom of choice has value in and of itself, regardless of results.

In pedagogical practices centered on democratic dialogue, the educator becomes attuned to her students' needs in order to facilitate meaningful discussion that allows students to understand their capabilities and what they must know and do in order to become whom they want to be. By engaging in democratic dialogue, the educator and her students arrive to its instrumental importance: that liberty of thought and action can lead to secondary positive results.

The constructive importance of democratic dialogue must be recognized as vital to education. Principles of this constructive importance are freedom as an educative process that plays a critical role in public discourse, and the community's negotiation for consensus on values, goals, and priorities (Drèze & Sen, 1995; Sen, 1999). As Drèze and Sen (1995) argued, democracy goes beyond the establishment of institutions and formalities so "the majority" is able to rule; it includes the guarantee of free discussions

and “the widespread actual participation of people” (p. 24). This is one of the reasons why using the HCA as conceptual framework for this is prudent for this research study, especially when it comes to educational assessment. Two studies found in this literature review serve as examples. These were Walker’s (2008) and Flores’s (2007).

Walker’s (2008) study had the purpose of proposing a human capabilities approach for evaluating students learning and the social and pedagogical arrangements that support equality in capabilities for all students. Using what she called “a grounded method of student voice” (Walker, 2008, p. 480), she conducted interviews among junior students in a university in South Africa.

Walker’s (2008) participants were enrolled in two different undergraduate classes, Southern Africa History and Town Planning. Through her work, Walker (2008) aimed to “illustrate how education might foster capabilities which are valued by students” (p. 480). She conducted individual interviews and focus groups and then “teased out” (p. 480) what was of value and importance for them in their learning in well-being themes. Using this approach, Walker was able to offer a different lens from the traditional measurable learning achievements approaches.

From the students’ input, Walker (2008) extrapolated “valued opportunities and achievements that suggest well-being impacts or changes in the capability sets of the students” (p. 481). The functional capabilities that the students identified were: knowledge, social relations, critical thinking, imagination and empathy, recognition and respect, active and experiential learning, autonomy, confidence, active citizenship, deliberative dialogues, and having economic opportunities. Each capability has its local, situated and “thick” discussion.

Walker considered her research to be an initial step at applying the capability approach to local learning activities, and that using human capabilities approach provided a dimension for assessing non-measurable valued and relevant achievements for students. Walker's work resonates with this dissertation. As shown in Walker's research, the HCA allows universities to consider beyond the traditional higher education aims and graduate attributes/outcomes, and possible neoliberal interpretations. Furthermore, Walker's work illustrates how the HCA allows students to identify and have deliberative and democratic dialogues on the kinds of capabilities they value.

The second article presented in this section is Flores' (2007), is a literature review and a discussion on philosophical foundations of the HCA and the link with indigenous students in Mexico. The purpose of Flores' article was to examine Sen's notion of identity in the contexts of the social and educational disadvantages that indigenous people face in Mexico. Flores (2007) suggests that a person's identity is a crucial aspect to articulating and understand research on education. Identity might influence school choice, career preferences, and classroom behavior, and even shape and modify values and beliefs. For Sen, pioneer of the HCA, identity can be chosen instead of just discovered, and there are facets of it. These facets are context-dependent for their relevance.

Pedagogical practices and institutional arrangements play a crucial role in allowing students to choose and develop their identity, and to some indigenous students, self-identifying as indigenous is a source of pride. In Mexico, however, not all indigenous people have had the "instrumental freedoms" to lead the lives they value, as Flores (2007) pointed out. Flores' article was informative for this study in strengthening

the understanding of the relevancy of the HCA to this research project. Flores' work allowed the researcher to identify the convergence between the IUs educational model, the sociopolitical context in which they exist, and the HCA.

Summary

This chapter presented a review of the literature relevant to this study. It was divided into four sections: (a) the IUs in Mexico, (b) research on assessment of competencies at the IUs in Mexico, (c) creating rubrics collaboratively in higher education, and (d) the human capabilities approach to education. The first section provided more background on the intercultural universities in Mexico. The second section presented the only research found on assessment of competencies at the intercultural universities of Mexico, by Caro. Caro's (2010) study allowed the researcher to have a point of reference from another IU. The findings from Caro's study were that the process used to assess students' artifacts was not consistent, and some students and tutors complained about it, but that its performance-based nature should remain. Caro's findings highlighted the need for a participatory, performance-based instrument with high validity and reliability estimates.

The third section analyzed several studies where researchers developed rubrics collaboratively. Some instrument-development practices were confirmed to be adequate, such as the use of inter-rater and intra-rater score reliability estimates, and the inclusion of experts. No other study was found to have the extensive inclusion of students and the use of Lawshe's CVR. This study addressed these gaps. The fourth section presented two articles on the human capabilities approach and higher education- one was a

qualitative study by Walker (2008), and the other one was a literature review by Flores (2007).

CHAPTER III

METHODS

This chapter opens with the purpose of the study and the research questions guiding it. The research design is described, followed by sampling strategies and data collection, sources, analysis, and management procedures. The chapter concludes with a description of integrity measures and limitations of the study.

Purpose of the Study

After an extensive review of the literature on performance-based assessment (PBA), an analysis of the IU's policy book, the findings of a study conducted at one of Mexico's IUs (see Caro, 2010), and personal communication with SEP and IU educators, the need for developing a PBA instrument was established. This study aimed to develop a process to create a performance-based assessment instrument with acceptable standards of validity and reliability estimates. The instrument was designed to measure the attainment of competencies outlined in the Intercultural Management for Development undergraduate program at one of Mexico's 10 IUs.

There were three central premises guiding this study. The first central premise of this study was that performance-based assessment is the desirable approach to assess student learning in institutions such as the IUs. This idea is supported by the Ministry of Education's official IU policy book, a research project conducted at one of the IUs (see Caro, 2010), and an extensive literature review. The IUs already require the completion of performance-based assessment products that students submit as a graduation requirement. These products are action-research projects called *documentos recepcionales* (DR). Caro's study suggested that, at her IU, some of her participants

expressed that the process to assess DRs or the attainment of competencies could be fine-tuned, and there was no rubric or other assessment instrument used throughout the process. These ideas led to the second central premise of this study.

The second central premise of this study was that rubrics are the preferred performance-based assessment instruments. This idea was also supported by an extensive literature review. It was confirmed through personal communication with assessment officials in the Ministry of Education and with numerous IU contacts that such instrument had not been created yet.

The third central premise of this study stemmed from the theoretical framework guiding this study, the human capabilities approach (HCA). As argued in Chapter 1, the HCA mirrors the IUs' mission and vision. In particular, the work of Walker (2008) of HCA and assessment in higher education was seminal for guiding and interpreting data from this study. In brief, HCA focuses on human development, giving individuals freedom to achieve, but also the potential to achieve.

In HCA terminology, the freedom to achieve and the potential to achieve are functionings and capabilities, respectively (Sen, 1992, 1999). Developing functionings and capabilities is undertaken through democratic dialogue, through constant negotiation of what kind of life is valued, and through people's empowerment (Drèze & Sen, 1995; Sen, 1992, 1999). In education, HCA fosters students' exercise of freedom over their valued functionings and capabilities (Nussbaum, 1997; Unterhalter, Vaughan & Walker, 2007; Walker, 2008). The articulation of this freedom in education translates into going beyond instrumental values in education; thinking beyond the fact of securing a job upon completion of one's studies (Walker, 2008).

The third central premise of this study, based on the HCA, was that students should guide the development of their rubric. This was achieved during two focus groups, as students negotiated and stated what evidence demonstrates the attainment of competencies of their program of study, Intercultural Management for Development. The resulting rubric should be subject to quality measures just as any high-quality assessment instrument should be (Messick, 1994). These measures are content validity and inter- and intra-rater score reliability (Moskal & Leydens, 2000; Onwuegbuzie & Daniel, 2002, 2004; Thompson & Vacha-Haase, 2000; Vacha-Haase, Kogan, & Thompson, 2000; Witta & Daniel, 1998).

Research Questions

This study was guided by two research questions:

- (a) According to IU seniors, what evidence demonstrates the attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?
- (b) Can acceptable standards of validity and reliability estimates be established for an instrument developed to measure students' attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?

Research Design

This study used sequential mixed methods, commonly used for developing and/or testing instruments, and especially advantageous for researchers creating a new instrument (Creswell, 2003). Sequential mixed methods is typically conducted in, but not limited to, two phases (Creswell, 2003). The sequence is comprised of qualitative and quantitative phases that happen in chronological order (Teddlie & Tashakkori, 2009).

Moreover, “[t]he research questions... depend on the previous strand” (p. 143). The first research question of this study established the evidence of attainment of competencies, organized in the form of a rubric. This evidence served as the basis for responding to the second research question, which was related to the instrument’s content validity and inter- and intra-rater score reliability estimates.

In sequential mixed methods, all phases are interrelated but may evolve in the course of the study (Teddlie & Tashakkori, 2009). In this study, for example, there were qualitative data collected through a focus group conducted in the early stages of the study; these data became the content for creating an initial draft of the rubric. Then, the first draft of the rubric was used to create a Table of Specifications (ToS), which is a mixed methods tool that yields quantitative and qualitative data (Newman et al., 2011a). The ToS allows the researcher to organize experts’ consensus on items of the instrument subject to score validation estimates in two ways. The first way that the ToS allows the researcher to organize consensus is numerically, through frequency count and percentage of agreement calculation. The second way is through obtaining oral/written feedback from participants. This feedback can then be used as qualitative data (Newman et al., 2011a). Using the ToS, quantitative and qualitative data were obtained in the form of content validity estimates and expert judge consensus or agreement.

For instrument development using sequential mixed methods, Creswell (1999, summarizing Caracelli & Greene, 1993, and Tashakkori & Teddlie, 1998), suggests to first “obtain themes and specific statements from participants in an initial qualitative data collection [phase]. In the next phase, use these statements as specific items and themes

for scales to create... [an] instrument that is grounded in the views of the participants. A third, final phase might be to validate the instrument..." (p. 221).

The following diagram (Figure 1) illustrates the sequence used in this study. Although the two main stages are qualitative and quantitative respectively, during the analysis, both types of data were obtained to inform consecutive steps within the stages. For example, during the second focus group, quantitative data in the form of content validity estimates were obtained, and qualitative data were also gathered as the discussion among participants took place. Another example is the collection of feedback from scorers during the scoring sessions. In some mixed methods research studies, integration or *mixing* of data occurs at several points during the process of the research study (Creswell, 2003) with the purpose of better understanding both types of data. The data collection procedures mentioned in the diagram, which are the focus groups, the Table of Specifications, the content validity and reliability estimates, are explained in detail in the Data Collection segment, available after the sampling section.

Figure 1. Sequential Mixed Methods utilized for the Creation of a PBA Instrument or Rubric.
 *During the researcher’s visit to conduct the first focus group, there was a recent graduate who expressed interest in participating, and was welcomed.

Sampling

Participants were selected using purposive sampling strategies. Purposive sampling allows the researcher to select participants based on specific purposes associated with answering the research question (Teddle & Tashakkori, 2009). Using

purposive sampling, the researcher can select individuals that meet a predetermined set of characteristics (Patton, 2002).

There were three sets of participants in this study. The first set was comprised of six IU senior students and a recent graduate who participated in the first focus group. The second set was comprised of five IU seniors who participated in a second focus group. The third set of participants was comprised of Spanish-speaking education professors in Mexico and the United States.

Three of the senior students that participated in this study were also participating in another study with another researcher at their university. Through personal communication, an IU researcher, Dr. Diaz², expressed support for this dissertation study, and assisted the researcher to establish the first contact with participants.

All participants were invited to be part of this study on a voluntary basis. For this first and second set of participants, the criteria to be part of this study were to be enrolled at or be a graduate from the IU, and not having transferred from another institution at any point. In the original dissertation proposal, it was projected that only IU graduates would participate in the focus groups, under the assumption that they had already completed the DR defense process. Contacting at least ten graduates to participate in person in this study was extremely challenging. However, senior students, who were approaching defense date soon, were very capable of expressing their views, as explained in the next section.

The second set of participants of this study was another group of IU seniors, who served as expert judges to conduct content validity estimates on the instrument developed

² Pseudonym

by the first group. For this second set of participants, the criteria to be part of this study were the same as for the first set.

A third group of participants was composed of seven professors of education in Mexico and the United States, contacted personally by the researcher. The criteria for selection were to hold graduate degrees in education, and to be a native speaker of or highly proficient in Spanish. These professors served as scorers, estimating the developed instrument's inter- and intra-judge score reliability, using IU alumni's DRs.

Data Collection

This section includes data collection procedures, introduced briefly in the Research Design section. Data collection entailed conducting two focus groups, and obtaining validity and reliability estimates of the instrument created.

The Focus Group

A focus group serves the purpose of data gathering through in-depth discussion and also observations during the session (Morgan & Spanish, 1984, as cited in Teddlie & Tashakkori, 2009).

Focus Group Administration

This study followed Krueger and Casey's (2000, as cited in Teddlie & Tashakkori, 2009) description of focus groups, who suggest that a focus group consist of five to ten participants, and that the group be homogenous. This group was homogeneous because all participants attended the IU. With the exception of an IU alumnus, focus group participants were in the process of completing their DRs.

Moreover, Krueger and Casey (2000, as cited in Teddlie & Tashakkori, 2009) suggest that there should be a moderator often accompanied by an assistant and that the

meeting last no more than two hours. During the first focus group, the researcher asked a colleague to take notes, and to serve as moderator if necessary. Unfortunately, her colleague declined the invitation a few minutes before the focus group took place. For the second focus group, the researcher had a colleague take notes and provide feedback upon completion of the meeting. Neither of the focus groups took longer than two hours. With permission of the participants, the focus groups were audio-recorded and the researcher managed the recorder. Only pseudonyms were used at all times. Participants were contacted through an IU researcher, in person, and/or via email.

Focus Group One

Six seniors and a recent graduate participated in the first focus group. Five participants were women, and two were men. Participants were labeled F1 through F5 for women, and M1 and M2 for men. The four campuses of the IU were represented. The research director of the IU helped the researcher contact three of the seniors via email, and then in February, 2012, she was introduced to three more seniors and a recent graduate during a conference that took place in their state's capital city. She invited the six seniors and the graduate to participate in her study, and all of them responded affirmatively.

The first focus group took place on February 17, 2012 in Mexico, during a week-long conference that several IU students attended in the capital city of their state. Dr. Diaz suggested that the researcher attend this conference to meet with the IU students. The focus group lasted an hour and thirty minutes. The purpose of this focus group was to collectively build the first draft of the IMD rubric.

In the dissertation proposal, it was projected that the sample would be comprised of IU graduates. The rationale was that they would be able to provide more informed feedback during the focus groups regarding the different levels of attainment of the IMD competencies. Having completed a bachelor's degree from the IU and a DR were considered unique factors for the participants' ability to dissect each competency and to provide feedback on their IU educational assessment experience.

Recruiting graduates for the first focus group was extremely challenging. Once students graduate from this IU, many have jobs in remote rural areas and their ties to their university are sporadic to none. The researcher obtained a directory of more than thirty recent graduates. Two graduates replied via email but their schedule and the researcher's did not allow them to meet during the first focus group. Only one graduate attended the first focus group when invited personally by the researcher.

The educational model of the IU is that students spend some days per month in the community working on their DRs and meet several times a semester with their tutors who orient them to the process of completing these (Casillas & Santini, 2006). By their last terms, these six participating seniors had had several semesters of practice and interaction with the creation process of their thesis.

The first focus group was guided by a set of questions previously reviewed by various "dissertation boot camp" peers who were all candidates or students in doctoral programs at the College of Education at FIU, the director of the IU, and approved by the researcher's dissertation committee. Each focus group participant received a copy of the focus group questions (Appendix A), the list of the IMD competencies obtained from the IU website, a rubric template (Appendix D), and an informed consent form (Appendix F)

that participants signed before the focus group session began. All the material provided to participants was in Spanish.

The researcher explained in detail the purpose of her study and guided students in the use of the rubric template (Appendix D) to facilitate the creation of the actual rubric and allowed the time to be used as efficiently as possible. She explained that the template would help participants organize their responses to the researcher's questions. For example, "What does attaining competency 1 look like, for a student that is performing at an Exemplary level?" During the discussions, the researcher explained that rubrics serve as anchors of understanding or communication tools when educators and students engage in performance-based assessments. Having the template of a rubric facilitated the creation of the actual rubric and allowed the time to be used as efficiently as possible.

During the focus group, participants discussed what evidence is appropriate to indicate each level of attainment for each of the competencies. As participants were discussing the levels of attainment, the researcher was taking notes on a copy of the rubric as well. The researcher transcribed the focus group recording, incorporated the students' discussions on each level of attainment, and confirmed on-site that her compilation of notes were accurate. She read back to the participants each of the sections of the rubric, checking the degree of consensus among participants regarding what she wrote. For example, the researcher said: "Let me read to you what I gathered from everyone's descriptions" and asked "Do you agree with what I wrote?"

A month after the focus group took place, for member-checking measures, participants received from the researcher a clean draft of the rubric and a copy of the

focus groups' transcripts via email. Participants were asked to check the documents for accuracy and provide any changes necessary. Five out of the seven participants replied confirming that these documents were accurate and that they agreed to the content. The other two participants did not reply to the researcher's email, even though she sent two weekly follow-up reminders. No further feedback or revisions were given to the researcher. Appendix H includes each of their rubric drafts and the researcher's compilation rubric.

Focus Group Two

The second focus group took place on March 14, 2012 at a meeting room in one of the IU campuses, in Mexico. It was comprised of five seniors from the same campus. Three of them were women and two were men. Participants were labeled as FA through FC for women, and MA and MB for men. Their participation was comparable to the first group's seniors. The procedures of the second focus group were almost the same as to the first focus group, except that this meeting had the main purpose of obtaining content validity estimates on the rubric, asking this second group of participants to take the role of experts.

During this second focus group, the experts received a copy of the guiding questions of the focus group (Appendix B). These questions were previously reviewed by various "dissertation boot camp" peers, all doctoral candidates/students at the College of Education at FIU, the director of the IU, and approved by the researcher's dissertation committee. Participants also received a copy of the rubric, which had already been submitted to and revised by the first group of seniors and the graduate for member-checking purposes. The researcher explained the purpose of her study, what the first

group of participants did, the concept and purpose of rubrics, the concept of content validity, the students' roles as expert judges, and the purpose of and procedures to complete the ToS. The ToS, which is a mixed methods tool to estimate content validity (Newman, Lim, & Pineda, 2011) is described later in this chapter, in the Content Validity section. All material provided to participants was in Spanish.

Similar to the first focus group, the researcher compiled notes on the students' discussions on each of the attainment levels, their consensus on the accuracy on the original descriptions, and any revision they agreed upon. For example, she said "you correct me if I'm wrong, OK? I take notes on what I hear you say," and after the attainment descriptions were revised, she would ask "Five out of five? All in agreement or some in disagreement?" The researcher's compilation notes during the focus group are available in Appendix I. For member-checking measures, participants received via email a clean draft of the rubric reflecting all the ToS feedback, along with a copy of their focus group's transcripts, transcribed by the researcher. Participants were asked to check the documents for accuracy and provide any changes necessary. Four out of the five participants replied and confirmed that the documents were accurate and that they agreed to the content. The fifth participant did not reply to the researcher's email despite her two weekly follow-up reminders. No further feedback was given to the researcher. Appendix J is the final version of the rubric.

Observation Session before the Focus Group Two

During her visit to one of the campuses of the IU, the researcher was given the opportunity to attend a senior students' meeting. This observation session was not planned, but the researcher considered it a good opportunity to better contextualize her

study. During this seniors' meeting, students presented their DR preliminary findings, or products almost completed, either individually or in pairs. All of the presentations the researcher observed were done using Power Point presentations.

During this observation session, the researcher noted that the feedback given to participants ranged from being based on the visual appeal of the Power Point Presentation, to some methodological aspects such as the strengthening of the use of their FODA (*Fortaleza's, Debilitates, Oportunidades and Amenazas*) strategy. Nevertheless, feedback on students' DRs was not structured or unified, and it was given only by the two advisors. The other students in the room did not give feedback to each other during the session.

The content of the DRs' presentations observed was of remarkable quality and sophistication. For example, one was a manual to manage a deer farm. The students justified the need for the manual, presenting the current feeding practices, which are wasteful and do not nourish the deer well. The manual also addressed necessary procedures to improve the infrastructure of the deer farm, which presently allows for predators to easily get in and eat the deer, or for small deer to escape. The students presented their management manual, proposing cost-efficient feeding practices and fencing alternatives, as well as community revenue projections of a well-managed deer farm.

Through these observations, the researcher inferred that these seniors were going to be able to share their experiences about the process of their almost finished *documento recepcional*. During this observation session, the researcher also considered the possible

usefulness of a rubric, made by students, and distributed among their team members and their mentors.

The Performance-based Assessment Instrument

According to Stevens and Levi (2005), “a rubric is a scoring tool that lays out specific expectations for an assignment” (p. 3). Rubrics include a rating scale with systematic guidelines for assessing performance on tasks, products as evidence of learning or attainment of goals, or even answers to open-ended questions (Teddle & Tashakkori, 2009). When constructing a rubric, the different components of an assignment should be presented with detail, along with a thorough description “of what constitutes acceptable or unacceptable levels of performance for each [component]” (Stevens & Levi, 2005, p. 3).

Rubrics are categorized as either holistic or analytic (Zimmaro, 2004; Jonsson & Svingby, 2007). A holistic rubric provides “a single score based on an overall impression of a student’s performance on a task” (Zimmaro, 2004, p. 2). Students’ work is assessed as a whole. On the other hand, an analytic rubric allows scorers to “provide specific feedback along several dimensions” (p. 2). An analytic rubric should be used when relative strengths and weaknesses should and can be compared, when detailed feedback is needed, when dealing with complicated skills or performance, and/or when students themselves want to self-assess their understanding or performance (Zimmaro, 2004). Moreover, rubrics can be either general or task-specific (Zimmaro, 2004). General rubrics contain task descriptions that are universal across tasks. Task-specific rubrics provide a more detailed performance description for every assignment (Zimmaro, 2004).

The rubric resulting from this study is analytic and general. This rubric is analytic because scorers provide a score per competency, comparing each level of attainment described in the rubric with the students' work. Because scorers were not directly involved in the actual process of developing the DRs with the students, their feedback was not detailed, but an analytic rubric has that potential. This rubric is general because the descriptions of the competencies and the levels of attainment are universal across other tasks, different from the completion of their DRs. In this study, the "task" for students was to complete their DRs.

Validity and Reliability Estimates

Validity Estimates

Especially in mixed methods research, validity estimates or trustworthiness measures should provide the framework for the transparent construction of principles of practice. These processes "increase the likelihood that mixed methods designs are scientifically sound" (Ridenour & Newman, 2005, p. 2). The study is more likely to be organized and systematic, sequential, and the processes and findings can be verified and probably replicated (Ridenour & Newman, 2005).

The rubric developed in this study had estimates of content validity with expert judges, which is a sophisticated variation of face validity (Newman et al., 2006). As opposed to face validity, content validity with expert judges assumes expertise when subjects judge the instrument (Newman et al., 2006). In quantitative terms, this is "judgmental validation," and in qualitative terms, this is like "peer debriefing" (Teddlie & Tashakkori, 2009, p. 210).

Content validity with expert judges was estimated using qualitative data from the first focus group, formatted into a rubric, and organized into a Table of Specifications (Appendix E). In this case, the content was the competencies, or outcomes, and the items were types of evidence, or behaviors, dispositions, that the IU students and the recent graduate during the first focus group outlined. Because the outcomes or competencies outlined in the IMD program had already been established by the university, no experts' consensus was needed there. However, it was necessary to calculate participants' agreement on the evidence that is appropriate for demonstrating the attainment of the IMD competencies.

The Table of Specifications to Estimate Content Validity

A Table of Specifications (ToS) is “operationally defined by a set of procedures that attempt to align a set of items, tasks, or evidence with a set of concepts that are to be assessed” (Newman et al., 2011b, p. 5). The ToS is a mixed methods tool because it allows the researcher to gather quantitative data, as the researcher calculates percentages of agreement on the appropriateness of each item of the instrument being validated, and gather qualitative data as participants express their feedback and negotiate, orally, their consensus (Newman et al., 2011b).

The ToS developed for this study includes 20 items to rate, which are four sets of attainable skills per each of the five competencies listed in the IMD program of study. The experts, the IU seniors, indicated whether they agreed with the adequacy of the description of each item, and then a percentage was calculated on their agreement. For example, if four of the five participants agree on the adequacy of the description of an

item, this represents 80% of agreement among participants. This is one of the strategies that Newman and McNeil (1998) suggest to calculate consensus.

According to Newman et al., (2011b), an item on an instrument should be removed if the item reaches lower than 80% of expert agreement, which is approximately what is required for estimates of content validity. Alternatively, experts could add or replace descriptions, which will also be subject to the same score validation procedure. The ToS used in this study can be found in Appendix E.

Lawshe's Content Validity Ratio

In addition to percentage of agreement among experts, for triangulation purposes, discussed later in this chapter, content validity was estimated using Lawshe (1975) content validity ratio (CVR) to add rigor to the validity procedures. During this study's review of the literature, no study was found discouraging the use of the Lawshe's CVR. Lawshe's CVR is a useful tool for assuring that each item on an instrument proposed for inclusion is judged relevant to the content, by a group of expert judges (Guion, 1977). Lawshe's CVR is used to quantify consensus among experts regarding the content of an instrument (Lawshe, 1975; Guion, 1977). The formula to calculate Lawshe's CVR is as follows:

$$CVR = \frac{n_e - N/2}{N/2}$$

n_e = the number of experts who agreed on the relevance of the item, behavior, or question

N = total members of the panel of expert judges

Total agreement will equal 1, the highest CVR value. Refer to Appendix F for the minimum CVR values and their significance for one-tailed tests.

Reliability Estimates

The reliability estimates included in this study were inter-rater score reliability and intra-rater score reliability. These are operationally defined as the percentage of score agreement between raters, or inter-rater score reliability, and agreement within each rater's score over a period of time, or intra-rater score reliability (Moskal & Leydens, 2000; Newman et al., 2006). Inter-rater and intra-rater score agreement of at least 80%, is sufficient according to Newman et al. (2011b) to meet the minimum reliability requirement.

The Scorers

There were seven scorers that used the rubric produced in this study. Each scorer was assigned a scorer identification number. These identification numbers were S1 through S7. Four scorers are Mexican citizens. Three of the Mexican scorers are currently at Mexican universities and one at an institution in the United States. Two of the Mexican scorers are researchers at one of the IUs, different from the one where this study took place. The performance-based assessment requirements and processes at these scorers' IU are similar, so they were able to contextualize the production and assessment of DRs with ease. The four non-Mexican scorers were either born or have heritage from Latin America. Scorers were all contacted personally via email by the researcher. The scorers' invitation letter(s) and the instructions given to them in English and Spanish can be found in Appendix K; summaries of scorers' experience and expertise are available in Appendix L.

Scorers were from different institutions and nationalities. The researcher considered that having diverse institutional view points throughout the scoring exercises

would be enriching for her discussion. The researcher also considered that having international scorers would give a global, transnational, intercultural, dimension to her study.

Inter-rater Score Reliability

To obtain inter-rater score reliability estimates, the seven scorers received each the same number of DRs to score during two sessions, with at least two weeks in between scoring sessions. If the scores differed by one level of attainment, that is, one point apart, the score was considered acceptable, or the equivalent to 100% agreement. If the scores differed by more than one level of attainment, the DRs were evaluated by a third scorer. This score discrepancies resolution is a type of parity model (Johnson, Penny, Fisher, & Kuhs, 2003) commonly used and often favored over others in validity and reliability studies (Doscher, 2012) for contributing to higher rates of agreement (Johnson, Penny, & Gordon, 2000, as cited in Doscher, 2012). Inter-rater score reliability is reported in Table 4, in Chapter 4.

Intra-rater Score Reliability

The scorers re-assessed the same DRs at least two weeks after their first assessment, and both scores were compared. Their first and second scores were recorded in a table. The intra-rater score reliability estimates were calculated using the same procedure as the inter-rater score reliability, except that, if a scorer's own scores differed by more than one point, their intra-rater score reliability estimate would have been reported as not having reached the minimum acceptable percentage of 80%. All raters reached acceptable standards of intra-rater reliability, as indicated in table 5, in Chapter 4.

The Documentos Recepcionales

There were ten DRs that were used to pilot the rubric developed. All participants, with the exception of one, stated during the focus groups that the DR, which is the centerpiece of student evaluation, is an appropriate vehicle to demonstrate evidence of the attainment of the five competencies. However, the student who disagreed attributed the inappropriateness of the DR as the only evaluation tool to the lack of assessment organization of the IU, but not strictly to the concept of developing a DR. This participant complained that the presentation of the DR becomes their “exam,” not the evaluation of the actual learning that took place.

The researcher obtained nine DRs from the IU website and one DR directly from a recent graduate via email. All the DRs are from graduates from the same IU where this study took place. The researcher prepared the DRs and submitted them to the scorers without identifiable information. The original titles were also changed for confidentiality purposes. The DRs ranged widely in terms of topics, length, and observable complexity. Some DRs were only in Spanish and others included translations from indigenous languages such as Nahuatl. The DRs were done either in pairs or in groups, but there was no identifiable information on the documents describing teamwork assignments, so only one score was given to each DR. DR10 was used for intra-rater as well as inter-rater score reliability estimates. The following is a general description of each of the DRs used in this study:

DR1: Flower Embroiderers. This DR presents a thorough description of the geographical location and history of the community, as well as an analysis of their present economic conditions. The students worked with a group of women who support

themselves primarily with their embroidery work. The students created a catalogue of embroidery products for the group of women. The purpose of this brochure is to disseminate their embroidery practices and hopefully expand the market for their products, using this brochure as a tool. Page length: 38.

DR2: Traditional Doctors' Catalogue of Plants. DR2 is a catalogue of 142 medicinal plants, with the plants' names in Nahuatl and Spanish, used by the healers of this community, in addition to and the "recipes" for several formulae. This DR includes a peer debriefing segment, because students immersed themselves with the healers for several months. The purpose of this catalogue is to provide the traditional healers and their apprentices with a centralized compilation of information. This DR is a tool for the group of healers to keep a log of their medicinal plants, both in Spanish and in Nahuatl languages and to describe the healing procedures that the plants are used for. Page length: 101.

DR3: Herbalist Shop. This project is a catalogue of traditional medicinal and beauty products. It includes ingredients used and the health issues each product can ameliorate, as well as a photographic registry on certain healing procedures. The purpose of this catalogue is to serve as a tool to pass on herbalist knowledge from mothers to daughters in written format, in addition to orally. This DR also addresses the importance of bridging traditional knowledge with scientific knowledge to create these herbalist products. Because this group of women commercializes their herbalist products, the brochure serves marketing purposes as well. Page length: 23.

DR4: The Group of Healers. This DR is devoted to a group of healers in one of the communities surrounding the IU that the students worked with. The purpose of this

DR is to create a brochure of healing plants to disseminate their use in traditional medicine. In a botanical analysis, each plant is introduced with their indigenous name and the scientific name. The plants' physical description and their uses are included in detail, bilingually –in Spanish and in Nahuatl. Page length: 36.

DR5: Flowers Cross-stich Designers. Very similar to DR1, this is a catalogue of products that women from the community produce. It provides an introduction of the socioeconomic context the group of women live in, and the challenges they face because of the migration of the men in their lives, who emigrate primarily for economic reasons. The purpose of this catalogue is to centralize the Flowers Cross-stich Designers' products for commercial purposes. The catalogue includes a description of the materials used, and the name of the designs in Spanish and in their indigenous language. Page length: 30.

DR6: Fireworks Catalogue. The purpose of this DR is to provide the organization of fireworks handcrafters in the community with a digital collection of their fireworks. This DR includes pictures and descriptions of 14 types of fireworks, as well as a contract template for future business opportunities the firework makers may have. Page length: 12.

DR7: Manual for Handling Fireworks. The purpose of this DR is to provide firework makers with a safety campaign manual for preparing and handling fireworks. This manual includes hand-made drawings and narratives demonstrating proper ways to handle fireworks, along with descriptions of precautionary measures firework makers should take. Page length: 10.

DR8: Photograph Registry for Handling Explosives to make Fireworks. The purpose of this DR is to create an organizational communication handbook used among

the fireworks handcrafters. It includes several photographs and the narrative of proper handling techniques. No further information is provided. Page length: 26.

DR9: Sweet Flavors. The purpose of this DR is to compile recipes and pictures of local, traditional, sweet beverages. This DR includes a description of the climate and biodiversity of the region where these beverages are popular. This description highlights the existence and availability of ingredients to prepare these traditional beverages such as fruits, and plants. The basic ingredient of the beverages is sugar cane, so this DR elaborates on the way it should be grown and processed. Some recipes have Nahuatl translations of the ingredients. Page length: 13.

DR10: Production and Commercialization of Palm Trees. The purpose of this DR is to provide a plan for the commercialization and production of a type of palm tree, which is an important commercial product in the region. This DR includes a theoretical framework, a problem statement and purpose, as well as a description of the methodology used. It presents in detail the extraction process of the palm and the challenges that the peasants encounter, using interviews' quotes to illustrate this. This DR also includes an analysis of the present demand and costs-benefits of the proposed plans, which include reaching more Mexican and international markets. An analysis on participatory action-research is included. Page length: 107.

There was no actual evaluation of the DRs at any point because the focus of this study was on the participatory nature of the process to develop and validate the instrument, and not on the institution or students' performance. The "grade" given by the scorers did not affect students or alumni in any way.

Data Analysis

Data were analyzed in an iterative and collaborative manner during the two focus groups, and systematically during the scoring sessions. Qualitative data gathered during focus groups were organized through the rubric, participants' feedback, and through open coding, which was used to answer the first research question. Open coding allows the researcher to analyze qualitative data through categorizing and describing phenomena, typically with the use of conceptual labels (Strauss & Corbin, 1990). The qualitative data gathered during the scoring sessions in the form of direct feedback from scorers was reported individually, and summarized. Quantitative data were analyzed using acceptable estimates of 80% or higher, and the Lawshe's CVR values and their significance for one-tailed tests.

Rubric Piloting

During the first focus group, the students received a print-out of the five competencies as outlined on the IU's website, and a template of a rubric. Collectively, participants suggested some evidence, behaviors, skills, dispositions, or outcomes that might characterize each competency's attainment. Participants filled out their rubric templates as they discussed each section, or each level of attainment. The researcher then used the data to draft the rubric and submitted it for feedback both on-site and electronically for member-checking to participants. She revised the rubric accordingly until the students indicated it was ready to be submitted to the "experts," who were the second focus group's IU senior students.

A month later, the researcher met with the second group of students who served as a panel of experts, providing feedback on the content of the rubric that their peers

developed. Content validity estimates were obtained. The researcher included the feedback the second group of participants gave and submitted the rubric for their feedback on-site and later electronically for member-checking. Once approved by the second group of participants, the researcher prepared the scoring sessions' materials.

The materials used for the scoring sessions were the final rubric version (Appendix J), the IMD competencies obtained from the IU website, the five-page summary of the researcher's proposal, a scoring sheet (Appendix K), and blind copies of the DRs, assigned to each scorer (table 3, available in Chapter 4). Scorers were contacted via email, informed of the purpose and processes that this study entailed, and asked for their feedback on the rubric and the process itself. In the course of a month, the scorers provided their results and feedback. These data were used to obtain inter- and intra-rater score reliability estimates, reported in Chapter 4. Figure 2 presents a summary of the steps and detailed procedures conducted for the development of the rubric.

Data Management

All data for this study, including transcripts, copies of the graduates' DRs, and data analysis materials, were organized and stored in a locked cabinet in the researcher's home office. All electronic files were organized and stored in the researcher's password-protected personal computer. These electronic files were saved on a flash drive as a back-up. Participants were given pseudonyms to protect their identity throughout the process of this dissertation and when reporting findings. All data will be kept for three years from the study completion.

Figure 2. The development of the IU rubric. This figure illustrates the step-by- step process to develop an instrument to measure the attainment of competencies at the IU.

Integrity Measures

Integrity measures represent a set of criteria used to enhance the rigor of a research study, particularly the accuracy, validity, and trustworthiness of the findings (Creswell, 2003; Newman et al., 2011b; Ridenour & Newman, 2008; Teddlie & Tashakkori, 2009). This research study promoted rigor by using integrity measures

including member-checking, researcher's reflexivity, peer debriefing, an audit trail, and triangulation.

Member-checking

Member-checking is a credibility measure (Creswell, 2003; Leedy, 1997; Newman et al., 2011b; Ridenour & Newman, 2008) defined as the process of having participants review the research report, checking for accuracy and completeness (Leedy, 1997). Participants of the first focus group received via email their focus group's transcripts, and a draft of the rubric which would be used during the second focus group to develop a Table of Specifications. The researcher received the confirmation of the adequacy and completeness of both documents from five of the seven participants. There were no revisions suggested. Participants of the second focus group received via email their focus group's transcripts and the final draft of the rubric which would be used to score DRs. The researcher received the confirmation of the adequacy and completeness of both documents from four of the five participants.

Reflexivity

Reflexivity is a legitimization consideration that allows the researcher to identify "the degree of potential bias in the characteristics of the data, in part, through a process of self-critique" (Newman et al., 2011b, p. 198). With this integrity measure, the researcher should maintain an attitude of transparency, keeping in mind that her position towards what is being studied is not neutral (Newman et al. 2011b). The researcher has an effect on that what is researched, and vice versa. With reflexivity, the researcher is open and makes public her biases in constructing meaning throughout the study (Newman et al., 2011b).

Dietz (2011) reminds us of the need for a “double reflexivity” when working in contexts such as the IUs. This takes, to begin with, being open and reflective about power relations, the role of the researcher, knowledge creation and representation, the “validating audiences” and the group studied (Dietz, 2011). The researcher attempted to maintain a two-way communication with participants to the best of her abilities, assuming a position of equality. Constant communication with professors and peers also allowed the researcher to have reflexivity. She attempted to maintain reflexivity by keeping a research experiences’ journal.

Peer Debriefing

Peer Debriefing is when “[t]he researcher engages other professionals to obtain their perspectives on the researcher’s perception of what was experienced or observed” (Newman et al., 2011b, p. 199). Onwuegbuzie, Leech and Collins (2008) suggest that peer debriefing is comparable to an interview of the researcher, to help her make transparent the processes and negotiations of meaning that took place during the study. The peer debriefers were (a) an instrument development expert, distinguished professor emeritus, presently working at a higher education institution in the United States and (b) a researcher, member of the IU community.

The researcher met with her peer debriefers throughout the development of this study. The researcher held frequent meetings in person with the instrument development, during the course of the study. These meetings were on a weekly basis at the beginning of the study, and on a monthly-basis towards the end of the study. Meetings in person with the researcher member of the IU community were less frequent due to international travel requirements. However, there was constant email communication. Meeting with

these two researchers helped enhance the integrity of the study in several ways: through accountability, mentorship, clarification on complex research procedures, and negotiation of meaning resulting in numerous revisions.

Audit Trail

To maintain an audit trail, the researcher should "... present enough detail, information, and documentation about the research study so that an independent researcher can attempt to replicate the methods used and the findings" (Newman et al., 2011b, p. 200). Audit trail in this study consisted of a thorough description of methodology that was followed with fidelity, and accompanied by rich narratives and explanations. Direct quotes and scanned original material from participants were also provided for transparent documentation.

Triangulation

This integrity measure requires that the researcher obtains data from multiple sources to make sure that there is consistency in the findings. Newman et al., (2011b) suggest that in some cases, triangulation is similar to quantitative concepts of validity and reliability. Triangulation also entails combining and comparing multiple sources of data, data collection and analysis procedures (Teddlie & Tashakkori, 2009).

There were several instances where triangulation took place. For example, while data were primarily collected via focus groups, a Table of Specifications, Lawshe's Content Validity Ratio, the scoring sessions, and direct feedback from participants, the researcher also used the SEP's IU policy book. She used the SEP's IU policy book to establish an initial understanding of the IU's justification of plans and procedures for performance-based assessment, and the role of competencies in the IU programs. The

plans described in the SEP's IU's policy book were corroborated with feedback from participants and personal communication with an IU researcher. It was confirmed that the IU's, in policy and practice, embrace performance-based assessment and used DRs as their final projects as graduation requirement.

Another example of triangulation was done using data from an observation session. During the researcher's visit to one of the IU campuses to invite seniors to participate in her study, there was a seniors' research project meeting. During this meeting, seniors presented their DRs to their classmates and two advisors. The researcher asked one of the facilitators if she could sit in the back of the room, observe, and take some notes, and the facilitator agreed. The observation session was not planned, but emerged as an opportunity when visiting the campus to meet with the seniors. The observations during that session became secondary or complementary data, the type of data that Teddlie and Tashakkori (2009) refer to as "unobtrusive measures" (p. 223). The researcher has the opportunity to examine aspects of the phenomenon studied "without interfering with or changing it" and "individuals under observation will not react to being observed" (p. 223). Her observations helped her to contextualize her research project better and triangulate with students' feedback during the second focus group.

Role of the Researcher

The researcher is Mexican and fully bilingual, Spanish being her mother-tongue. She has studied English throughout her academic education, from primary school to higher education. She has completed a bachelor's and a master's degree in the United States, satisfying all English-proficiency requirements.

Data collected for this research study were obtained in Spanish. The researcher analyzed the material in Spanish and translated it into English. The researcher asked an IU researcher, who is fully proficient in English and Spanish, for feedback on the translation of the IMD competencies. The researcher's mother, medical editor and translator with over 30 years of professional experience, was asked to back-translate the informed consent that the IU students received and signed.

The researcher's personal and professional interest in the IUs emerged from her upbringing in a *mestizo* household where their country's diversity, in all its forms and expressions, was celebrated and respected. The researcher's father, while serving in the House of Representatives of his state, Guerrero, chaired the Education committee. Among the committee's projects was to advocate for and strengthen intercultural education initiatives in a state where 14% of the state's population speak an indigenous language. Speaking an indigenous language is the self-reported indicator the 2000 national census used to report indigenous population representation (INEGI, 2004).

During his term as a government official, the researcher's father advocated together with educators and leaders of the indigenous communities for an Intercultural University to be established in their state. The researcher's father shared with her his vision of empowerment and development of young people, especially of ethnic minority origin, through high-quality education, a vision that had guided his work as a civil servant for over five decades. When the IU in his state³ opened its doors in 2006, the researcher's father invited her to the inauguration ceremony, an event that sparked her personal interest in the IUs.

³ This study did not take place in Guerrero's IU

Summary

In this chapter, the research design for this dissertation was explained, followed by a description of criteria regarding sampling strategies, data sources, and data collection. A detailed process of the creation and piloting of the rubric was presented. Finally, the researcher presented procedures for data management and integrity.

Chapter IV

RESULTS

The purpose of this study was to develop a process to create a performance-based assessment instrument with acceptable standards of content validity and inter- and intra-rater score reliability estimates, to assess students' attainment of competencies at an Intercultural University in Mexico, in the IMD major. The two research questions guiding the study were: (a) According to IU seniors, what evidence demonstrates the attainment of competencies of the IMD academic program of study at an intercultural university in Mexico? and (b) Can acceptable standards of validity and reliability estimates be established for an instrument developed to measure students' attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?

This chapter presents an analysis of the data collected to (a) determine the evidence of attainment of competencies, and (b) estimate content validity and inter- and intra-rater score reliability of the rubric.

Data collection followed the sequence discussed in Chapter 3 (Figure 1) and the steps to develop the rubric (Figure 2). Feedback from focus group participants and scorers is presented and discussed, followed by an analysis of each step of the rubric-development process. The chapter concludes with a summary of the findings.

Data Analysis

The results and data analysis are presented in six sections. The first section of the data analysis was devoted to the students' evidence of attainment of competencies obtained during the two focus groups, evidence that became the content of the instrument

developed. The second section is devoted to the estimates of content validity of the instrument, which were obtained by using a Table of Specifications and calculating Lawshe's CVR. The third section is devoted to focus group participants' feedback on PBA and the rubric-development process, and includes feedback via email of an additional graduate unable to attend the focus groups. Although no changes were made to the rubric based on his comments, the graduate's feedback on the attainment of competencies is also included. His feedback was informative and useful for articulating implications for practice, as discussed in Chapter 5. In the fourth section, estimates of inter- and intra-rater score reliability are presented, followed by a fifth section that includes scorers' feedback on the rubric-development process. The sixth section includes the researcher's reflections on the entire instrument-development process.

The IU Students' Evidence of Attainment of Competencies

During the two focus groups, participants described the evidence of attainment of the five pre-established competencies outlined in the Intercultural Management of Development major. These competencies indicate that students should be able to (a) articulate regional/indigenous knowledge and initiatives, (b) facilitate advocacy for, and management of resources and information, (c) strengthen and design intercultural initiatives, (d) highlight regional/indigenous knowledge and initiatives, and (e) generate diagnostic and proposal-oriented knowledge.

Guided by the researcher's questions (Appendix A and Appendix B) during the two focus groups, evidence provided orally by the participants was organized into a rubric format. Examples of these questions are: "In your own words, what does Competency 1 mean, and what evidence indicates that you have achieved this outcome?"

and “What does it look like to attain Competency 1 at the Beginner level?” Using open coding, the researcher used the final draft of the rubric, representing the input of both focus groups’ participants, to find evidence of attainment of competencies. Additionally, further evidence, examples, and quotes to illustrate the findings were found in the focus group transcripts and the DRs. Open coding is a process that allows the researcher to analyze qualitative data through categorizing and describing phenomena, often with the use of conceptual labels (Strauss & Corbin, 1990). In this case, the “phenomena” are the actual or potential valued performances found in the student-developed rubric produced during the two focus groups. Seven categories emerged through open-coding, discussed below.

According to the 11 participating IU seniors and the graduate, the evidence that demonstrates the attainment of competencies outlined in the IMD major was a complex set of acquired, potential, and/or desired performances or achievements. These performances and/or achievements pertain to professional, intellectual, and personal areas of students’ life, and reflect students’ values and goals for their education and completion of their degree.

The acquired, potential, and/or desired performances or achievements can be reflected in the final research projects that students produce as a graduation requirement, even when a student might still be in the process of attaining the competencies fully. Once democratically-negotiated by participants, each of the acquired, potential, and/or desired performances and/or achievements can be coded into levels of attainment and used to develop a rubric such as the one created in this study (Appendix J).

The following categories are performances or achievements that represent the evidence of attainment of competencies that students either possess, or value and therefore desire to attain. This list of categories is not exhaustive, and there is overlap among them. Nevertheless, the categories provided a clear language to refer to the attainment of competencies, either achieved or valued, by the IU students and recent graduate.

To illustrate each of the categories, the researcher provided examples from the rubric's content, the focus groups transcripts, and DR content that exemplified the categories below. Participants from the first focus group were labeled F1 through F5 for women, and M1 and M2 for men. Participants from the second focus group were labeled FA through FC for women, and MA and MB for men.

Open Coding Categories

Collaboration and outreach. Students expressed the need and desire to develop through the IMD program better collaboration and outreach skills and dispositions, such as creating links with institutions and non-governmental organizations, or networking. Examples of achieved or valued performance in this category are, in Competency 1, Satisfactory Level: “[The student] creates links with different consultants,” and in Competency 1, Exemplary Level: “[The student] creates regional and inter-regional links in addition to creating links with institutions.” These practical examples from the student-developed rubric illustrate the IU students’ awareness of the importance of collaboration and outreach in an increasingly interconnected world.

Analytical and critical thinking. Students understand the need to consider a community development opportunity or problem from a variety of perspectives, and to

demonstrate their ability to see and act upon the interconnectivity of those perspectives. Examples of these achieved or valued performances are in Competency 4, Beginner Level: “[The student] starts to visualize and compare regions,” and in Competency 4, Satisfactory Level: “He/she criticizes and analyzes in proposal-like ways, and with theoretical and practical foundations that are reflected in practice.”

Acquisition and appreciation of knowledge. Epistemology is a prevalent theme in the IMD curriculum. As students acquire indigenous and academic knowledge, they seek to understand more about the difference, and the ways they converge. As they design community projects that create bridges connecting different ways of knowing, students value their role as knowledge advocates. The students’ acquisition and appreciation of knowledge is exemplified in Competency 5, Developing Level: “The student fosters valuing those [different] types of knowledge... [and] is able to generate new knowledge without forgetting, or considering of less value, the knowledge that had already been established.”

Planning and implementation skills. At the IU, students are evaluated through their final research projects (DRs), which are action-research, community development initiatives. Therefore, planning and implementation skills are essential for completing DRs and for professional development. These skills are referenced in the rubric in various ways, under several general competencies. Examples are Competency 3, Beginner Level: “He/she is able to see areas of lacking and weaknesses.” and Competency 2, Satisfactory Level: “[The student] designs and carries out his/her project, and evaluates its impact.”

Creativity and resourcefulness. Students value creativity and want to be able to explore different ways of designing and implementing community development projects. Some examples of creativity and resourcefulness are in Competency 4, Exemplary Level: “[The student] is able to go beyond what was learned in the classroom and is open to new knowledge,” and “[the student] sees preliminary results and endeavors to undertake new community development projects. He/she is able to design and re-design plans of action.”

Empowerment and autonomy. Students expressed valuing empowerment and autonomy, both potential and achieved. The IU offers students an open initiative to be advocates for their community, and to attain this, empowerment and autonomy are fundamental components. Autonomy combined with empowerment allows students to develop professionally because it results in productivity and consultancy capacities. This is exemplified in Competency 5, Exemplary Level: “[The student] is able to advise/serve as a consultant to whoever comes to him/her for services.”

Empathy, solidarity, and advocacy. Students view their work as more than project design and implementation. They look forward to becoming advocates for their community’s development, improvement, and empowerment and appreciate being part of their community. One of the students (F2) commented the following: “[I am drawn to the ability] to work in the field with the community, which is precisely something that differentiates us from other universities... [the] exchange we have within the communities, the ability to share with them, and learn about their needs, [and] the fact that you often times become the one that gets them out of the problem they... [are] in.” In the rubric, there are three examples of empathy, solidarity, and advocacy. The first is

Competency 1, Satisfactory Level: “[The student] is able to advocate in favor of the community. His/her techniques are truly participatory and this participation genuinely includes community, students, and institutions.” The second example is Competency 3, Exemplary Level: “He/she inserts him/herself and feels totally [a] part of the community.” The third example is Competency 2, Exemplary Level: “[The students] design projects from the ‘bottom up’ and not ‘top down’. He/she not only advocates, but empowers others in the community, teaching them to advocate. He/she intervenes in returning empowerment to the people.”

In summary, according to focus group participants, the evidence of attainment of competencies in the IMD major was a complex set of acquired, potential, and/or desired performances or achievements. These performances and/or achievements pertain to professional, intellectual, and personal areas of students’ life, and reflect what students value and aspire to have as outcomes of their education and completion of their degree. Once democratically-negotiated by participants and coded into levels of attainment, each of the acquired, potential, and/or desired performances and/or achievements can be used to develop a rubric. These acquired, potential, and/or desired performances or achievements can be reflected in the final research projects that students produce as a requirement for graduation.

Estimates of Content Validity

Content validity in this study was operationally defined as the estimate of how representative are items of the content the instrument is designed to measure (Newman et al., 2006). Estimates of content validity of the rubric were obtained during the second focus group, where participants served as expert judges. After the first group of

participants (six IU seniors and one graduate) created the rubric, the second group of participants (five IU seniors) was asked to estimate how representative the rubric's descriptions of attainment of competencies were of the attainment levels. Participants' responses in consensus form were used to report content validity estimates. Below is a detailed description of the content validity estimates' procedures and analysis.

Content validity estimates were obtained in two ways: through expert judge agreement using a Table of Specifications (ToS; Appendix E), and by calculating Lawshe's Content Validity Ratio, which is a formula to quantify consensus among experts regarding the content of an instrument (Guion, 1977; Lawshe, 1975). The generally acceptable estimate with expert judges' agreement is .8, or 80% (Newman et al., 2011b). There were five senior students serving as "experts" in the study; having consensus of at least four students yields 80% agreement. For the Lawshe's content validity ratio, when all experts say the description is appropriate, the computed CVR is 1 which indicates total agreement. With five experts, a Lawshe's CVR of .99 is considered to be an acceptable standard. See Appendix G for Lawshe's CVR values and their significance. The formula is as follows:

$$CVR = \frac{n_e - N/2}{N/2}$$

n_e = the number of experts who agreed on the relevance of the item, behavior, or question

N = total members of the panel of expert judges

Using the first group of participants' rubric and their ToS, the experts, the second group of participants, or "experts," discussed and evaluated the content of all competencies, at all four levels of the rubric: Beginner, Developing, Satisfactory, and

Exemplary. Although some of the students' ToS boxes were not filled in completely, students' agreement and negotiation was reflected in the researcher's audio recording and written transcript of the focus groups. Because the researcher observed the students more engaged by conversations about different levels of attainment - than by filling out the ToS, she made sure that they responded orally and spoke their agreements, in addition to filling out their ToS form.

The following excerpt from the second focus group transcript is a representative example of how an attainment description was revised, demonstrating that the researcher sought oral and written consensus from students regarding the content of each of the rubric's levels of attainment. The competency being discussed was Competency 1, Exemplary level:

Original description:

The [student] demonstrates expert-like diagnostic capabilities and goes beyond what is taught in the classroom. He/she does not romanticize the IMD practice. He/she articulates regional initiatives; creates regional and inter-regional links in addition to creating links with institutions and with non-governmental organizations. He/she shows the potential of being self-employed.

Revised description:

The [student] demonstrates expert-like diagnostic capabilities and goes beyond what is taught in the classroom. He/she articulates regional initiatives; creates regional and inter-regional links in addition to creating links with institutions and with non-governmental organizations. He/she shows the potential of being self-employed.

The students were discussing if the phrase “does not romanticize” should be included in the description. Their discussion went as follows:

MA: [Reading original description of competency] “Does not romanticize,” what’s that?

FC: “Does not romanticize”... They mean that, we don’t speak nice things about our program of study... Or rather, that we go straight to the facts, or... or... I don’t know.

FB: I would remove that part only

FC: The “goes beyond what was learned in the classroom”...?

FB: that of “does not romanticize”

FP (researcher): So I cross that out?

Voices: Yeah, uhuh

FP: And what should go in its place? I just cross it out and goodbye? [Voices: Uhuh].

So then, do you think it now aligns with what Competency 1, Exemplary level should be?

In this little box? Should it be there?

MA: Yes, right? [to peers]

FC: Mmm... wait, not yet [student reads original description of Competency 1 again]

[students talk among themselves]

FP: [pause] If I remove the part where it says “does not romanticize,” would you accept this [description] in the Exemplary level, with that edition?

FA: Yes

FB: Well, yes, it goes there

FP: Do you all agree? [researcher looks for all answers, or students’ nod]

Voices: Uhuh

There were three descriptions of attainment levels in the rubric that were accepted unanimously with no revisions, yielding 100% agreement. These were: Competency 1 Developing level, Competency 2 Beginning level, and Competency 4 Exemplary level. The group initially rejected competency 2 Developing level. However, the students provided a new description for Competency 2 Developing level which was then accepted unanimously, yielding 100% agreement. The remaining 16 attainment descriptions were unanimously accepted, with feedback. Feedback included some suggested changes, and once revised, the competency attainment description was accepted. Once consensus on the feedback pertaining to these descriptions was reached and included, the 16 attainment descriptions yielded 100% agreement. The original version of each of the competency attainment level descriptions, with all the researcher's notes on the students' feedback, is available in Appendix I.

After estimating content validity using the ToS, the researcher calculated content validity on the final draft of the rubric with a second measure, Lawshe's Content Validity Ratio (CVR). Three of the 20 descriptions were accepted as they were originally, 16 out of 20 were accepted with feedback, and one was revised entirely and then accepted once the revisions were incorporated, and then accepted.

Lawshe's CVR was calculated only once because all the levels' descriptions were eventually accepted. Lawshe's CVR calculation is as follows, where n_e is the number of experts (five students) who agreed on the relevance of the item, behavior, or question, and N is the total number of members of the panel of expert judges (five students). A CVR of 1 represents total agreement among expert judges:

$$CVR = \frac{n_e - N / 2}{N / 2} = \frac{5 - 2.5}{2.5} = 1$$

The evidence based on the calculations of agreement of the IU students on the content of the rubric using a ToS and the calculation of Lawshe's CVR indicate that the rubric yielded acceptable standards of content validity. The final version of the rubric is available in Appendix J.

Focus Group Participants' Feedback

Feedback from students was gathered during two focus groups, as the researcher introduced the purpose of her study and explored the students' experiences with performance-based assessment at the IU. The open-ended questions that elicited this feedback included: "How would you describe your assessment experience at the IU?" and "What do you think is the best way to show attainment of competencies at the IU?"

Participants from both focus groups discussed several issues pertaining to the process of completing their DRs, a process evaluated only through performance-based assessment. Participants discussed how their evaluations are accomplished, where they see room for improvement in the evaluation process, and suggestions they have for implementing those improvements. A concern about fairness in grading when developing and submitting their DRs was discussed, as well as the need for organization, independence, and the freedom to choose projects in diverse subjects and formats.

Additionally, participants talked about peer-evaluation as a possible way to improve DR assessment and completion. Although peer-evaluation had been tried before at the IU, participants did not explicitly suggest peer-evaluation was the solution to what some of them perceived as an often disorganized, and sometimes unfair or arbitrary

scoring process. The following is feedback from both focus groups' participants regarding PBA and the process of completing and scoring DRs at this IU.

On different occasions, participants mentioned the need for a more organized system that allows them to flourish and have freedom, but that also imparts fairness in grading. The students' concerns included their tutors' busy schedules, the fact that one's grade sometimes "depends on the professor" (F3), or that "professors don't come to agreement as to what they will evaluate [students] on" (F2). F4 explained the way DRs are created, and her concerns:

... every semester we conduct participatory research during our fieldwork, and then we systematize the information, and then we submit a final report, and really... I have realized that, yes, well... there are several classmates that are truly giving their all and I feel we are missing, yes, we are missing that systematization of information, and that is why some of them perform a little poorly... And I don't like that and there have even been... meetings about that... and I feel that sometimes we are evaluated well, and in some instances as I told you, we... attain many theoretical abilities and we do heuristic work but I feel that even then, there is a certain... well, on behalf of professors, right? Some evaluate some students well and others not and I don't like that. I feel that we all know.

For F5, it is important not only to be organized, but also to be able to evaluate projects fairly. Fair grading should be able to be implemented despite the different designs or approaches the DRs might have:

a lot of times our [DRs] are being evaluated in the same manner than others, right? When precisely, our projects are very different, right? Maybe some are

qualitative, others quantitative, and then, ehm, the way already established of wanting to evaluate them all the same way affects us because they are not the same. We don't use the same methodologies, the same techniques... And often times we let evaluation of professors and students be guided by 'I like you,' or 'this,' or 'that.'

Addressing more organizational issues, students expressed that an "anchor" of understanding, a tool to centralize communication, could be beneficial to address concerns about the quality of their work and fairness in grading. The researcher explained to participants the purpose of rubrics, and the purpose of her study, and the students asserted that a more organized system would be beneficial. Students mentioned that school officials had produced an earlier initiative to create a centralized guide that outlined the competencies for each of the IU's major tracks, but that it did not yield decidedly productive results. "Supposedly, there was consensus," said F3, "and the professors said... 'we went to such meeting and here are all the agreements signed.' Well, I mean, of what good is your [agreement] document, and that you have it signed and that all the competencies are outlined there if everyone is going to have his/her own requirements?"

For student FB, the lack of consistency and organization when evaluating the DRs was problematic because it resulted in a limited view of her learning. She mentioned that when the time comes to present her DR, her presentation performance becomes the "exam." She said: "Our exam is the presentation of our paper. That's the evaluation we get here."

Moreover, participants pointed out the need for independence and diversity of projects when submitting their DRs. F4 suggested having alternative ways to present their DR's, not only written: "maybe a video, I don't know, other ways to graduate."

Although participants did not consider peer-evaluation as the solution to assessment issues, they did talk about issues and possible benefits of peer evaluation. An issue expressed by student MA was the fairness of teamwork distribution and performance. MA said, "on many occasions we work in teams and I consider that some... ehm, some [peers] work more or work less and that is what... ehm, I am not satisfied with, with the way we are evaluated, right? I think that there are some that do not deserve the grade and despite this, they are there... they receive it."

Some students commented that they had tried peer-evaluation before, although it wasn't clear whether or not it was successful. They would base each other's grades "on the performance of your work... it was a process" (FB), and their experiences throughout their college years. A student mentioned:

You have seen a trajectory of work, and you already know the way[your peers] work. For example, me with her (points at FB), since we enrolled we have been a, a research team. Then, for example, we distribute the work. It is my task to write up and edit, and her... she takes care more of the format of the paper, like that. Therefore, based on that, I can say, 'well, she did it and I like the end product' and I can then evaluate her. For me this is the best type of evaluation.

In summary, focus group participants discussed their concerns about organization and fairness in grading when developing and submitting their DRs, as well as the need for independence and having the freedom to work on diverse projects presented in varied

formats. Participants also talked about peer-evaluation as an option to assess each other's work, although this was not suggested as the solution to a disorganized or unclear assessment system.

Feedback from Another IU Graduate

At the time in which this study took place an IU graduate pursuing a master's degree in another Latin American country expressed a desire to participate in this study. Although this graduate was contacted during initial stages of this study when the researcher was looking exclusively for graduates to participate in focus groups, this graduate was not in Mexico and therefore could not participate. Nevertheless, he wanted to share his feedback on the process proposed in this study. The researcher shared with him the rubric that his peers had created, and he submitted a page-long response of positive feedback. He highlighted the importance of clearly defining the competencies and the attainment levels, which is seldom discussed at the IU: "It is a topic not made sufficiently explicit in the program of study at the IU." The graduate also provided feedback on the attainment levels' descriptions, using his and some of his colleagues' post-graduation experience. These comments were received after the scoring sessions so they were not included in the rubric. The following are some highlights of the graduate's feedback.

Competency 1:

Specifically in the Exemplary Level, the graduate highlighted the importance of developing networks with institutions and governmental entities. Developing professional networks was something that his younger colleagues had added to their rubric but not with the emphasis that he called for. He wrote:

let me share with you what I have found in my field work, and that is that the linkages one creates as an expert [alumnus] are still linked to a conceptual notion of interculturalism as a process of mere fellowship and connections with social actors, robbing importance away from the linkages made with institutions and governmental entities [for future employment purposes] because the ideology that they are ‘the bad guys of the story’ still prevails... This opinion might be prejudicial but these are the comments I have heard... from graduates that have not found employment yet and they point out that it is due to the lack of these linkages with different institutions.

Competency 2:

The graduate mentioned that students learn to make oral and/or formal agreements when they engage in fieldwork at the IU, and pointed out that this is missing from the descriptions of the attainment of competencies in the rubric.

Competency 3:

According to him, students should go through a phase of personal analysis of their professional profile. He mentioned that “the students know [or should know] the strengths their major offers and therefore should be generating networks of knowledge with other... disciplines, in addition to reflecting on his/her process before, during and after [their studies].” Students had mentioned during the first focus group that competency #3 was “vague” to them, and even one of them, FB, suggested merging all into only the first two. The students did not do this, but FB’s comment was seconded by at least another student. FB said the following: “I feel that the first one and the second

one could be one [other student: “aha”] and supersede the others [other student: “yes”] because if we look, everything is repeated, right?”

Competency 4:

The graduate suggested adding to this competency more description regarding intercultural relations. If not made more explicit, intercultural relations might be reduced to “mere fellowship of equals.” The students had briefly alluded to this in their “Exemplary” level description, saying that work should be “horizontal” between them and the community. The graduate suggested describing these relations as sources of “synergy and alternative solutions to problems and needs” identified.

Competency 5:

In terms of this competency, the graduate suggested adding the students’ “development of organizational and classifying skills of [the community’s] needs and knowledge.” With practice the student should also demonstrate the ability to link theoretical and empirical concepts, allowing for the “dialogue of knowledges,” to take place between Western and indigenous ways of knowing. These recommendations had previously been suggested by focus group participants and were already present throughout the rubric (Competency #5) in a variety of phrasings.

In summary, feedback was submitted by a graduate who could not be present during the focus groups. Among the arguments he shared with the researcher was that at the IU, discussions on the attainment of competencies are not sufficiently emphasized, resulting in a lack of shared expectations throughout the learning process. Although the graduate also suggested additional wording for each competency, neither his feedback nor that of the scorers resulted in changes made to the student-generated rubric. Only the

focus group participants' suggested changes were included. However, in Chapter 5, scorers' or the alumnus' feedback is considered for discussing implications for practice and suggestions for further research.

Estimates of Reliability

After content validity estimates were obtained, the rubric was subject to reliability estimates using an archive of DRs. All identifiable information from the DRs was removed. Estimates of score reliability in this study were obtained in two ways: with the percentage of agreement between scorers, or inter-rater score reliability (Newman et al., 2011b), and within each rater's score over a period of time, or intra-rater score reliability (Moskal & Leydens, 2000). If a measurement is accurate - in this case, the scores that the rubric yields- as Teddlie and Tashakkori (2009) point out, it should be repeatable over time, or obtainable with an identical measurement strategy. Inter-rater and intra-rater score agreement of at least 80%, is sufficient according to Newman et al. (2011b) to meet acceptable standards of score reliability.

Inter-rater Score Reliability

Upon completion of the two focus groups, content validity estimates were obtained, and the final draft of the rubric was produced. Seven scorers were contacted and were given identifications of "S1" through "S7." The researcher obtained and prepared nine DRs from an archive available on the IU website and one DR that the author shared directly with the researcher. The researcher removed the authors' names and IU information before submitting the DRs to the scorers.

The rubric includes the five IMD competencies and four descriptions of levels of attainment for each competency. The levels of attainment are Beginner, Developing,

Satisfactory, and Exemplary. The Beginner level represents a 1-point score, the Developing level a 2-point score, the Satisfactory level a 3-point score, and the Exemplary level a 4-point score. The highest score a DR could receive on each competency was 4, and the highest total score a DR could receive was 20.

Table 2 includes the assigned DRs and the return date. If the results of the first two scorers were more than one point apart, this was considered a scoring discrepancy. In case of scoring discrepancy, the DRs were then submitted to be scored for a third time, with the third scoring conducted by a different rater. Table 2 includes the instances in which a third scorer was needed. The score agreement calculations are the number of scorers who considered the score to be within one attainment level difference. For example, if scorers gave the same score, this represents 100% agreement; if the first scorer gave a 4 and the second scorer gave a 3, this also represents 100% agreement because there is only one point difference. This score discrepancies resolution is a type of parity model (Johnson, Penny, Fisher, & Kuhs, 2003) commonly used and often favored over others in validity and reliability studies (Doscher, 2011) for contributing to higher rates of agreement (Johnson, Penny, & Gordon, 2000, as cited in Doscher, 2011).

Table 2

Documentos Receptionales Scored and Return Date

Scorer	Round One	Round Two or Three, if Applicable
	DR Assigned	DR Assigned
Scorer 1	DR4, DR9, and DR10 July 29	DR2, DR3 (third round), and DR10 Aug 12
Scorer 2	DR3 and DR10 July 25	DR4, DR9, and DR10 Aug 14
Scorer 3	DR5, DR9, and DR10 Aug 1	DR1, and DR10 Aug 24
Scorer 4	DR10 July 18	DR10 July 27
Scorer 5	DR1, DR5, and DR10 July 23	DR6, DR8, and DR10 Aug 6
Scorer 6	DR2, DR7, DR8, and DR10 July 28	DR1, DR4, and DR10 Aug 18
Scorer 7	DR3, DR6, and DR10 July 27	DR2 (third round), DR4, DR7, and DR10 Sept 15

Table 3 includes total scores for each DR in each scoring round. Score agreement was determined through consensus on each competency's score, and calculating the resulting percentages. This calculation was done for each competency, and then an overall percentage was reported as indicated in Table 4. The original distribution of DRs was random. Scorers submitted their results back to the researcher and she then recorded the scores in the order in which they were received.

Table 3

Total Scores in each Scoring Rounds

DR	Round One	Round Two	Round Three	Average Score
	Total Score	Total Score	Total Score	
DR1	12	18	n/a	15
DR2	18	10	7	11.6
DR3	10	5	13	8.3
DR4	5	13	12	10
DR5	13	13	n/a	13
DR6	5	11	n/a	8
DR7	10	5	n/a	7.5
DR8	6	6	n/a	6
DR9	5	7	n/a	5.5
DR10	18, 17, 20, 17, 20, 20, 5	18, 15, 20, 17 20, 20, 5	n/a	16.4

One of the non-Mexican scorers, Scorer 4 (S4), expressed divergent negative opinions regarding the process and the grading exercise, as discussed in the next section, and only provided final scores for the DR10 on both rounds. Because each DR had to be scored at least twice, the researcher distributed the two DRs that Scorer 4 received, which were DR7 and DR8. These two DRs were assigned to Scorer 6 and Scorer 7. DR8 was given to Scorer 6 in the first round, and DR7 was given to Scorer 7 in the second round.

This re-distribution was not random. The researcher saw this as an opportunity to have IU's scorers score more DRs should further analysis of inter-rater reliability estimates be needed. Scorer 7, who is a Mexican citizen from an IU, also expressed negative opinions, but Scorer 7 did complete the scoring of all the DRs assigned and seemed to be personally interested in the topic of assessing DRs.

Scorer 4 and Scorer 7 had critical opinions about assessment in general and about using this rubric in particular. For Scorer 7, this process represented "Western dominion, where Western values are privileged over those of the communities that are to be served." For Scorer 4, the process assumed that "an external observer, foreign to the context, can evaluate a program [sic] looking only at a final product," in this case the product being the DRs. In terms of the rubric itself, Scorer 4 expressed that "there was nothing in the [DRs] that could allow [him] to evaluate" the competencies. Scorer 7 in particular expressed that the rubric's descriptions of attainment did not reflect the high level of complexity of the IMD competencies, and that the rubric had a more technical 'paperwork' character to it.

Scorer 4 gave the researcher the option to exclude this scorer from the scoring exercise. The researcher acknowledged Scorer 4's philosophical objections, and reiterated that the scoring was not going to affect any student, because the authors of the DRs have graduated, and that scores were going to help determine the instrument's stability independent of philosophical objections. Scorer 4 completed the scoring but, because he/she considered that he/she did not have enough background information, Scorer 4 could only give a final score on DR10. On the other hand, although Scorer 7

had negative opinions on the process and the rubric, Scorer 7 did not consider not completing the scoring rounds.

The researcher decided to keep the two scorers for three reasons. First, their participation was helpful to see if the instrument was stable over time in terms of statistics, although having five scorers would have still satisfied the minimum number sufficient, according to the literature (Newman et al., 2006). Second, the researcher considered the intellectual exchange with the two scorers to be formative for her, and informative to her study. Third, she considered it inappropriate to remove scorers that did not find her process or the rubric to be totally adequate or acceptable. Their involvement remained voluntary at all times.

Table 4 includes the scores given to each competency and the agreement calculations. As indicated in Table 2, some DRs were scored more than once in case there were discrepancies between the first two scorers. The final percentage agreement was calculated based on the number of competencies that achieved the minimum of 80% agreement. For example, if competencies 1, 2, 3, and 4 had a 100% agreement and 5 had less than 100%, the final percentage of agreement is 80%. This meant that four out of five competencies complied with the minimum criterion.

Table 4

Scores on each Competency and the Inter-rater Agreement Calculations

DRs and Rounds of Scoring		Competency Number and Scores					% Agreement
		#1	#2	#3	#4	#5	
DR1	First Scorer	3	2	2	2	3	80%
	Second Scorer	4	3	3	4	4	
	% of agreement	100	100	100	50	100	
DR2	First Scorer	3	4	4	3	4	40% (third score needed)
	Second Scorer	2	2	1	3	2	
	% of agreement	100	50	50	100	50	
	Third Scorer	2	1	1	2	1	40%
	% of agreement	100	66.6	66.6	100	66.6	
DR3	First Scorer	2	1	1	3	3	60% (third score needed)
	Second Scorer	1	1	1	1	1	
	% of agreement	100	100	100	50	50	
	Third Scorer	2	3	3	3	2	20%
	% of agreement	100	66.6	66.6	66.6	66.6	
DR4*	First Scorer	1	1	1	1	1	40% (third score needed)
	Second Scorer	3	2	3	3	2	
	% of agreement	50	100	50	50	100	
	Third Scorer	2	2	2	3	3	40%
	% of agreement	66.6	100	66.6	100	66.6	
	Fourth Scorer	4	3	4	4	4	
DR5	First Scorer	2	3	3	2	3	100%
	Second Scorer	3	2	2	3	3	
	% of agreement	100	100	100	100	100	

Table 4

Scores on each Competency and the Inter-rater Agreement Calculations

DR6	First Scorer	1	1	1	1	1	80%
	Second Scorer	2	3	2	2	2	
	% of agreement	100	50	100	100	100	
DR7	First Scorer	2	2	2	2	2	100%
	Second Scorer	1	1	1	1	1	
	% of agreement	100	100	100	100	100	
DR8	First Scorer	1	2	1	1	1	100%
	Second Scorer	1	2	1	1	1	
	% of agreement	100	100	100	100	100	
DR9	First Scorer	1	1	1	1	1	100%
	Second Scorer	1	1	2	1	2	
	% of agreement	100	100	100	100	100	
DR10	Scored by the seven scorers	4, 4, 4, 3, 4, 4, 1	3, 3, 4, 4, 4, 4, 1	4, 4, 4, 4, 4, 4, 1	3, 2, 4, 3, 4, 4, 1	4, 4, 4, 3, 4, 4, 1	80%
	% of agreement	85.7	85.7	85.7	71.4	85.7	
DR10	Scores excluding the two scorers with negative opinions (S4 and S7)	4, 4, 4, 4, 4	3, 3, 4, 4, 4	4, 4, 4, 4, 4	3, 2, 4, 4, 4	4, 4, 4, 4, 4	100%
	% of agreement	100	100	100	80	100	

Note * DR4 was given to all the Mexican scorers. The first scorer is a professor at a non-IU university in Mexico. The second scorer is a professor at an institution in the US. The third and fourth scorers are professors at an IU in Mexico.

Results indicate that the inter-rater score agreement was 80% or more in seven of the ten student projects scored. This estimate of 80% or higher inter-rater score reliability fulfills acceptable standards for reliability (Newman et al., 2011b). The three DRs that did not reach inter-rater agreement above 80% were DR2, with 40% agreement, DR3, with 60% agreement, and DR4 with 40% agreement. DR2 was scored by three Mexican scorers: one faculty member at a non-IU university in Mexico, and two faculty members at an IU. DR3 was scored by two Mexican scorers: one faculty member at a non-IU university in Mexico, and one faculty member at a non-IU institution in the US. DR4 was scored by the four participating Mexican scorers. These three DRs, DR2, DR3, and DR4, were very different from each other. The topics were very different, as was the length and extent of the narrative.

Consensus Frequencies

Consensus frequencies were calculated for each competency, and on each of the ten DRs⁴. Of the ten scorings, scorers reached consensus of 80% or higher nine in Competency 1, seven times in Competency 2, seven times in Competency 3, six times in Competency 4, and seven times in Competency 5.

Intra-rater Score Reliability

The seven scorers received DR10 for estimating intra-rater reliability. The first round of scoring took place between July 16 and August 1, 2012 and the second round took place between July 27 and September 15, 2012. Table 5 shows the intra-rater reliability estimates for each scorer.

⁴ Only the first scoring of the DR10 was considered for this frequency count

Table 5

Intra-rater Reliability Estimate

Scorers & Scoring Rounds		Competencies and Scores					Total	% of Agreement
Scorer 1	Round 1	4	3	4	3	4	18	100%
	Round 2	4	3	4	3	4	18	
Scorer 2	Round 1	4	3	4	2	4	17	100%
	Round 2	4	2	3	2	4	15	
Scorer 3	Round 1	4	4	4	4	4	20	100%
	Round 2	4	4	4	4	4	20	
Scorer 4	Round 1	3	4	4	3	3	17	100%
	Round 2	3	4	4	3	3	17	
Scorer 5	Round 1	4	4	4	4	4	20	100%
	Round 2	4	4	4	4	4	20	
Scorer 6	Round 1	4	4	4	4	4	20	100%
	Round 2	4	4	4	4	4	20	
Scorer 7	Round 1	1	1	1	1	1	5	100%
	Round 2	1	1	1	1	1	5	

Results indicate that seven out of seven scorers had 80% or higher intra-rater agreement, fulfilling the minimum criterion for establishing significant agreement (Newman, et al., 2011b).

Feedback from Scorers

As discussed earlier, seven scorers from diverse backgrounds, four women and three men, participated in this study. They all hold advanced university degrees in education and all are presently working in the field of higher education, including two scorers who work at one of the IU network institutions. Scorers received the DRs, the rubric, the list of IMD competencies, and the researcher's five-page proposal summary, and were asked to provide feedback on the process proposed in this study. The amount of feedback the researcher received ranged from one sentence to two pages.

Scorer 1

Profile: Mexican Professor in Mexico at a non-IU institution

Scorer 1 provided very brief feedback. Scorer 1 found the process and the instrument to be "useful" and "informative." Regarding the DRs, Scorer 1 mentioned that the communities where the DRs took place were portrayed positively and as hard-working, entrepreneurial people.

Scorer 2

Profile: Mexican Professor in an American Institution

Scorer 2's feedback focused on weaknesses in the structure of the rubric, not on its development. Scorer 2 suggested the need for a clearer continuum from one level to the next:

The rubric seems to be clear in the beginning as to what the objectives for the works [sic] are supposed to comply [sic; comprise]. Nevertheless, when it is applied in the assessment it seems inadequate as the linear progress from 1 to 4 explains, it makes it confusing because the projects show levels of good performance at 3 or 4 but they miss some points enlisted in the lower scores. An example of this would be in the project DR3 on competency 4. The score that I marked was 3 but it does not present any information pertaining the points listed in scores 1 and 2.

Scorer 3

Profile: Non-Mexican Professor in an American Institution

Scorer 3's feedback suggests that the process yielded a clear rubric, but using it without more familiarity with the DRs' contexts could be challenging. Scorer 3 wrote the following:

The rubric was rather clear. It was difficult to apply the rubric (at first) without having had practice with other texts or materials. I had to read all three texts with the rubric and then go back and evaluate them using the rubric.

Scorer 4

Profile: Non-Mexican professor at an American University

Scorer 4 did not provide a complete score for any of the assigned DRs, with the exception of DR10. Scorer 4 reluctance to complete this assignment stemmed principally from his philosophical stance on educational assessment. Scorer 4 believed that a lack of context prevented him from making an evaluative statement: "Since I don't know the context, starting point or goals for this catalogue [the DR], I cannot score this

competency.” Fundamentally Scorer 4 believes that the problem resides in “the idea of performance-based assessment, that assumes that an external observer, foreign to the context, can evaluate a program only seeing the final product.” To value the process, Scorer 4 would need to be familiar with the students’ starting point, their surroundings, the resources available, and pressing issues for students.

Scorer 5

Profile: Non-Mexican professor at an American University

Scorer 5 found the process to be good in general, but suggested that the structure of the rubric be revised, which is similar to what Scorer 2 said. Scorer 5 wrote: “The rubric is generally good, although it would [also] be good if there is more differentiation between the four levels presented.” Scorer 5 suggested revisiting the process to calibrate the competencies with the descriptions on the rubric: “the rubric used may need to be re-evaluated to properly assess the projects presented.”

Scorer 5 also suggested that some phrases used in the attainment levels’ descriptions be explained further. The focus group students used phrases that could be considered “insider’s phrases.” These phrases could be misunderstood by those unfamiliar with the context. The example she gave was in Competency 2, Exemplary level: “[the student] designs projects from the ‘bottom’ and not from the ‘top’.” Among IU students or educators, this sentence might not need further clarification, but it could be confusing for a reader foreign to concepts of hierarchical work in intercultural settings.

Scorer 6

Profile: Mexican professor at an IU

Scorer 6 provided extremely thorough feedback on the process, discussing the interconnectedness/relationship/connection between the competencies and the students' projects. Scorer 6 wrote that this process "allows professors to know, identify, and clarify the learning stages or phases that students are developing in an integral manner, moment by moment, towards attaining the competencies established in the bachelor's program."

Moreover, Scorer 6 emphasized the adequacy of this kind of processes in a context where performance-based assessment is in place:

I believe that an integral evaluation through research projects is more beneficial for all those involved in the learning process. It is the evaluation itself that allows us to realize whether the project meets the set expectations, the competencies offered in the program of study.

Scorer 6 also pointed out that developing tools like the one resulting from this process helps students portray the complexity of their learning processes as they complete their DRs:

[This] reflects in a more complete fashion the connection with the competency that one wants to achieve, it generates learning as it links the student with the life in his/her community, making him more aware of his origins, his identity, the richness of the communities, enriching relationships and fostering culture. The projects students complete reflect the commitment, the way they deal with values, the ethics and practice of daily reality mixed with the innovation of students, and

above all, it is pleasant to recognize the spirit of solidarity and the effort to carry out projects for their community. Moreover, it is important that they handle methodology according to the [desired] attainment levels.

About the rubric itself, Scorer 6 noted that the labels of the attainment categories were not punitive: “The rubric is more appropriate and acceptable as it uses the terms ‘beginner’, ‘developing’, ‘satisfactory’, and ‘exemplary’ as opposed to the traditionally used ‘approved’ or ‘failed.’ The rubric is more viable and innovative to establish a grade for a student.”

Finally, Scorer 6 pointed out the benefits that implementing this process give to the community:

The community is aware of the ways of learning that go on at the school, what the school fosters in their children, which will therefore allow for more youth to pursue professional studies with their parents’ endorsement and the community’s as well. Moreover, this instrument [can] help communities to make their knowledge known to the world; the richness of their people, their products, and even its dissemination in a local, national, and international way.

Scorer 7

Profile: Mexican Professor at an IU

Although from the same institution as Scorer 6, the feedback from Scorer 7 differs markedly from Scorer 6’s feedback. Scorer 7 wrote that this process:

...is framed in a logic of interculturalism of Western dominion, where Western values are privileged over those of the communities that are to be served, for example: to highlight indigenous knowledge could be desirable for the Western

mestizo, but not for many ethnic groups of Mexico. The secret of their knowledge is part of their cultural resistance. On the other hand, no correlation is shown between the rubric and the competencies of the LGID; these latter ones comprise a higher level of complexity. An observation is that the rubric has a more technical ‘paperwork’ character to it.

During the second scoring round, Scorer 7 expressed again that this process “had a very strong Western load.”

The seven scorers that participated in this study submitted their feedback on the process to develop the instrument and the instrument, and their feedback was diverse. Some scorers’ feedback included words such as “useful” and “informative” in extensive and quite positive feedback, such as that provided by Scorer 6, who considered this process and the resulting instrument to be very useful and beneficial to students, educators, and the community. There was feedback on technical aspects of the rubric such as the inclusion of insiders’ phrases that needed to be explained, or the need to revisit the process to calibrate the competencies with the descriptions in the rubric. Two of the seven scorers expressed negative opinions, principally because they considered the process to be influenced by Western values or because the scorer was an outside evaluator, detached from the students’ processes.

The Researcher’s Analysis of the Instrument-Development Steps

Below are the proposed steps for the development of the instrument first introduced in Chapter 3, in Figure 2. Each step is presented individually, followed by the findings and the researcher’s analyses associated with that step.

Step 1: “During Focus Group (FG) 1: FG1 participants negotiate the definition of each IMD competency, and the definitions become the content of the rubric. Participants provide feedback on PBA at the IU and on this study.”

Findings and Further Analysis:

The first step entailed conducting a focus group with six IU senior students and a recent graduate. This focus group was very productive, although under time constraint. During this focus group the students agreed that their final research projects, or DRs, were appropriate tools to express their attainment of competencies, but they also raised concerns about the process of preparing and scoring their DRs, and not having an anchor of understanding such as a collective rubric. Their feedback on PBA and the DR process is available in this chapter.

Step 2: “The first draft of the rubric is produced. All material is submitted to FG1 participants for member-checking. The researcher organizes the data into a Table of Specifications.”

Findings and Further Analysis:

Transcribing and listening to the meeting’s material and creating a Table of Specifications were crucial activities for the researcher to better understand the data. Organizing data into a cleaner version of the rubric and submitting it for feedback and for member checking purposes was straightforward and productive.

Step 3: “During FG 2: Each FG2 participant receives a copy of the Table of Specifications and collectively provide estimates of content validity. Participants provide feedback on PBA at the IU and on this study.”

Findings and Further Analysis:

The analysis of this step is manifold. First, a discussion is presented on general observations about the day on which the second focus group took place. Second, a discussion on the ToS and the revisions to the rubric are presented.

On the day of the second focus group, when content validity estimates were obtained, the researcher had the opportunity to observe and take notes during a DR presentation session. During this session, seniors met with their methodology advisor and their academic advisor. The observations during that session became secondary data, the type of data that Teddlie and Tashakkori (2009) refer to as “unobtrusive measures” (p. 223). With this type of measure, the researcher has the opportunity to examine aspects of the phenomenon studied “without interfering with or changing it”, and “individuals under observation will not react to being observed” (p. 223).

During their session, students presented their DR’s preliminary findings, or projects almost completed, either individually or in pairs. All of the presentations the researcher observed were done using Power Point presentations. During this observation session, the researcher noted that the feedback given to participants ranged from being based on the visual appeal of the Power Point Presentation to methodological revisions. Feedback was given by the students’ methodology advisor and their academic advisor. The other students in the room did not give feedback. The content of the DRs presented that the researcher observed was of remarkable quality and sophistication. For example, one was a manual to manage a deer farm. The students justified the need for the manual, presenting the current feeding practices, which are wasteful and do not nourish the deer well, and the infrastructure, which presently allows for predators to easily get in and eat

the deer, or for small deer to escape. The students presented their management manual, proposing cost-efficient feeding practices and fencing alternatives, as well as community revenue projections of a well-managed deer farm.

Through these observations, the researcher confirmed that these seniors were going to be able to share their experiences about the process of their almost finished DRs. During this observation session, the researcher also considered the possible usefulness of a rubric, made by students, and distributed among their team members and their mentors. These aspects will be discussed in the final chapter of this study.

Following the DR presentation session, the second focus group took place. During the second focus group, the researcher observed two of the students had difficulty using the ToS, constantly opting to give the researcher oral feedback without filling out the ToS. The researcher explained the purpose of a ToS, the concept of content validity, and the students' role as expert judges. The students were unfamiliar with these concepts, but their oral performance was comparable to the other participants. Once the feedback was included, the rubric's levels were accepted or revised and accepted. This could have been a reason why the two students were hesitant when filling out their ToS. For this reason, the researcher confirmed the ToS results orally throughout the entire process. Additionally, similar to the first focus group's experience, the researcher was forced to conduct the second focus group under a time constraint.

Step 4: "The researcher gathers estimates of content validity via the Table of Specifications. If 80% agreement on the rubric's content is achieved, the item stays and is included in the rubric; otherwise, the researcher seeks FG2 participants' feedback, during

the focus group, revises the items, and submits to FG2 participants for member-checking. Researcher then calculates Lawshe's Content Validity Ratio.”

Findings and Further Analysis:

Students were not very familiar with the concept of content validity, and had never used a ToS before. The researcher explained to the students their role as expert judges and the purpose of the rubric and the ToS in very accessible terms, and she also accommodated for the students’ unfamiliarity with the concepts of content validity and a ToS. She accomplished this by discussing every level of the rubric orally and agreeing to guarantee understanding and consensus on the revised rubric, and therefore agreement on the content validity estimates. She recorded feedback and consensus in her ToS, and played the focus group recording twice to confirm accuracy.

Findings during Step 4 suggested two possible future scenarios: (a) students acting as expert judges should be involved in a separate training session to practice using ToS for content validity estimates, or (b) the ToS should not be used and the researcher should encourage students to debate, deliberate, decide, and reach a consensus, while recording this feedback herself. Findings during Step 4 also confirm the adequacy of the focus group format to allow students to gather and discuss the content of the rubric, as opposed to contacting them individually. The students displayed very productive group work and interaction.

Step 5: “Once the adequacy of each item on the Table of Specifications was agreed upon (80%), the researcher organizes students' responses to create the final draft of the rubric after member-checking. Scorers are contacted and receive DRs to score using the rubric.”

Findings and Further Analysis:

The researcher took notes on the students' deliberations and sought oral consensus at all times in addition to what was recorded on the students' ToS forms. Later, the researcher listened to the transcripts to corroborate the accuracy of the final version of each description, confirming her records. The findings from Step 5, like those from Step 4, attest to the suitability of the focus group format for seniors to deliver their judgments on the rubric content.

The researcher's decisions regarding the selection of scorers was based on her qualitative judgments on the richness of having multiple perspectives. She believed strongly in the importance of having scorers from diverse backgrounds: different national contexts, academic affiliations, and philosophical views.

Step 6: "First scoring round of inter- and intra-rater reliability estimates takes place. Scorers provide feedback on the process."

Findings and Further Analysis:

Even though scorers were in various places geographically, this step yielded productive results in part because the researcher composed a letter that clearly outlined the requirements of their participation and the scorers were helpful and cooperative. As seasoned educators, the scorers were able to complete their scoring without incident. There were two scorers with critical views on the process this study proposed. The researcher recorded the scores and organized the scorers' feedback on the process proposed in this dissertation, allowing her to explore different opinions. Findings on this step support the feasibility of having scorers from diverse backgrounds participate in piloting an IU rubric.

Step 7: “Second scoring round of inter- and intra-rater reliability estimates takes place. Percentages of agreement are calculated. Scorers provide feedback on the process.”

Similarly to Step 6, this step yielded productive results because scorers were helpful and cooperative, and because as seasoned educators, they were able to complete their scoring without incident.

Findings and Further Analysis:

Feedback from scorers unveiled several aspects of rubric usage and development. Without these qualitative components, “findings” would only help the researcher infer mathematically the adequacy of the estimates obtained and no information on usefulness for daily use would have emerged.

Summary

This chapter was devoted to results and analysis of data. The data collection process followed sequential mixed methods, as discussed in Figure 1, and the steps to develop the instrument, as outlined in Figure 2; both Figures are available in Chapter 3. The results and data analysis were presented in six sections. The first section of the data analysis was devoted to the students’ evidence of attainment of competencies, obtained during the two focus groups. This evidence became the content of the instrument developed. Using open coding, the researcher organized the data into seven categories. These seven categories reflect the complex set of acquired, potential, and/or desired performances or achievements that pertain to professional, intellectual, and personal areas of students’ life, and reflect what students value and aspire to, or want to possess in their personal, educational, and professional journey.

The second section was devoted to the estimates of content validity of the rubric that participants developed. Content validity estimates were obtained by using a Table of Specifications and calculating Lawshe's CVR, and substantiated by evidence resulting from these two content validity strategies. The resulting rubric's content validity estimates reached 80% or higher, sufficient to meet acceptable standards (Newman et al., 2011).

The third section was devoted to focus group participants' feedback on PBA and the rubric-development process. Participants welcomed the instrument-development process, and shared some of their concerns about organization and fairness in grading when developing and submitting their DRs. Focus group participants also discussed their desire to have the freedom to present their DRs in creative and diverse ways, having an organized and fair scoring process. Participants also mentioned peer evaluation possibilities.

There was an additional graduate who wanted to participate in the study but could not attend the focus groups. His feedback on his peers' rubric and the process in general was received via email. His feedback on the attainment of competencies was discussed. Although no changes were made to the rubric based on this feedback, his contributions were useful for articulating implications for practice, discussed in Chapter 5.

The fourth section presented the estimates of inter- and intra-rater score reliability. Seven scorers in Mexico and the US used an archive of DRs and the participants-developed rubric to estimate the instrument's reliability. Seven out of ten DRs yielded an inter-rater score agreement of 80% or greater, which served as the evidence to indicate that the instrument's inter-rater score reliability fulfilled acceptable

standards (Newman et al., 2011b). For intra-rater reliability score estimates, DR10 was used to estimate this. All scorers attained 80% or higher scoring agreement, which is consistent with acceptable standards for reliability (Newman et al., 2011b).

The fifth section included scorers' feedback on the rubric-development process. During and after the scoring sessions, scorers provided feedback on the instrument-development process and the instrument. The scorers' feedback ranged from positive and supportive, to technical calibration suggestions, to critical. The positive feedback included mention of the process being helpful and enriching for students, educators, and communities. The technical feedback included suggestions such as aligning the competencies and the descriptions better and using less local terms or phrases. The critical feedback came from two scorers that hold philosophical views that consider this process Western-value laden, in that it makes an outsider evaluate a process from which they are detached. Although these two scorers had expressed critical opinions, their intra-rater score reliability was shown to be stable.

Finally, the sixth section of this chapter included the researcher's reflection on the entire instrument-development process. She presented each step of the process and provided her analysis of each phase. Her reflections helped her to articulate suggestions for further research, available in Chapter 5.

Chapter V

DISCUSSION AND RECOMMENDATIONS

This chapter is devoted to an in-depth analysis and a re-examination of the data to arrive at a meta-interpretation. Mixed methods research allows for this re-examination (Teddlie & Tashakkori, 2010) for a more comprehensive discussion of results. The chapter opens with a summary of the study and the answer to the research questions, followed by an analysis of results. This chapter also includes a discussion on the limitations of the study, the implications for practice, recommendations for future research, and conclusions.

Summary of the Study

This study took place at one of the intercultural universities of Mexico. These are institutions of higher education that have been created by the Ministry of Education in conjunction indigenous leaders. The IUs focuses primarily on serving indigenous students from different ethnic groups in the country. The IUs' existence and operations are invariably challenging (Fábregas, 2008; Schmelkes, 2009). The IUs' assessment strategies and self-assessment pointers, which might pave the way to meeting many of their quality educational needs or challenges, were missing. A performance-based instrument with acceptable standards of validity and reliability estimates to assess students' attainment of competencies was not found at the writing of this study. Additionally, research on the topic of performance-based assessment itself was scant.

The purpose of this study was to develop a process to create a performance-based instrument, a rubric, with acceptable standards of validity and reliability estimates to assess students' attainment of competencies outlined in one of the IU's undergraduate

degrees, the Intercultural Management for Development. Using sequential mixed methods, this instrument was created with the participation of IU senior students, a graduate, and the researcher, and piloted by seven Spanish-speaking professors in higher education in the United States and Mexico.

This study aimed to answer two research questions:

- (a) According to IU seniors, what evidence demonstrates the attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?
- (b) Can acceptable standards of validity and reliability estimates be established for an instrument developed to measure students' attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?

The IMD program is guided by five competencies that students should develop and attain upon graduation. These five competencies indicate that students should be able to: (a) articulate regional/indigenous knowledge and initiatives, (b) facilitate advocacy for and management of resources and information, (c) strengthen and design intercultural initiatives, (d) highlight regional/indigenous knowledge and initiatives, and (e) generate diagnostic and proposal-oriented knowledge.

The researcher based the creation of the performance-based assessment instrument on these five competencies.

This study used sequential mixed methods (Creswell, 2003; Tashakkori & Teddlie, 2009), where qualitative and quantitative phases in the study take place chronologically: one stage takes place after another. A meta-interpretation is inferred at

the end based on the findings (Newman et al., 2011b). The human capabilities approach (HCA) was the conceptual framework guiding this study. Education articulated through the HCA is empowering and transformative (Unterhalter & Brighouse, 2007), and allows us to carry out assessment based on students' input and negotiation (Walker, 2008).

It was established through the review of literature that the instrument developed in this study would be an analytic, general rubric. Rubrics are the most common tool when engaging in performance-based assessment (Jonsson & Svingby, 2007; Moskal, 2000; Reddy & Andrade, 2009). Several studies suggested the invariability of the benefits of rubrics as performance-based assessments (Jonsson & Svingby, 2007; Moskal, 2000; Reddy & Andrade, 2009). Among the benefits were the promotion of learning and achieving among students and better ways to give and receive timely and thorough feedback. Rubrics have potential benefits for teaching and program evaluation (e.g., Dunbar, Brooks, & Kubicka-Miller, 2006; Knight, 2006; Oakleaf, 2006).

The rubric was created with the purpose of assessing students' levels of attainment of competencies outlined in the IMD major. In brief, the procedure for the creation of the rubric, as outlined in Figure 2 available in Chapter 3, was as follows: first, a group of senior students and a graduate from the IU met during a focus group to develop, upon democratic deliberation, the initial content of the rubric. Participants used the IMD major's five competencies, to be attained upon graduation, as the basis for their discussion. During the focus group, participants also gave feedback to the researcher on PBA and PBA at the IU.

The second step included member-checking measures. Approximately a month later, participants from the first focus group received a copy of the rubric they produced

during the focus group, as well as the transcripts of the focus group's recording for member-checking. Participants approved the accuracy of the rubric and the transcripts. The second step also consisted of the creation of a Table of Specifications (ToS), a mixed-methods tool to estimate content validity (Newman, Lim, & Pineda, 2011) and the organization of a second focus group with different IU seniors, which took place a month later.

The third step consisted of conducting a second focus group where participants determined content validity estimates and provided feedback on performance-based assessment and on this study. The second focus group consisted of a different group of five senior students from the IU who met to estimate content validity as expert judges using the ToS as well as the rubric that their peers from the first focus group had developed. Specifically, this group provided an estimate of agreement of the appropriateness of the evidence used to reflect the attainment of the five competencies, and also gave feedback on PBA at the IU. Approximately a month later, the five seniors received the material produced during the second focus group for member-checking and they approved the material. The final draft of the rubric was a product of the second focus group. During the fourth step, the researcher completed the calculations to determine whether the content validity estimates reached acceptable standards of 80% agreement or higher.

During the fifth step, the final version of the rubric was created. Seven Spanish-speaking professors of education in Mexico and the United States piloted the rubric using nine final research projects (DRs). The DRs used in this study were obtained from an

archive available via the IU website and one DR that the author shared directly with the researcher. All identifiable information on the authors or the institution was removed.

The sixth step consisted of conducting the first round of reliability estimates with scorers, who also provided feedback on performance-based assessment and on this study. The scorers' collective responses were used to estimate the reliability of the rubric in two ways: inter-rater and intra-rater score reliability. The final step entailed completing the scoring sessions to obtain reliability estimates. Inter-rater score reliability was estimated to find the percentage of score agreement between scorers, and the intra-rater score reliability was estimated to check, with a minimum of two weeks in between, the stability of scorers' judgments over time. Results from both reliability measures yielded acceptable standards of 80% or higher.

This study aimed to contribute to the quality of the assessment of student learning by providing a process to develop an instrument with acceptable content validity and reliability estimates that can be used across the IU network. Part of the intent of this study was to provide a more flexible, constructivist, and participatory process to assess student learning in the form of performance-based assessment, which is well-documented to be a useful alternative to standardized testing (Anderson, 1998; Moskal, 2000). It was hoped that the process and the resulting instrument would allow IU students to demonstrate their attainment of competencies in a performance-based and participatory manner.

Answer to Research Questions

Question #1: According to IU seniors, what evidence demonstrates the attainment of the competencies of the IMD academic program of study at an intercultural university in Mexico?

Performance-based assessment is the IU's method to assess students' learning (Casillas & Santini, 2006). For this study, students outlined the evidence of attainment of competencies in their own words, using their own conceptualizations and interpretations. The practice of empowering students and giving them the freedom to express the education they value and desire to receive is characteristic of the human capabilities approach (Nussbaum, 1997; Walker, 2008). This approach advocates that people state and pursue the life they value (Sen, 1999).

According to participating IU seniors and a graduate, the evidence that demonstrates the attainment of competencies outlined in the IMD major is a complex set of acquired, potential, and/or desired performances or achievements related to the academic, professional, intellectual, and personal arenas of students' life. These performances or achievements reflect that students possess or developed the *know-how*'s, and the *know-what*'s outlined in their major's competencies, but also the potential achievements that will come in the course of the person's development.

The evidence of competency attainment also reflects what students value and seek to gain from the process of their education and degree completion. At the time in which the two focus groups took place and the evidence of attainment was discussed, the students might or might not have fully attained all competencies. Nevertheless, students were able to collectively envision what the evidence of having attained their major's

competencies looked like. Moreover, this evidence of attainment of competencies can be reflected in the final research projects or DRs that students produce as a graduation requirement. Once coded into levels of attainment, each of the acquired, potential, and/or desired performances or achievements were used to develop a rubric (Appendix J). For example, Competency 3 indicates that “The student will be able to strengthen and design intercultural initiatives.” The students outlined what the evidence of having attained that competency looked like, in a spectrum of attainment from beginner level to exemplary level, regardless of where in that spectrum the participating students were. Some participants might be able to design small-scale intercultural initiatives for development at the time of the focus groups, but at the time of the completion of their DRs, they might be able to implement robust and sound intercultural projects in their communities.

Using open coding, the researcher organized and analyzed the evidence of attainment of competencies outlined by the students. Seven categories emerged. These categories, discussed in Chapter 4, are complex set of acquired, potential, and/or desired performances or achievements related to the academic, professional, intellectual, and personal arenas of students’ life, that represent the evidence of attainment of competencies that the students either already possess, or value aspire to attain. These categories were obtained through open coding. Open coding is a strategy that allows the researcher to analyze qualitative data through categorizing and describing phenomena, typically using conceptual labels (Strauss & Corbin, 1990). In this case, the “phenomena” were the actual or potential valued performances found in the student-developed rubric.

The list of categories that emerged is not exhaustive, and there is overlap among them, but they provided a clear language to refer to the attainment of competencies, either achieved or valued, by the IU students. The categories were (a) collaboration and outreach, (b) analytical and critical thinking, (c) acquisition and appreciation of knowledge, (d) planning and implementation skills, (e) creativity and resourcefulness, (f) empowerment and autonomy, (g) empathy, solidarity and advocacy.

The articulation of the IMD competencies in the rubric can be considered functional capabilities, as defined by Walker (2008), and can be portrayed in the students' final research projects. Functional capabilities in education are both achievements and opportunities (Walker, 2008). Capabilities are potential functionings, and functionings are actual doings, or realizations (Sen, 1999; Walker, 2008). Both capabilities and functionings are important to discuss in assessment of higher education (Walker, 2008), and the notion of functional capabilities captures the importance of both. As Walker and Unterhalter (2007) argue, assessment entails a dynamic relationship between capabilities and functionings.

Question #2: Can acceptable standards of validity and reliability estimates be established for an instrument developed to measure students' attainment of competencies of the IMD academic program of study at an intercultural university in Mexico?

There is ample evidence that the instrument resulting from this process, an analytic general rubric, yielded acceptable standards of validity and reliability estimates. The evidence is based on the content validity and inter- and intra-rater reliability estimates, presented in Chapter 4. In brief, estimates of content validity were obtained

through a Table of Specifications (ToS; Appendix E), with a group of IU seniors serving as expert judges. The ToS was created based on the rubric created by the first group of IU seniors and a graduate.

Once acceptable estimates of content validity were obtained from the second focus group participants, the researcher incorporated their feedback, prepared the final draft of the rubric, and submitted to the second focus group participants for member-checking. The final draft of the rubric was then used by seven scorers. The scorers piloted the rubric scoring students' final research projects. The percentage of agreement on the content validity and the reliability estimates met the necessary 80%. These estimates are consistent with the literature on recommended percentages of agreement. According to Newman et al. (2011b), inter-rater and intra-rater score agreement of at least 80% is sufficient to meet the minimum reliability requirement.

Analysis of Results

Validity and Reliability

The estimates of content validity of the rubric developed in this study were high. These estimates were obtained through the use of a ToS with expert judges and the calculation of Lawshe's CVR, which yielded a value of one, representing total agreement among judges. Additionally, face validity emerged as a result of estimating content validity, as participants gave feedback on the instrument's development process. Therefore, although at a small scale, this process had content and face validity for participating students and graduates, and some professors who served as scorers, from Mexico and the United States, with the exception of two scorers, as discussed later.

One of the two scorers with critical views on this study's process is a Mexican professor at one of the IUs and the other one is a non-Mexican professor in the United States. The scorer who is a faculty member at one of the IUs said that this process "had very strong Western load" (S7), that the rubric-development process "is framed in a logic of interculturalism of Western dominion," and that the rubric "has a more technical 'paperwork' character to it." The scorer who is a non-Mexican citizen and a faculty member at an American University said that "the problem is in the idea of performance-based assessment that assumes that an external observer, foreign to the context, can evaluate... seeing only the final product." The two scorers' perceptions, however, are not consistent with what students expressed during the focus groups. This process was welcomed among students, as discussed below.

A possible interpretation of the two scorers' feedback is that they perceived this process as a standardization of student assessment, an imposition of Western values by an outsider. Their opinion resonates with what Carney (2003) argues, that an uncritical emphasis on cognitive achievement without a consideration of what local values might define as "good education" results in "Westernization." The researcher agrees philosophically to a certain degree with the second scorer (see Pineda, 2010); however, the purpose of his participation was to explore the stability of the rubric, not to provide a direct statement on the students' work. The act of attempting to better organize an institution's assessment practices, placing students at the center of the process, should not translate into an oppressive practice, but an opportunity for empowerment and improvement. Having a disorganized process of assessment, therefore giving room to

disparities could be, on the other hand, a form of injustice. For example, as a student expressed regarding the lack of organization:

... you start developing your DR and it is a mess because, well, your work does not have like a defined structure starting from your first semester, or second, or third... every DR is floating [or wandering] around because it does not have like a clear structure of what the research was about and of course this creates a conflict in you because although you know how you have developed your project, you don't know how to transcribe all that into your DR.

Assessment should provide opportunities for schools to engage in “self-reflection, self-critique, self-correction, and self-renewal,” (Darling-Hammond , 1994, p. 5). The lack of instruments with acceptable qualitative or quantitative estimates of validity and reliability and a lack of active collaboration among stakeholders in the assessment process can hinder those opportunities. This study attempted to offer a process for the development of a performance-based assessment instrument. It is hoped that institutions could replicate and adapt this process to their organizational and/or pedagogical needs, should they consider this instrument-development process valuable.

The Process and the Students

Students welcomed the process in several ways. They expressed the need for a more organized, participatory, and fair assessment experience, pointing out that a previous attempt to standardize attainment descriptions in which students did not have input, failed. Education articulated through the HCA is empowering (Unterhalter & Brighouse, 2007). The process proposed in this study was based on the students' feedback only, either as originators of the description of attainment during focus group

#1, or as expert judges, giving feedback on those descriptions during focus group #2. By doing so, students were given the opportunity to determine what constitutes a credible evidential basis (Messick, 1994) of the attainment of competencies.

This study's findings are similar to what Caro (2010) found in her research on performance-based assessment among undergraduates at another IU, described in Chapter 3. Some of Caro's participants considered that the assessment process lacked clarity because there were no uniform criteria. They pointed out that their good communication skills and confidence when defending their thesis seemed to help more than their actual content knowledge. Only few students considered that the assessment process at their IU was tailored towards understanding the learning process, in knowing the "how" as much as the "what."

Another study in rubric development in higher education in which the results corroborate this dissertation's results was Mansilla, Duraisingh, Wolfe, and Haynes' (2009). Among their study's strengths was their systematic process for validation of their rubric, strengthened even more by the large number of students' projects scored. Nonetheless, among their weaknesses were the lack of explanations of procedure for their reliability estimates and the lack of student input. The procedure to develop the rubric at the IUs did not have a large sample like the Mansilla et al.'s study, but there was an attempt to include students and education experts in two different countries, and to provide a detailed narrative of the validity and reliability estimate procedures.

The Process and the Scorers

There is also supportive evidence that the process used in this study to develop an instrument can produce a rubric that allows for stable scoring independent of

national/cultural/institutional settings. When piloted by seven scorers from different academic and national backgrounds, five of them considered the rubric positively through their feedback, and their scoring suggested acceptable consistency.

The rubric remained stable with and without the score of the two scorers who had critical views on this process. With their scores, agreement decreased from 100% to 80%, yielding in both instances high inter-rater score agreement. Additionally, their intra-rater reliability score agreements were 100%. This evidence supports that the instrument developed yielded acceptable standards of reliability estimates, independent of two scorers' critical opinions of this process. Moore and Young (1997) argue that one of the major sources of unreliability in performance-based assessments is inter-rater score disagreement. In this dissertation, inter-rater score reliability estimates were acceptable, and agreement was reached despite two scorers' critical opinions.

The Five IMD Competencies, the Evidence of Attainment of Competencies, and the Human Capability Approach

This section is devoted to a discussion of the role of the five IMD competencies, the evidence of their attainment, and the connection these competencies and their attainment have with the human capabilities approach.

Two main concepts of the human capabilities approach are functionings and capabilities. Functionings are achieved outcomes, and capabilities are opportunities, or potential functionings (Sen 1992, 1999; Walker & Unterhalter, 2007; Walker, 2008). The competencies, as the main goals of the IMD major, can be conceptualized as functionings. Although some students might not have attained all competencies yet, the competencies represent the achieved and achievable outcomes that students aspire to

attain from their years of enrollment at the IU.

The five IMD competencies served as the springboard for the detailed description and discussion of the functional capabilities (Walker, 2008) that represent the evidence of attainment of those competencies. Students received the list of the IMD competencies obtained from the IU website, and used this list to guide their discussions and the creation of the rubric. Within a HCA framework, educational assessment “is concerned with a dynamic relationship between opportunity and outcome, or capabilities and functionings” (Walker & Unterhalter, 2007, p. 4). As Terzi (2007) argues, in the HCA, “the use of [the words] ‘being able to’ in expressing capabilities [and also in the wording of the competencies] implies... the opportunity and possibility entailed by the concept of capability, rather than simply the common understanding of ‘to be able to’ in terms of ability” (p. 37). These functionings –the five IMD competencies- are related to the actual learning process, opening possibilities for other functionings and capabilities (Terzi, 2007).

Limitations of the Study

This study had three main limitations. These limitations were (a) the process’ usefulness and structure, (b) the study’s sampling, and (c) the researcher’s limited time spent at the IU. Each of these limitations is explained below.

Mixed methods research allows for the emergence of a greater assortment of divergent views (Teddle & Tashakkori, 2009). This study provided the opportunity to analyze the significance of the IMD rubric development process in qualitative as well as quantitative terms. Although the findings suggest that the rubric developed with this process yielded acceptable standards estimates of validity and reliability, the sources of

evidence for the actual process' usefulness were limited. The sources of evidence in this study were predominantly the IU participating students' and scorers' testimonials and feedback on the process and the rubric. Increasing the sources of evidence when building performance-based assessment instruments results in stronger claims about meaningfulness and usefulness (Messick, 1996). Some examples of other sources of evidence that can be used to determine the usefulness of this process at the IU are (a) interviews with tutors and professors, (b) more focus groups with IU students of all semesters, (c) an opinion tool/survey about this process, and (4) interviews with education officials.

The human capabilities approach highlights the empowering purpose of education and assessment (Walker, 2008; Walker & Unterhalter, 2007), especially pertaining to students. Students were at the center of this study and they guided the development of the IMD rubric. Keeping students at the center of the instrument development process does not mean the exclusion of other stakeholders like elders, parents, and other community members, possible future employers, other school officials, and more IU educators.

In terms of usefulness, this study is limited in that the resulting rubric does not explicitly provide score meaning and does not differentiate team scores from individual scores. Although inter- and intra-rater agreements were calculated and yielded acceptable standards, scorer agreement does not inform constituencies about score meaning (Messick, 1994). This study suggests that the process proposed herein yields stable reliability estimates, but it is limited in providing an explanation beyond scorers' numerical reliability estimates. To remediate this, diverse stakeholders need to engage in

a participative process (Walker, 2008) and determine score meaning on their own, should this study's process be established. Furthermore, the rubric that resulted from this study does not have the function to differentiate team effort scores from personal attainment of competencies scores. As the IU students expressed, projects at the IU are usually done in teams. Presently, this process focuses on attainment of competencies and the rubric provides a single final score.

Another limitation of this study was the lack of a large sample of anchor papers to share with scorers. Anchor papers are selected student papers that serve as examples of performances at the different attainment levels of the rubric, allowing scorers to compare levels of attainment as they score other papers (Moskal, 2003). Having anchor papers is a common practice when piloting rubrics and is beneficial for reliability purposes (Moskal, 2003).

Moreover, this study was limited by the students' unfamiliarity with developing or validating rubrics as well as their time availability. As an IU graduate expressed, discussing the attainment of competencies at the IU is not a common practice. Regarding time availability, students had agreed to two hours of participation. Covering all the material with time constraints during the focus groups was challenging. If not for time constraints, a more thorough content validity analysis could have emerged.

Finally, this study was further limited by short field engagement. The researcher was able to visit the students only twice and no more than a week each time. Prolonged engagement in the field would have helped her to observe DR procedures longer, interact with more stakeholders, and gain deeper understanding of performance-based assessment at the IU.

Implications for Practice

The results of this study generated practice implications for student assessment in higher education. These implications were derived on the basis of their consistency with the study's results and with theory of and research on formative assessment. The implications were also consistent with human capabilities approach literature.

Conduct Performance-based Assessment with a Formative Assessment Framework

The process that this study used yielded confirmatory data on the applicability of performance-based assessment in higher education. This is shown by this study to be particularly salient in programs where students have community-involvement and community-development appointments as part of their formation. This study also yielded data to substantiate the need for performance-based assessment with a formative assessment framework. Formative assessment is assessment for learning (Bloom, Madaus, & Hastings, 1981), over the course of a program or academic year.

Performance-based assessment is an alternative assessment approach (Darling-Hammond, 1994). This approach is a more flexible, constructivist, and participatory approach for assessing student learning (Anderson, 1998; Moskal, 2000), promoting active learning, and facilitating accurate judgment of complex competencies (Jonsson & Svingby, 2007; Messick, 1994; Stevens & Levi, 2005). For educators and students that desire an assessment model where power and control are not unilateral, but negotiated and shared democratically (Anderson, 1998), performance-based assessment is a comprehensive option.

With performance-based assessments, students have the ability to be active stakeholders in the assessment of their learning and in exercising self-assessment, both

individually and as a group (Anderson, 1998; Stevens & Levi, 2005). Moreover, performance-based assessment allows students and educators to move away from a model of educational accountability based on high-stakes testing, and perpetuated by bureaucratic systems (Mathison & Ross, 2008).

Performance-based assessment, as with any other type of assessment, does not succeed in a vacuum. It requires that educators (a) become competent masters of the required educational standards, (b) make attainment targets clear, and (c) use assessment data over time in collaboration with students to scaffold and motivate students' learning (Stiggins, 2007). These are fundamental characteristics of formative assessment (Bloom, Madaus, & Hastings, 1981; McMillan, 2007; Stiggins, 2007).

The results of this study demonstrated in numerous ways the importance of conducting formative assessment. Students expressed the need for an organized process from beginning to end to complete their performance-based assessment sole instrument, their DR. Students also mentioned that professors and students should communicate effectively regarding expectations and attainment levels. As an IU scorer and a graduate both noted, discussing attainment of competencies clearly is crucial for a successful assessment experience. Formative assessment using the established performance-based assessment instruments is beneficial in contexts such as the one described above.

If performance-based assessment outlines what students do, formative assessment outlines what educators and students do together as a team, over time. Formative assessment entails "using student achievement data to inform instruction" (McMillan, 2007, p. 1), ideally blending assessment and instructional functions effectively (Brookhart, 2007). Formative assessment has ramifications for students' motivation,

metacognition, and peer- and self-assessment (Bloom, Madaus, & Hastings, 1981; Brookhart, 2007; Stiggins, 2007).

Formative assessment is assessment for learning (Bloom, Madaus, & Hastings, 1981). Quality assessment is not merely a function of validity of scores or results. Quality is judged in terms of the impact of those results on student learning (Stiggins, 2007). Bloom argued that assessment “in relation to the process of learning and teaching can have strong positive effects on the actual learning of pupils as well as on their motivation for the learning and... [their] self-concept in relation to school learning” (Bloom, 1969, p. 50, as cited in Wiliam & Leahy, 2007). Formative assessment’s benefits include improving and increasing student learning and ownership of the learning process, as well as metacognition, motivation, and self-regulation (Bloom et al., 1971, as cited in Dann, 2002; Bloom, Madaus, & Hastings, 1981; Brookhart, 2007; McMillan, 2007).

Use Student Assessment Data to Yield Formative Data

A characteristic of formative assessment is that student assessment data are used by both educators and students to inform learning and teaching (Bloom, Madaus, & Hastings, 1981; Brookhart, 2007; Guskey, 2007; McMillan, 2007). The process presented in this study has the potential to yield formative data. Throughout this study students, educators, and graduates provided feedback on the process to develop the rubric. Some participants were supportive and others objected philosophically to the process.

Not all educators hold educational assessment values that empower and include students, provide transparency and fairness, and inform teaching for improvement. Some

educators oppose “measuring” or quantifying learning; nevertheless, the process presented in this study does have formative assessment potential for those that consider assessment as a vehicle for learning. If used as a learning exercise, this study’s process could be used to guide class discussions, and clarify expectations generated from these discussions. This practice is consistent with literature on formative assessment. Bloom, Madaus, and Hastings (1981) argue that educators need to make objectives explicit, apply them consistently, and provide feedback throughout the process. As a result, the assessment process can substantially improve student learning (Bloom, Madaus, & Hastings, 1981).

Moreover, this participatory and empowering assessment practice is also consistent with one of the principles of the HCA, which is that individuals should have the freedom to determine what they value (Sen, 1999). Students’ capacity for critical examination of themselves and their traditions which in this case are related to assessment and educational goal-setting, are essential to develop education for freedom (Nussbaum, 2006). Freedom is a pillar of the HCA (Sen, 1999) because it is a principal determinant of individual initiative and social effectiveness. Using the process of this study, through group deliberations students and educators will have the opportunity to collectively engage in democratic dialogue to determine educational expectations for the attainment of competencies. Formative assessment is also consistent with HCA principles because it highlights the process of creation and deliberation, and not just a product, as a utilitarian approach to education would.

Include Students as Stakeholders when determining Assessment Content and Procedures

Another implication for practice that emerged from this study is the importance of including students when determining assessment content and procedures. This study, although at a small scale, provided evidence that students are able to create their own instrument that is stable over time when used by scorers. Moreover, because their program involves constant teamwork, IU students expressed interest in exploring this assessment process collectively and implementing it with peers. These students' perspectives point to the feasibility and desirability of having students' proactive participation in determining assessment content and policy at their institution.

Nicol and McFarlane-Dick's (2007) work on formative assessment and self-regulated learning suggests that when "students are already assessing their own work and generating their own feedback, higher education should build on this ability" (p. 199). Bloom, Madaus, and Hastings (1981) highlight benefits of the use of formative assessment by students: "[t]he most important value of formative evaluation, in our view, is the aid it can give students in learning the subject matter and behaviors" (p. 167). In addition to self-regulation, peer-assessment is particularly useful in contexts such as the one presented in this study. Bloom, Madaus, and Hastings (1981) suggest that the most effective procedure for peer-review is to have small groups of students meet to review their formative assessment material as they help each other overcome difficulties.

In the HCA, students exercise choice over their own valued capabilities and functionings, or beings and doings (Nussbaum, 1997; Unterhalter, Vaughan, & Walker, 2007; Walker, 2008). Framing an institution's assessment practices with capabilities and

functionings allows educators and students to go beyond instrumental values of education: thinking beyond securing a job (Walker, 2008). Evaluating educational effectiveness based on resources and inputs alone is a limited view. The focus should be on whether students are able to turn resources into capabilities and potential functionings (Unterhalter, Vaughan, & Walker, 2007).

Use Assessment Processes to inform and include Constituencies

Accessing government funding and ensuring degree accreditation (Schmelkes, 2009) is a challenge for institutions such as the IUs. Increased and more effective communication with various stakeholders would be beneficial in addressing this challenge. The process this study proposed has the potential to be useful in making public to educators and other stakeholders the goals and academic achievements of institutions.

As one scorer noted, as students immerse themselves in the community while completing their DRs, the community grows in awareness of what is being taught at the IU. Consequently, this awareness develops community members' trust in the institution. The resulting rubric embodies the desirable evidence of competency attainment which, according to an IU graduate, is seldom discussed in detail.

The IUs are characterized by their stated goal to include the communities around them (Casillas & Santini, 2006) as active stakeholders. The process proposed in this study can be used to inform multiple audiences and stakeholders. This inclusion could include inviting the community for public consultation and scrutiny of the attainment of competencies at the school. Inclusive activities could include holding open-houses, consulting sessions with community elders, and organizing town hall meetings. Just as

Bloom argued that formative assessment provides opportunities for students to show improvement and demonstrate success (Guskey, 2007), through making the process of attainment of competencies public and transparent, institutions would have a similar opportunity.

Recommendations for Future Research

Based upon the results of this study, there are three formal research recommendations. These are (a) to introduce this study's instrument-development process to other IUs, (b) to conduct parallel mixed design studies to further explore assessment using the HCA as a framework, and (c) to conduct a case study with in-depth interviews to explore assessment as a concept in intercultural settings.

Introduce this study's Instrument-development Process to other IUs

There is ample evidence resulting from this study to suggest the relevance of introducing this study's instrument-development process to other IUs. This instrument entailed conducting focus groups with students using a program of study's competencies to initiate discussions and create a rubric, using mixed methods to collect and analyze data, and obtaining estimates of validity and reliability with scorers using the resulting rubric.

This study's instrument-development process will allow institutions to establish collaborative processes for assessment procedures with and among faculty, students, and as many other stakeholder as possible. Using this process, rubrics with acceptable reliability and validity estimates and that reflect the desired competencies can be developed. Having more stakeholders' involvement can help strengthen, with

participatory democratic dialogue (Sen, 1999), the content representativeness of the construct measure (Messick, 1994), or content validity, and its usefulness.

The IUs, in their stated desire to equip young leaders to collaborate and create fruitful initiatives (Casillas & Santini, 2006), are fertile soil for a process such as the one proposed in this study. Moreover, peer-evaluations should be added to the instrument-development process, given that many participants expressed the potential usefulness of such an exercise.

Conduct Parallel Mixed Design Studies that Explore Assessment Using HCA

The second recommendation entails conducting parallel mixed design studies exploring assessment using a HCA framework. HCA fosters -among many aspects of human and community development- democratic dialogue, community participation, social arrangements conducive to justice, and individual empowerment (Sen, 1999; Walker & Unterhalter, 2007; Walker, 2008). However, there is an underexplored potential between education and the HCA (Saito, 2003) in general, and the HCA and educational assessment in particular (Walker, 2008).

To explore assessment using the HCA, parallel mixed design studies should be conducted, where at least two parallel and relatively independent strands, one qualitative and one quantitative, are established. This design allows for investigation of exploratory and confirmatory questions simultaneously. For example, the qualitative strand could include interviews with educators and focus groups with students, exploring their views on the HCA and its potential as an assessment framework. The quantitative strand could be the implementation of an opinion survey within the institution, and the piloting of previously designed HCA-based assessment tools. These tools could be rubrics,

observation lists, or inventories. Different teams of researchers may be required to complete parallel mixed design studies (Teddlie & Tashakkori, 2009).

Conduct a Case Study with In-depth Interviews to Explore Assessment as a Concept in Intercultural Settings

Evidence resulting from this study suggests that the concept of “assessment” might be problematic in some contexts, particularly when considered as an accountability measure and considered as a byproduct of exercise of power. These contexts, as in the case of this study, might be intercultural contexts. In these intercultural contexts, where there is a conscious and purposeful differentiation of actors, with full disclosure of ethnic, political, sociocultural, epistemological, and identities, structures of power or perceptions thereof are invariably challenged. Conceptually, assessment might be equivalent to imposed hierarchical structures, and therefore be approached cautiously and skeptically.

Notions of political power associated with testing and accountability cannot be disassociated (Apple, 2000; Gabbard, 2000). Notions of individual selection, social control, and hierarchical systems of accountability (Mathison & Ross, 2008) among non-mainstream culture groups might discourage assessment practices, and this discouragement is well founded (Schmelkes, 2009; Treviño, 2006). Top-down policies in intercultural settings hinder the full potential of meaningful learning (Dietz & Mateos, 2011). Nonetheless, consistent with Caro’s (2010) findings, evidence resulting from this study suggests that a lack of or having a disorganized structure to assess attainment of competencies is actually detrimental for students.

The process proposed in this study, and the IUs’ assessment processes in general advocate for students and educators to own these processes, and therefore not embracing

top-down policies. Further research should explore assessment as a concept in intercultural settings, exploring possibilities of collective agreements on the proper approach, level, and type of assessment that is welcomed. A case study that includes in-depth interviews conducted at one of the IUs or a similar institution is suggested.

Conclusions

The Intercultural Universities are higher education institutions in Mexico that embrace performance-based assessment and advocate for the empowerment of indigenous youth (Casillas & Santini, 2006). Nonetheless, a review of research suggested that a performance-based assessment instrument to assess the attainment of competencies was missing. The purpose of this study was to develop a process to create such an instrument, in which students are empowered in and through their educational assessment. The process proposed in this study included primarily students, who are at the center of the IUs' educational model. The human capabilities approach (HCA) was the theoretical framework guiding this study because of its emancipatory, participatory, and flexible approach to human development and assessment.

In the HCA, participatory dialogue and democratic deliberation are essential in determining people's well-being and the kind of life they value (Sen, 1999). Only through "openness of communication and arguments" (p. 152), can collective values be articulated. In Sen's (1999) words, "to express publicly what we value and to demand that attention be paid to it, we need free speech and democratic choice" (p. 152). In education in general, and educational assessment in particular, students' input is valued and negotiation should be at the forefront of decision-making (Walker, 2008). The human capabilities approach is on display at the IU, and evidence supports the feasibility

of having the IUs' performance-based assessment, articulated through functionings and capabilities.

This study revealed the feasibility of the creation of an instrument-development process in which students develop and validate a rubric, and their final research projects are used by international scorers to estimate reliability. While this process was welcomed among IU participants, the scorers' feedback was not uniformly supportive. The results of this study suggested that the rubric yields acceptable estimates of content validity and inter- and intra-rater score reliability, independent of scorers' philosophical objections.

Although this process was student-centered, there were several limitations to this approach regarding other stakeholders. Sample-size, time allocation, the inclusion of many other stakeholders in determining what the attainment of competencies entail, and researcher's prolonged field engagement are among the limitations. Implications for practice became evident based on this study's findings. These implications were to conduct performance-based assessment with a formative assessment framework, to use student assessment data effectively to yield formative data, to include students as stakeholders when determining assessment content and procedures, and to use assessment processes to inform and include constituencies.

There were several formal research recommendations as a result of the completion of this study. These recommendations were to (a) introduce this instrument-development process to other IUs, (b) to conduct parallel mixed design studies that explore the intersection between HCA and assessment, and (c) to conduct a case study with in-depth analysis to explore assessment as a concept in intercultural settings.

REFERENCES

- Ahuja, R., Berumen, G., Casillas, M. de L., Crispín, M. L., Delgado, A., Elizalde, A., . . . Schmelkes, S. (2004). *Políticas y fundamentos de la educación bilingüe* [Policies and foundations of bilingual education]. Mexico City, Mexico: Secretaría de Educación Pública (SEP).
- Anderson, R. S. (1998). Why talk about different ways to grade? The shift from traditional assessment to alternative assessment. *New Directions for Teaching and Learning*, 1998(74), 5-16.
- Apple, M. (2000). *Official knowledge: Democratic education in a conservative age* (2nd ed.). New York, NY: Routledge.
- Benveniste, L. (2002). The political structuration of assessment: Negotiating state power and legitimacy. *Comparative Education Review*, 46(1), 89-118.
- Bertely Busquets, M. (1998). Educación indígena del siglo XX en México [Indigenous education in the 20th century in Mexico]. In P. Latapí Sarre (Ed.), *Un siglo de educación en México* [A century of education in Mexico] (pp. 74 -101). Mexico City, Mexico: Fondo de Estudios e Investigaciones “Ricardo J. Zevada,” Consejo Nacional para la Cultura y las Artes, Fondo de Cultura Económica.
- Bloom, B., Madaus, G., & Hastings, J. T. (1981). *Evaluation to improve learning*. New York, NY: McGraw-Hill.
- Brighouse, H. (2000). *School choice and social justice*. New York, NY: Oxford University Press.
- Brookhart, S. M. (2007). Expanding views about formative assessment: A review of the literature. In J. H. MacMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 43-62). New York, NY: Teachers College.
- Caro, M. A. (2010). *Semejanzas y diferencias interculturales del concepto de acreditación en los titulares académicos de la UAIM*. [Intercultural similarities and differences of the concept of accreditation among graduating seniors of the UAIM] (Unpublished Master’s Thesis). Universidad Autónoma Indígena de México, Mexico.
- Carney, S. (2003). Globalisation, Neo-liberalism and the limitations of school effectiveness research in developing countries: The case of Nepal. *Globalisation, Societies and Education*, 1(1), 87–101.

- Casillas Muñoz, L., & Santini Villar, L. (2006). *Universidad Intercultural: Modelo Educativo. Serie Documentos* [Intercultural University: Educational Model. Document Series]. Mexico City, Mexico: SEP.
- Coordinación General de Educación Intercultural Bilingüe (CGEIB). (2011). *Investigación y Evaluación Educativas* [Educational research and evaluation]. Retrieved from: <http://eib.sep.gob.mx/index.php/investigacion-y-evaluacion-educativas>
- Creswell, J. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review*, 64(1), 5-30.
- Dann, R. (2002). *Promoting assessment as learning: Improving the learning process*. London, UK: RoutledgeFalmer.
- Dietz, G. (2008). La experiencia de la Universidad Veracruzana Intercultural [The experience of the Intercultural University of Veracruz]. In D. Mato (Ed.), *Diversidad cultural e interculturalidad en educación superior: Experiencias en América Latina* [Cultural diversity and interculturalism in higher education: Experiences in Latin America] (pp. 359-370). Caracas, Venezuela: IESCAL/UNESCO.
- Dietz, G. (2011). Hacia una etnografía doblemente reflexiva: Una propuesta desde la antropología de la interculturalidad [Towards a doubly reflective ethnography: A proposal emerging from the anthropology of interculturalism]. *Revista de Antropología Iberoamericana*, 6(1), 3-26.
- Dietz, G. & Mateos, L. S. (2011). *Interculturalidad y educación intercultural en México: Un análisis de los discursos nacionales e internacionales en su impacto en los modelos educativos mexicanos* [Interculturalism and intercultural education in Mexico: An analysis of the national and international discourses and their impact on Mexican educational models]. Mexico City, Mexico: Secretaría de Educación Pública (SEP).
- Doscher, S. (2012). *The development of rubrics to measure undergraduate students' global awareness and global perspective: A validity study*. (Doctoral dissertation). Retrieved from FIU Electronic Theses and Dissertations. (Paper 588)
- Drèze, J., & Sen, A. (1995). *India: Economic development and social opportunity*. New York, NY: Oxford University Press.

- Fábregas Puig, A. (2008). La experiencia de la Universidad Intercultural de Chiapas [The experience of the Intercultural University of Chiapas]. In D. Mato (Ed.), *Diversidad cultural e interculturalidad en educación superior: experiencias en América Latina* [Cultural diversity and interculturalism in higher education: Experiences in Latin America] (pp. 339-348). Caracas, Venezuela: IESCAL/UNESCO.
- Flores Crespo, P. (2007). Ethnicity, identity, and educational achievement in Mexico. *International Journal of Educational Development*, 27(3), 331-339.
- Gabbard, D. A. (2000). Accountability. In: D. Gabbard (Ed.), *Knowledge and power in the global economy: Politics and the rhetoric of school reform* (pp. 53-61). Mahwah, NJ: Lawrence Erlbaum Associates.
- García Segura, S. (2004). De la educación indígena a la educación bilingüe intercultural: La comunidad p'urhepecha, Michoacán, México [From indigenous education to intercultural bilingual education: The p'urhepecha community, Michoacan, Mexico]. *Revista Mexicana de Investigación Educativa*, 9(20), 61-81. doi: ISSN 1405-6666
- Guerra García, E. (2005). La aneregogía de la voluntad: Propuesta educativa sociointercultural de la Universidad Autónoma Indígena de México [The anaerogy of the will: Sociointercultural educational proposal of the Automous Indigenous University of Mexico]. *Ra Ximhai*, 1(1), 15-38.
- Guion, R. (1977). Content validity: The source of my discontent. *Applied Psychological Measurement*, 1(1), 1-10.
- Guskey, T. (2007). Formative classroom assessment and Benjamin S. Bloom: Theory, research, and practice. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 631-78). New York, NY: Teachers College.
- Instituto Nacional de Estadística, Geografía e Informática (INEGI). (2004). *La población hablante de lengua indígena de Guerrero* [The indigenous language-speaking population of Guerrero]. Retrieved from: http://www.inegi.org.mx/prod_serv/contenidos/espanol/bvinegi/productos/censos/poblacion/poblacion_indigena/PerLi_Gro.pdf
- Johnson, B., & Christensen, L. (2004). *Educational research: Quantitative, qualitative, and mixed approaches* (2nd ed.). New York, NY: Pearson.
- Johnson, B., & Onwuegbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.

- Johnson, R., Penny, J., Fisher, S., & Kuhs, T. (2003). Score resolution: An investigation of the reliability and validity of resolved scores. *Applied Measurement in Education*, 16(4), 299-322.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144.
- Kamens, D., & McNeely, C. (2010). Globalization and the growth of international educational testing and national assessment. *Comparative Education Review*, 54(1), 5-25.
- Landorf, H., & Pineda, M. F. (2010). *Global learning and the human capability approach*. Paper presented at the XIV Congress of the World Council of Comparative Education Societies, Istanbul, Turkey.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563-575.
- Leedy, P. D. (1997) *Practical research: Planning and design* (6th ed.). Upper Saddle, NJ: Prentice-Hall.
- Linn, R., Baker, E., & Dunbar, S. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20(8), 15-21.
- Mansilla, V., Duraisingh, E., Wolfe, C. R., & Haynes, C. (2009). Targeted assessment rubric: An empirically grounded rubric for interdisciplinary writing. *The Journal of Higher Education*, 80(3), 334-353.
- Mato, D. (Ed.). (2008). *Diversidad cultural e interculturalidad en educación superior: Experiencias en América Latina* [Cultural diversity and interculturalism in higher education: Experiences in Latin America]. Caracas, Venezuela: UNESCO.
- Mathison, S., & Ross, E. W. (2008). The hegemony of accountability: The corporate-political alliance for control of schools. In D. Gabbard & E. W. Ross (Eds.), *Education under the security state: Defending public schools* (pp. 91-100). New York, NY: Teachers College Press.
- Mayer, A. (Ed.). (2007). *México en tres momentos, 1810-1910-2010* [Mexico in three periods: 1810-1910-2010]. Mexico City, Mexico: Instituto de Investigaciones Históricas, Universidad Autónoma de México (UNAM).
- McMillan, J. (2007). Formative classroom assessment: The key to improving student achievement. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 1-7). New York, NY: Teachers College.

- Mendoza Zuany, R. G. (2009). Building hybrid knowledge at the Intercultural University of Veracruz, Mexico: An anthropological study of indigenous contexts. *Intercultural Education*, 20(3), 211-218.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Messick, S. (1996). Validity of performance assessments. In G.W. Phillips (Ed.), *Technical issues in large scale performance assessment* (pp. 1-18). Washington, DC: National Center for Educational Statistics.
- Moore, A. & Young, S. (1997). *Clarifying the blurred image: Estimating the inter-rater reliability of performance assessments*. Paper presented at the Annual Meeting of the Northern Rocky Mountain Educational Research Association, Jackson, WY. Retrieved from <http://www.eric.ed.gov/PDFS/ED414319.pdf>
- Morales Garza, S. (1998). La educación indígena, especial e inicial: De “modelos complementarios” a estrategias claves para reivindicar la igualdad en la diferencia [Indigenous, special, and elementary education: From “remedial models” to key strategies to revindicate equality in diversity]. In P. Latapí Sarre (Ed.), *Un siglo de educación en México* [A century of education in Mexico] (pp. 141-172). Mexico City, Mexico: Fondo de Estudios e Investigaciones “Ricardo J. Zevada,” Consejo Nacional para la Cultura y las Artes, Fondo de Cultura Económica.
- Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, 7(3). Retrieved from <http://PAREonline.net/getvn.asp?v=7&n=3>
- Moskal, B. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation*, 8(14). Retrieved from <http://PAREonline.net/getvn.asp?v=8&n=14>
- Moskal, B., & Leydens, J. A. (2000). Scoring rubric development: validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10). Retrieved from <http://PAREonline.net/getvn.asp?v=7&n=10> .
- Newman, I. & Covrig, D. (forthcoming, 2012). Building consistency between the title, problem statement, purpose and research questions to improve the quality of research plans and reports. *New Horizons in Adult and Human Resource Development*.
- Newman, I., Lim, J., & Pineda, M. F. (2011a). *Content validity using mixed methods approach: Its application and development through the use of a table of specifications methodology*. Paper presented at the meeting of the American Evaluation Association, Anaheim, CA.

- Newman, I., & McNeil, K. (1998). *Conducting survey research in the social sciences*. New York, NY: University Press of America.
- Newman, I., Newman, C., Brown, R., & McNeely, S. (2006). *Conceptual statistics for beginners* (3rd ed.). New York, NY: University Press of America.
- Newman, I., Newman, D., & Newman, C. (2011b). Writing research articles using mixed methods: Methodological considerations to help you get published. In T. S. Rocco & T. Hatcher (Eds.), *The handbook of scholarly writing and publishing* (pp. 191-208). San Francisco, CA: Jossey-Bass.
- Newman, I., Ridenour, C. S., Newman, C., & DeMarco, G. M. P. (2003). Typology of purpose. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 167-188). Thousand Oaks, CA: Sage Publications.
- Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Nussbaum, M. (1997). *Cultivating humanity: A classical defense of reform in liberal education*. Cambridge, MA: Harvard University Press.
- Nussbaum, M. (2006). Education and democratic citizenship: Capabilities and quality education. *Journal of Human Development*, 7(3), 385-395.
- Onwuegbuzie, J., Bustamante, R., & Nelson, J. A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research*, 4(1), 56-78.
- Onwuegbuzie, A. J., & Daniel, L. G. (2002). A framework for reporting and interpreting internal consistency reliability estimates. *Measurement and Evaluation in Counseling and Development*, 35, 89-103.
- Onwuegbuzie, J., Leech, N., & Collins, K. (2008). Interviewing the interpretive researcher: A method for addressing the crises of representation, legitimation, and praxis. *International Journal of Qualitative Methods* 7(4), 1-17.
- Onwuegbuzie, A. J., & Daniel, L. G. (2004). Reliability generalization: The importance of considering sample specificity, confidence intervals, and subgroup differences. *Research in the Schools*, 11(1), 61-72.

- Parker, S., Ruvalcaba, L., & Teruel, G. (2002). Schooling inequality among the Indigenous: A problem of resources or language barriers? Research network working paper #R-438. Washington, DC: Inter-American Development Bank.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Pineda, M. F. (2007). Unity and diversity: High school students' perceptions on multiculturalism and the Intercultural University of the State of Guerrero, Mexico. In S. M. Nielsen & M. S. Plakhotnik (Eds.), *Proceedings of the Sixth Annual College of Education Research Conference: Urban and International Education Section* (pp. 90-95). Miami, FL: Florida International University.
- Pineda, M. F. (2010). Standardized Tests in an Era of International Competition and Accountability. In Wiseman, A. (Ed). *The Impact of International Achievement Studies on National Education Policymaking*, Volume 13, International Perspectives on Education and Society Series (pp. 331-353). UK: Emerald Publishing.
- Pineda, M. F., & Landorf, H. (2010). Historical and philosophical foundations of intercultural education in Mexico. In C. Grant (Ed.), *Intercultural and multicultural education: Enhancing global interconnectedness* (pp. 138-153). New York, NY: Routledge.
- Quinlan, A. M. (2006). *A Complete guide to rubrics: Assessment made easy for teachers, K-College*. Lanham, MD: Rowman & Littlefield Education.
- Ramírez Castañeda, E. (2006). *La Educación Indígena en México* [Indigenous Education in Mexico]. Mexico City, Mexico: Universidad Nacional Autónoma de México.
- Reddy, Y. M., & H. Andrade (2009). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435-448.
- Ridenour, C., & Newman, I. (2005). *Implementing mixed methods research design in the real world: Purposes, dilemmas, and new perspectives*. Paper presented at the annual conference of the American Education Research Association, Montreal, Canada.
- Ridenour, C., & Newman, I. (2008). *Mixed methods research: Exploring the interactive continuum*. Carbondale, IL: Southern Illinois University Press.
- Ryan, G., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods*, 15(1), 85–109.

- Saito, M. (2003). Amartya Sen's capability approach to education: A critical exploration. *Journal of Philosophy of Education*, 37(1): 18-33.
- Schmelkes, S. (2008). Creación y desarrollo inicial de las universidades interculturales en México: Problemas, oportunidades y retos [Creation and initial development of the intercultural universities in Mexico: Problems, opportunities, and challenges]. In D. Mato (Ed.), *Diversidad cultural e interculturalidad en educación superior: Experiencias en América Latina* [Cultural diversity and interculturalism in higher education: Experiences in Latin America]. (pp. 329-338). Caracas, Venezuela: IESCAL/UNESCO.
- Schmelkes, S. (2009). Intercultural universities in Mexico: progress and difficulties. *Intercultural Education*, 20(1), p. 5-17.
- Sen, A. K. (1992). *Inequality reexamined*. Cambridge, MA: Harvard University Press.
- Sen, A. K. (1995) Gender, inequality and theories of justice. In M. Nussbaum & J. Glover (Eds.), *Women, culture and development: A study of human capabilities* (pp. 259-273). New York, NY: Oxford University Press.
- Sen, A. K. (1997). *On economic inequality*. New York, NY: Oxford University Press.
- Sen, A. K. (1999). *Development as freedom*. New York, NY: Oxford University Press.
- Spring, J. (2008). Research on globalization and education. *Review of Educational Research*, 78(2), 330-363.
- Stevens, D., & Levi, A. J. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning*. Sterling, VA: Stylus.
- Stiggins, R. (2007). Conquering the formative assessment frontier. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 8-28). New York, NY: Teachers College.
- Strauss, A., & Corbin, J. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13(1): 3-21.
- Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Los Angeles, CA: Sage.

- Terzi, L. (2007). The capability to be educated. In M. Walker & E. Unterhalter (Eds.), *Amartya Sen's Capabilities Approach and Social Justice in Education* (pp. 25-43). New York, NY: Palgrave.
- Thompson, B., & Vacha Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*, 60, 174-195.
- Treviño Villareal, E. (2006). Evaluación del aprendizaje de los estudiantes indígenas en América Latina: desafíos de medición e interpretación en contextos de diversidad cultural y desigualdad social [Evaluation of the learning of indigenous students in Latin America: Challenges of measurement and interpretation in contexts of cultural diversity and social inequality]. *Revista Mexicana de Investigación Educativa*, 28(11), 225-268. doi: ISSN 1405-6666
- Unterhalter, E., & Brighouse, H. (2007). Distribution of what for social justice in education? The case of education for all by 2015. In M. Walker & E. Unterhalter (Eds.), *Amartya Sen's capabilities approach and social justice in education* (pp. 67-86). New York, NY: Palgrave.
- Unterhalter, E., Vaughan, R., & Walker, M. (2007). *The Capability Approach and Education*. Retrieved from <http://www.nottingham.ac.uk/educationresearchprojects/developmentdiscourses/papersandresources.aspx>
- Vacha-Haase, T., Kogan, L. R., & Thompson, B. (2000). Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. *Educational and Psychological Measurement*, 60, 509-522.
- Walker, M. (2008). A human capabilities framework for evaluating student learning. *Teaching in Higher Education*, 13(4), 477-487.
- William, D., & Leahy, S. (2007). A theoretical foundation for formative assessment. In J. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 29-42). New York, NY: Teachers College.
- Witta, E. L., & Daniel, L. G. (1998). *The reliability and validity of test scores: Are editorial policy changes reflected in journal articles?* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Zolla, C., & Zolla, E. (2004). *Los pueblos indígenas de México: 100 preguntas* [Indigenous groups of Mexico: 100 questions]. Mexico City, Mexico: UNAM. Retrieved from: http://www.nacionmulticultural.unam.mx/100preguntas/pregunta.html?num_pre=1

Zimmaro, D. M. (2004). *Developing grading rubrics*. Retrieved from
<http://www.utexas.edu/academic/mec/research/pdf/rubricshandout.pdf>

APPENDICES

APPENDIX A

First Focus Group's Questions

First Focus Group's Questions

1. Why did you decide to come to the IU?
2. How would you describe your assessment experience at the IU?
3. What do you think is the best way to show attainment of competencies upon graduation for IU students?

The IU's graduates are supposed to achieve these five competencies upon graduation:

- articulate regional/indigenous knowledge and initiatives
 - facilitate advocacy for, and management of resources and information
 - strengthen and design intercultural initiatives
 - highlight regional/indigenous knowledge and initiatives, and
 - generate diagnostic and proposal-oriented knowledge
4. Competency #1 is "To be able to articulate knowledge (scientific and indigenous) and bridging initiatives." In your own words, what does this mean and what evidence indicates that you have achieved this outcome?
 5. Competency #2 is "To be able to facilitate advocacy for, and management of resources and information." In your own words, what does this mean and what evidence indicates that you have achieved this outcome?
 6. Competency #3 is "To be able to strengthen and design intercultural initiatives." In your own words, what does this mean and what evidence indicates that you have achieved this outcome?
 7. Competency #4 is "To be able to highlight regional/indigenous and initiatives." In your own words, what does this mean and what evidence indicates that you have achieved this outcome?
 8. Competency #5 is "To be able to generate diagnostic and proposal-oriented knowledge." In your own words, what does this mean and what evidence indicates that you have achieved this outcome?
 9. What was challenging, if anything?
 10. What would you do differently if you were to start over at the IU?
 11. Would you like to add anything else?

APPENDIX B

Expert Judges' Focus Group/Working Session Questions

Panel of Experts' Focus Group/Working Session Questions

1. How would you describe your assessment experience at the IU?
2. Please refer to the list of competencies handed to you-

The IU's graduates are supposed to achieve these 5 competencies upon graduation:

- articulate regional/indigenous knowledge and initiatives,
- facilitate advocacy for, and management of resources and information,
- strengthen and design intercultural initiatives,
- highlight regional/indigenous knowledge and initiatives, and
- generate diagnostic and proposal-oriented knowledge.

3. Competency #1 is "To be able to articulate regional/indigenous knowledge and initiatives." In your own words, what does this mean and what evidence indicates that the student has achieved this outcome?

4. Competency #2 is "To be able to facilitate advocacy for, and management of resources and information." In your own words, what does this mean and what evidence indicates that the student has achieved this outcome?

5. Competency #3 is "To be able to strengthen and design intercultural initiatives." In your own words, what does this mean and what evidence indicates that the student has achieved this outcome?

6. Competency #4 is "To be able to highlight regional/indigenous knowledge and regional initiatives." In your own words, what does this mean and what evidence indicates that the student has achieved this outcome?

7. Competency #5 is "To be able to generate diagnostic and proposal-oriented knowledge." In your own words, what does this mean and what evidence indicates that the student has achieved this outcome?

8. Would you like to add anything else?

Appendix C

Newman and McNeil (1998) Checklist for Instrument Development

Newman and McNeil (1998) Checklist for Instrument Development

The researcher used this checklist as an audit trail tool, for transparency purposes.

Stage 1: Specify the intent of the instrument

Question 1: What are the basic questions I want my instrument to answer?

Though this was not an evaluative study, the question related to the purpose of the instrument, which was to measure students' attainment of the IMD competencies

Question 2: How do I plan to use the information?

The goal was to develop a process to create a performance-based assessment instrument, or a rubric, with acceptable estimates standards of content validity and score reliability, a process that hopefully other IUs could replicate. If the instrument were to be found useful, it should inform students, educators, and other stake-holders about the educational process of the IU students.

Question 3: Who do I plan to collaborate with to obtain those answers?

Primarily IU seniors and/or recent graduates

Stage 2: Assess the Available Resources

Question 4: What assistance do I need in creating this instrument?

Participation during focus groups, review of different research projects related to this one, and peer debriefing.

Question 5: What types of analyses are required?

- (1) Open coding of the qualitative data obtained through focus groups, deductive analysis of scorers' feedback (and any other participant that provides feedback),
- (2) estimates of content validity and inter- and intra-rater reliability, using an archive of students' research projects to pilot the rubric.

Question 6: Are computer programs available? N/A

Stage 3: Define the population

Question 7: What is the population to which I am interested in generalizing? My study did not seek generalization, but I would like the IU students to use my instrument development process in the future should they find it useful.

Question 8: What are the characteristics of the population to which I am interested in generalizing? This study did not aim to generalize.

Question 9: How does the accessible population differ from the population which I want to generalize? This study did not aim to generalize.

Question 10: How would volunteering affect my results?

Volunteering should actually benefit the process of making the instrument, because participants should come with the intent to collaborate and provide the researcher with rich data. There were no monetary incentives to participate in this study, but scorers received a gift certificate to a bookstore as a token of appreciation for their weeks of work.

Stage 4: Review of the Pertinent Literature

Question 11: Where can I look for information?

Libraries, online databases and through my contacts.

Question 12: What are the key terms for my literature search, for the section of instrument making? (a detail description of other sections of the literature review is available in Chapter 2)

“Rubrics,” “Performance-based assessment in higher education, and “collaboration in making rubrics.”

Question 13: What are the best sources of information?

Peer reviewed educational journals (including Mexico’s renowned journal RME), books, official websites, and dissertation repositories.

Stage 5: Determine the data collection techniques

Question 14: What instrument procedure should I use?

I will use the Human Capabilities Approach as a philosophical framework. The HCA highlights participatory dialogue and democratic deliberation, and active participation of stakeholders, where the researcher serves as facilitator only. Once the early draft of the instrument was completed, I facilitated the procedures for estimating content validity, Lawshe’s CVR, inter-rater reliability and intra-rater reliability.

Question 15: Should I use an existing rubric or develop my own?

The purpose of this study was to develop a process for an instrument. Other rubrics could serve as a reference for format purposes only.

Question 16: Who should administer the rubric?

This instrument will be used only by experts and scorers at the moment

Stage 6: Develop the instrument

Question 17. What are the item format considerations?

Each level descriptor of the rubric (beginning, developing, accomplished, exemplary) should be comprehensible and be clearly distinct from other levels.

Question 18. What are the psychometric considerations?

Expert Judge Content Validity, Inter-rater reliability and intra-rater reliability estimates.

Question 19. What are the piloting considerations?

Content experts will participate in estimating validity. Scorers (Spanish speaking doctoral students) will assist in estimating reliability. They will receive the same number of *documentos recepcionales* and each one will be scored at least twice.

Question 20. What are the training considerations?

Experts will receive instructions along with the Table of Specifications for content validity and a rating scale for each artifact that the graduates submit. No complex training is necessary for these tasks. The researcher will debrief them in person and will be present at all times should any question emerge. Scorers will receive instructions via email.

Stage 7. Determine the sampling procedure

Question 21. How should I sample? This study uses purposive sampling

Stage 8. Analyze results and prepare a report

Question 22. What analyses should I perform? Validity and Reliability Estimates for quantitative data, member-checking for qualitative data

Question 23. How should I write my report? In a Dissertation Format

Appendix D
Rubric Template

IMD RUBRIC				
Competency Description & Level of Attainment	BEGINNING (1)	DEVELOPING (2)	ACCOMPLISHED (3)	EXEMPLARY (4)
Competency 1:				
Competency 2:				
Competency 3:				
Competency 4:				
Competency 5:				

APPENDIX E

Table of Specifications in Spanish and English

TABLA DE ESPECIFICACIONES PARA ESTIMADOS DE VALIDEZ

INSTRUCCIONES:

Basado en las descripciones de cada competencia y nivel de aprovechamiento de la “matriz de descripción de competencias,” por favor marca (✓) el nivel que te parezca apropiado para cada uno, ya sea “principiante,” “en desarrollo,” “satisfactorio” u “ejemplar,” si le falta algo y si se debe aceptar (SI/SI, CON RETROALIMENTACION/NO). Después, indica si todas las evidencias/descripciones de cada nivel se deben aceptar (SI/SI, CON RETROALIMENTACION/NO). Escribe cualquier comentario/retroalimentación que tengas sobre cada una de las descripciones. Son 25 descripciones. ¡Gracias!

Fernanda Pineda (2012)

COMPETENCIA / ALIENAMIENTO	PRINCIPIANTE	EN DESARROLLO	SATISFACTORIO	EJEMPLAR	¿Qué le falta a la descripción?	Las evidencias/descripciones son suficientes y se debe aceptar? SI/SI, CON RETROALIMENTACION/ NO
Competencia 1/1						
Competencia 1/2						
Competencia 1/3						
Competencia 1/4						
Competencia 2/1						
Competencia 2/2						
Competencia 2/3						
Competencia 2/4						
Competencia 3/1						

Competencia 3/2						
Competencia 3/3						
Competencia 3/4						
Competencia 4/1						
Competencia 4/2						
Competencia 4/3						
Competencia 4/4						
Competencia 5/1						
Competencia 5/2						
Competencia 5/3						
Competencia 5/4						

GRACIAS!

TABLE OF SPECIFICATIONS FOR ESTIMATING CONTENT VALIDITY

INSTRUCTIONS:

Based on the descriptions of each competency and the attainment levels of the “Rubric for the Attainment of Competencies” please put a check mark (√) by each of the attainment level that you think it is appropriate to align it (that it best describes it), either Beginner, Developing, Satisfactory, or Exemplary, and also indicate if there is anything missing to it, and whether it should be accepted (YES / YES, WITH FEEDBACK / NO). Then, indicate if all the evidences/descriptions of each level should be accepted (YES / YES, WITH FEEDBACK / NO). Write any comment/feedback you may have about each of the descriptions. There are 25 descriptions. Thank you!

Fernanda Pineda (2012)

COMPETENCY / ALIGNMENT	BEGINNER	DEVELOPING	SATISFACTORY	EXEMPLARY	What is missing in this description?	Are the evidences/descriptions sufficient, and therefore, should be accepted? YES / YES, WITH FEEDBACK / NO
Competency 1/1						
Competency 1/2						
Competency 1/3						
Competency 1/4						
Competency 2/1						
Competency 2/2						
Competency 2/3						
Competency 2/4						
Competency 3/1						

Competency 3/2						
Competency 3/3						
Competency 3/4						
Competency 4/1						
Competency 4/2						
Competency 4/3						
Competency 4/4						
Competency 5/1						
Competency 5/2						
Competency 5/3						
Competency 5/4						

THANK YOU!

Appendix F
Informed Consent

Consent to Participate in a Research Study

Title: The Development of an instrument for the attainment of competencies at an Intercultural University In Mexico

You are being asked to participate in a research study. The investigator of this study is Martha Fernanda Pineda, a doctoral candidate at Florida International University. This study may include approximately 15 recent graduates from the Intercultural University (IU) and education professors in the United States and Mexico. Your participation will require approximately 2 hours of your time in person during a focus group and if necessary, some electronic communication with the researcher. The purpose of this study is to develop a process to create an instrument for the attainment of competencies at the IU. The copyright of the instrument will belong to the researcher and unlimited use will be available to the IUs. The objective of the focus groups is that participants provide input as to what each competency entails, and what evidence is appropriate to show that the competency has been achieved. "Evidence" is understood as the artifact developed during your years at the IU, submitted for graduation.

If you decide to be a part of the study, you will participate in a semi-structured focus group along with other graduates and the focus group will be recorded. You will receive the list of questions that will be used as a guide for the focus group. We do not expect any harm to you by being in the study. You may skip any questions that you do not want to answer. If you get upset or feel discomfort during the focus group, you may ask to take a break. There is no cost or payment to you as a participant. You will not get any direct benefit or actual reward from being in the study. However, your participation will be useful in the making of an educational assessment instrument that current IU students and educators might find useful.

Your responses will be identified by a pseudonym, and not your name. All your answers are private and will not be shared with anyone unless required by law. The artifact you share will be photocopied and you will retain all original work. Your name will be erased from the photocopied artifact and all subsequent photocopies of your work (if it is necessary to make more photocopies). For this study, the researcher and IU experts will use these artifacts only to estimate reliability and validity of the instrument. Once the

study is completed, the researcher will store the copies of the artifacts along with the transcripts and audio-recordings obtained during the focus group in a safe place. The photocopies of your artifacts might be used for subsequent research endeavors that the researcher might have, and the same confidentiality protocol will be followed. You may ask questions about the study at any time. If you choose not to participate, no one will be upset with you. You may also choose to stop your participation before the focus group is finished.

If you would like more information about this research study after you are done, you can contact Fernanda Pineda at 01 (305) XXX-XXXX or Fernandapineda@xyz.com. If you would like to talk with someone about your rights as a participant in this study, you may contact Dr. Patricia Price, the Chairperson of the FIU Institutional Review Board at 305-348-2618 or 305-348-2494. Your signature below indicates that all questions have been answered to your liking. You are aware of your rights and you would like to be in the study.

Signature of Participant	Printed Name	Date
--------------------------	--------------	------

I have explained the research procedure, subject rights, and answered questions asked by the participant. I have offered him/her a copy of this informed consent form.

Signature of Witness	Date
----------------------	------

Appendix G

Lawshe's CVR Minimum Values and their Significance for One-tailed Tests

Lawshe's CVR minimum acceptable values and their significance for one-tailed tests when all experts or panelists say the description is appropriate, the computed CVR is 1. When it is more than half of them, but less than all, the CVR goes between zero and .99

Number of Panelists	Minimum CVR Value
5	.99
6	.99
7	.99
8	.75
9	.78
10	.62
15	.49
20	.42
25	.37
30	.33

APPENDIX H

Rubric Drafts of Each Participant of Focus Group One and the Researcher's Compilation
Rubric

Student 1

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE 1	EN DESARROLLO 2	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: Saber vincularse con la comunidad y las iniciativas que tienen, apoyar, aprender con y de ellos.	Realización de Diagnósticos comunitarios, aplicación de técnicas básicas.	Aplicación de técnicas participativas Diagnósticos participativos, identificación de problemas comunitarios	Priorizar Problemas juntos en comunidad Vinculación y apoyo con iniciativas	Articulación de iniciativas comunitarias y regionales
Competencia 2:	Conocer instituciones que apoyan a las comunidades. Que hacen y como	Disear proyectos	Gestionar y evaluar proyectos Gestionar y evaluar proyectos	Apoyar a las comunidades enseñando a gestionar.
Competencia 3: Fortalecimiento de iniciativas.	Apoyo a iniciativas	Grupo de actores, Búsqueda de apoyo en lo que se haga Faltan personas.	Poner en práctica las iniciativas.	
Competencia 4:	Conocer las iniciativas / a fondo / o Generar iniciativas	Diffundir las iniciativas existentes. y concentrar sobre problemáticas	Criticar y analizar las iniciativas propositivo	ser propositivo rediseñar junto con la comunidad las iniciativas que existen, difundir info
Competencia 5:	Conocer las cosmovisiones de los agentes, donde trabajan / porque piensan	Reconocimiento de los saberes locales, investigar prácticas culturales	Diffundir el conocimiento de los saberes, Integrar el diálogo de saberes.	Apoyo a iniciativas culturales / Difusión del

Prácticas culturales y formas de convivencia

en los problemas Desarrollo de políticas de trabajo y participación

Student 2

MATRIZ DE DESCRIPCIÓN DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: Es la parte donde conocemos a los actores org. y sociales de una comunidad para poder generar la vinculación y promover la intervención.	Realizar de Diag. Comunitario, Identificación de actores, salir a conocer la comunidad a través de Observar y Diagnosticar de varios Diagnosticos	Integrar un grupo de trabajo. Conocer perfectamente la comunidad a través de varios Diagnosticos	Realización de Talleres, Proceso de Vinculación con los actores sociales. Encaminar la investigación hacia la teoría o un proyecto.	El trabajo con ese grupo de actores, poner en práctica algún proyecto, Poner en práctica todos sus conocimientos en la práctica.
Competencia 2: Poder buscar apoyo de diferentes grupos para poder realizar los trabajos. Recursos no solo económicos.	Identifica las necesidades problemáticas y conoce que son los recursos que necesita.	Aplica técnicas para poder obtener recursos, ya sea de manera general o individual.	Conoce las Org. Instituciones. Sabe buscar sus actores sociales que pueden brindarle apoyo y conoce las bases para gestionarse apoyo.	Sabe buscar sus proyectos, es el dueño o diseña sus bases de trabajo grupal o individual.
Competencia 3: Fortalecer la vinculación, aplicación de técnicas e iniciativas.	Diseñar, conocer diagnosticar, formular.	Evalúa las iniciativas que aplicó acompaña procesos, conoce más sobre políticas técnicas.	aplica lo que aprendió, diseña nuevas y mejores herramientas.	Genera nuevas ideas, aplica nuevas y mejores las pone en práctica en conjunto y evalúa.
Competencia 4: Conocer, Analizar y Sistematizar los conocimientos.	Observa, conoce y sabe sobre las iniciativas y los saberes comunitarios.	Emplean instrumentos, diagnósticos y técnicas de manera participativa para conocer esos saberes.	Es mucho más analítico sobre esos conocimientos, los analiza para mejorarlos, genera opiniones y es abierto a nuevos saberes, aspectos.	Es capaz de analizar diferentes situaciones de vista en varios aspectos.
Competencia 5: Generar más conocimiento propositivo y analítico para propositivo.	Conocer los saberes y conocimientos que ya existen.	Aplica técnicas y herramientas para poder ser analítico y propositivo generando nuevos conocimientos sin olvidar o hacer menos los que ya están.	Propone en conjunto para mejorar algo. Adquiere más conocimiento lo transmite y lo valora.	Propone acciones de mejora en conjunto y los lleva a la práctica y los evalúa.

Student 3

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: Constar los conocimientos comunitarios para propiciar los.	Diagnóstico comunitario	acercamiento a las personas a través de grupos de análisis	asesoramiento en las diversas manifestaciones o causas ya sea problemática u otros	empleo de los saberes a diversos ambitos
Competencia 2: búsqueda de apoyos y facilitación en los ambitos.	enfoces de dificultades	realización y aplicación de P	detección y beneficio de progreso	Fortalecimiento de gestión.
Competencia 3: búsqueda de estrategias para apoyar	visualiza la realidad para mejorar el ambito.	Refuerza las iniciativas dadas para el beneficio del grupo del.	Mediación y detección para la aplicación de conocimientos	inserción en los ambitos a desarrollarse
Competencia 4: Ver y comparar regiones	detectar las diferentes formas de aplicación	reforzamiento en la difusión	plata a conocer las formas metodológicas	emplear el conocimiento no juzga -criticas propositivas.
Competencia 5: detección de problemáticas a partir de su entorno social/cultural.	questionar los problemas para propiciar las estrategias	Exemplificación de su cultura	Anticipación y jerarquización de las propuestas.	valoración de las formas de vida.

Student 4

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: Articulación entre estudiante y comunidad Vinculación acción	Se relaciona con los diversos actores	Aplicación de herramientas metodológicas	Es crítico a partir de la experiencia vivida	Genera comunicación con los diversos actores
Competencia 2: Herramientas adquiridos para aplicarlos en la práctica	Es compañero de los grupos/comunidades Ha generado confianza	Empieza ser más consciente	Es explorando los recursos.	Como que quiere valores (se identifica)
Competencia 3:	Adjunto con la comunidad analiza y identifica.	Hacer con la comunidad y para la comunidad a partir de necesidades.	Empieza preocuparse para generar acciones	Es flexible, comprensivo.
Competencia 4:	Tiene el interés de iniciar.	Dialoga a través de las necesidades comunitarias.	Es el conocimiento adquirido que el estudiante aplica y genera profesionalmente.	Analiza - Encuentra la iniciativa sin prejuicios.
Competencia 5:	Identifica, visualiza, propone: (Quiere)	Es capaz de desarrollarse en el ámbito comunitario	Se ha generado el conocimiento y la formación del estudiante. Propone, actúa.	Es la lucha para poder generar, proponer, actuar. Criticar

Student 5

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: - Acercamiento a las comunidades para identificar iniciativas, problemáticas, necesidades.	- observación, entrevistas sobre temas específicos diagnósticos.	- ^{análisis de información} Búsqueda de herramientas metodológicas y teorías para llevarlas a la práctica.	- Implementación de propuestas viables en conjunto con la comunidad.	- Alguien que ha adquirido conocimientos y para ponerlos en práctica en el campo laboral.
Competencia 2: - Gestión de recursos e información	- investigación - Investigación de dependencias.	- Hacer un diagnóstico, en conjunto, con las comunidades.	- Trabaja soluciones - intermedia, trata de	- Persona que comparte y en conjunto con sus colegas pone en práctica en el contexto laboral.
Competencia 3:	- Elaboración de herramientas metodológicas.	- Formar parte de las comunidades en quien se trabaja.	- Cronograma de actividades - soluciones	- inserción campo laboral.
Competencia 4:	- comparación regiones	- concientización de la problemática, necesidad o iniciativa	- Reforzar iniciativas con actividades y prácticas A.C. ó NGS para la promoción de promoción.	- Combina prácticas de otras A.C. ó NGS para en otras actividades.
Competencia 5: - tomar en cuenta los saberes tradicionales de la investigación específica	- Justificación y argumentación del proyecto o tema específico.	- combinación - Diálogo de saberes, visibilización de la cultura.	-	- generación de posibles soluciones - Alternativas viables.

Student 6

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: Saber acercarse a la comunidad, para ver como es, que pasa en ella. una mini vinculación	Deriva a pen de en breuista, un primer diagnóstico. Empleando la observación	La realización de diagnóstico, conocer y saber más fondo el enfoque de una comunidad es decir sus características	La seguridad de realizar diagnóstico. Saber hacer una vinculación. Implementar la gestión	Poder ser capaz de bajar proyecto. una A.C. Dominio de comunidad y aplicar el saber-hacer
Competencia 2: Saber vincularse con actores y asociaciones.	• Vinculación con organizaciones. Y saber que hacen. Y estudio a los grupos sociales	* Se arriesga a hacer p. diseñar los proyectos	Se atreve a ver las intenciones para poder insertar su proyecto y evaluarlo	Pues lo acreditación del proyecto y trabajar y que se le den seguimiento. intermedios
Competencia 3: Tener la capacidad de ser un facilitador dentro de las comunidades regionales	Se aplica en proceso de comunicación	Aplica un diseño de iniciativas	Saber y hacer algo por la comunidad. Integrar los proyectos	Pone en práctica sus iniciativas
Competencia 4: Saber identificar los inicios y los problemas	Conoce los Potenciales y problemáticas. Iniciales.	Aplicación de como por hacer algo como la concentración	hace ya propuesto y analítico	saber y proponer algo según los problemas difunde
Competencia 5: Tener la capacidad de compartir los conocimientos con la comunidad.	• Saber el por qué. Cimiento de la gente. Que como la comunicación diagnóstico	Saber respetar los conocimientos y compartir lo que sabe. Valora. todo	Saber compartir los conocimientos que se tienen y tro. En conjunto	Tener algo bueno en beneficio los difunden.

Student 7

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: Aptitud de Vinculación a la comunidad-intervención social.	Primero se plantea la necesidad de conocer la comunidad y la vida propia.	Se plantea la necesidad de conocer la situación actual de la comunidad en sus problemáticas	Se toman acciones que puedan ayudar a mejorar la situación actual de la comunidad	Sera capaz de tener una apropiación de las necesidades de la comunidad y poder intervenir para mejorar o apoyar a las necesidades encontradas.
Competencia 2:	Investigación programas y comunidades/Intervención	enlace de programas e iniciativas a la comunidad	Como intervenir para que los des pueden servir para un bien común.	aprovechar y realizar proyectos desde la visión de la comunidad región.
Competencia 3:	Establecer una comunicación con la comunidad	información y conocimiento teórico/comunitario	Establecer y conectar los dos conocimientos	realiza proyectos con los dos visiones
Competencia 4:	Explorar los conocimientos de los miembros de la comunidad	conoce y pone en práctica teóricos/prácticos de la comunidad	aprende a conectar los tres cosas.	los pone en práctica.
Competencia 5:	incita convivencia con comunidad y conoce sus prácticas culturales	realiza el diagnóstico para saber como se presenta la problemática de la comunidad	Sabe como manejar lo teórico/práctico para intervenir en la problemática.	conoce y aplica los conocimientos que tiene e incluso busca apoyo para intervenir.

Researcher's Compiling Rubric

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1: ① > lo inmediato ②	observar - integrarse conocer las personas explorar - conoce contextualizar - conoce no hace gestiones herramientas básicas	analizar identif. temas incluir técnicas participativas contextualizar y usar vinculos para trabajo	tiene la capacidad de acción de la comunidad - verdaderamente participativa - gestión - identifica y asesora a nivel - aplica herramientas necesario	- son capaces de autoemplearse - hace diagnósticos - más allá de lo inmediato - en la escuela - no romantiza - expectativas - analizar iniciativas regionales - más allá de "mi comunidad"
Competencia 2: acerca a las comunidades la vida que resisten	conoce dependencias, asociaciones, personas - introduce a las metod. - estudia lo que hacen - identifica problemáticas de comunidad/familia - imagina intervenciones	- se acerca a conocer sus propios proyectos - protege - ve qué proyectos han fracasado y en cuáles se pueda basar - hace diagnósticos	- ya basados, los hace y los evalúa con el impacto - genera NTD, institutos, sus recursos, vincula con sus pares - conoce los buses - gestiona - encamina proyectos - intermedia algunos	- aprende con cursos B, HR - no proyecta desde arriba - escucha a gestores - empodera a la comunidad - intermedia en conflicto - participa en la escuela - interviene en mejorar el empoderamiento
Competencia 3: * us-	identifica, conoce las problemáticas - no fuerza lo que no le conoce - ve señales de ese grupo - ve carencias y debilidades	- busca saber que lo apoya - discute sus iniciativas - identificar herramientas - HUMANIZACIÓN	- implementa iniciativas del grupo comunitario - analiza	- promueve y facilita - se instala en el ámbito local
Competencia 4:	- empieza a utilizar - compara regiones - ve conocer potencialidades de una comunidad - juzga planes de la - define de fondo la vida	consolidación de... problemática - es más analítico, no juzga - crítico, video	- no juzga, valora conciencia - critica y analiza de manera propositiva conflict. + el paz	- empieza a ver resultados y extiende nuevas propuestas - acciones - genera más, abre a nuevos saberes - rediseña con la comunidad los planes.
Competencia 5:	conoce cosmovisión que, cómo, piensan? - mitos, prácticas, cosmov. - diagnóstico de cultura - identifica problemas desarrollo, propone	- aplica metodologías - no olvida los conocimientos que existen y valor - difusión - anima a la valoración	- justifica y argumenta del xs. - conforme avanza conoce, valora más, define prop. y la de - enlace de saberes y visible valoración de saberes - interacción genera desarrollo político político	- se propositivo de caracterización - sabe que las comunidades su diferencia - sistematizar cosmov. - definiciones obvia - lo más viable - hace la - define - genera proyectos 5/2 - proyecto a la prácticas

Appendix I

Researcher's Notes during Content Validity Session, Focus Group Two

fernanda

MATRIZ DE DESCRIPCION DE COMPETENCIAS LGDI				
Descripción de la Competencia y el Nivel de Desempeño	PRINCIPIANTE (1)	EN DESARROLLO (2)	SATISFACTORIO (3)	EJEMPLAR (4)
Competencia 1:	El LGDI empieza contextualizar su trabajo, a integrarse, a observar, a explorar y a conocer a las personas con las que va a trabajar. Conoce las herramientas básicas de un LGDI. No tiene aún la capacidad de gestionar.	El LGDI empieza a hacer análisis sobre sus observaciones. Identifica temas que se puedan trabajar. Contextualiza y crea vínculos de trabajo. Empieza a incluir técnicas participativas.	El LGDI tiene la capacidad de acción a favor de la comunidad. Sus técnicas son verdaderamente participativas, y la colaboración es genuinamente comunidad-estudiante-institución. Comprende a fondo el concepto de "gestión". Identifica oportunidades de gestión y asesora <u>cuando se necesita</u> . Aplica las herramientas de gestión.	El LGDI muestra disposición de experto en diagnósticos y va más allá de lo aprendido en las aulas. <u>No romantiza la GDI</u> . Articula iniciativas regionales. Crea vinculaciones regionales e interregionales además de con instituciones y con organizaciones no gubernamentales. El LGDI muestra capacidad de autoemplearse.
Competencia 2: Es saber acercar a las comunidades la información que requieren y que va más allá de recursos económicos y recursos humanos.	El LGDI conoce dependencias, asociaciones y personas. Su introducción y conocimiento de la metodología es básica. Empieza a estudiar lo que hace un LGDI. Identifica problemáticas de comunidad y familia e imagina intervenciones.	El LGDI se arriesga a diseñar sus propios proyectos y los pilotea. Ve qué proyectos han funcionado en el pasado y en cuales es factible basarse para futuros proyectos. Hace diagnósticos. <i>Conoce dep., identifi proyectos. Tiene conoc. de las bases de los campos de un proyecto</i>	Ya diseñado su proyecto, el LGDI lo lleva a cabo y evalúa su impacto. Conoce las bases de ONGs, instituciones y sus recursos y vincula esto con su proyecto. Gestiona. Encamina otros proyectos y sirve de intermediario para otros.	El LGDI aprende con recursos económicos y humanos. Diseña proyectos desde "abajo" y no desde "arriba". No solamente gestiona sino que enseña a gestionar, empoderando a la comunidad, e interviene en regresar el empoderamiento. Es intermediario en conjunto. Practica en la realidad.
Competencia 3: [alumnos consideran esta competencia como "vaga"]	El LGDI identifica y conoce su comunidad y las problemáticas. Sabe de observación y sabe de diagnósticos. Identifica las iniciativas a nivel de las problemáticas. No fortalece iniciativas aún porque no las está implementando. Se siente <u>parte de ese grupo</u> (comunidad). Ve carencias y debilidades.	El LGDI a logrado identificar a fondo su comunidad y sus problemáticas. Busca gente que lo apoye. Diseña sus iniciativas e identifica herramientas. <i>identifica "grupos focales"</i>	El LGDI indaga más a profundidad cuales son las necesidades e <u>implementa</u> iniciativas del grupo <u>propio</u> comunitario.	El LGDI promueve y facilita soluciones, no nada más comprende problemáticas. Se inserta con la comunidad o en el ámbito laboral.
Competencia 4:	El LGDI empieza a visualizar y compara regiones. Reconoce	El LGDI tiene conscientización de las problemáticas	El LGDI no juzga sino valora el conocimiento comunitario. Critica	El LGDI al ver resultados y se empieza a extender nuevas

	<p>Adaptación y comunidad. Juzga desde su punto de vista cultural los saberes diferentes (le falta el respeto total). Comienza pláticas de cómo difundir la información.</p>	<p>la gente; es más analítico y no juzga. Produce materiales como trípticos y videos para empezar a difundir información.</p> <p><i>Empieza a hacer soluciones que pueden ser fáciles</i> <i>Busca sumarse a los</i></p>	<p>con fundamentos teóricos y prácticos y se refleja en la práctica. Hace más materiales y hasta talleres.</p>	<p>acción con la comunidad. Se basa en acciones. Genera más conocimientos más allá de lo adquirido y está abierto a nuevos saberes. Ve el trabajo totalmente "horizontal" entre gestores y comunidad.</p>
Competencia 5:	<p>El LGDI empieza a conocer la cosmovisión de la comunidad: el qué y cómo piensan. Conoce los conceptos de sus mitos, prácticas culturales y maneras de convivir. Hace diagnósticos de cultura. Identifica problemas y diseña y propone soluciones a niveles básicos.</p>	<p>El LGDI aplica metodología/técnicas. Valora los conocimientos que existen y los difunde; anima la valoración de los conocimientos; comienza a generar nuevos conocimientos sin olvidar o hacer menos los que ya estaban antes.</p>	<p>El LGDI justifica y busca los argumentos sobre el porqué es necesario fortalecer esa(s) iniciativa(s). Conforme adquiere conocimiento, valora más su cultura propia y la de otros; hace un enlace de saberes y hace visible la valoración de saberes; practica una interculturalidad genuina.</p>	<p>El LGDI es propositivo de conocimientos; sabe que las comunidades son dueños de los conocimientos y todo lo hace en conjunto; sistematiza el conocimiento y lo difunde (lo hace visible); da posibles soluciones o busca lo más viable; desarrolla políticas públicas; critica proyectos y sabe cuáles se pueden llevar a la práctica.</p>

comunitarios
(cultural, social, eco, político)

realiza
determina
sabe qué proyectos
- Identifica sus los más viables
- Asegura

Fernanda

Appendix J

Final Version of the Rubric

RUBRIC TO MEASURE STUDENTS' ATTAINMENT OF COMPETENCIES OF THE IMD (LGID) MAJOR AT THE IU				
Description of the Competency and the Level of Attainment	BEGINNER (1)	DEVELOPING (2)	SATISFACTORY (3)	EXEMPLARY (4)
<p>Competency 1:</p> <p>“The student will be able to articulate regional/indigenous knowledge and create initiatives”</p>	<p>The LGID* begins to contextualize his/her work, integrating into the community, observing, exploring and knowing the people he/she will be working with. He/she knows the basic theoretical tools of the LGID major. The student does not have the ability of doing intercultural management.</p>	<p>The LGID starts to analyze his/her observations; identifies themes that can be productive to work on; contextualizes and creates work connections. He/she starts to include participatory techniques in his work.</p>	<p>The LGID is able to advocate in favor of the community. His/her techniques are truly participatory and this participation is genuinely including the community-student-institutions. The student understands the concept of “intercultural management” in depth. He/she identifies management opportunities and creates links with different consultants. He/she applies intercultural management tools.</p>	<p>The LGID demonstrates expert-like diagnostic capabilities and goes beyond what is taught in the classroom. He/she articulates regional initiatives; creates regional and inter-regional links in addition to creating links with institutions and with non-governmental organizations. He/she shows the potential of being self-employed.</p>
<p>Competency 2:</p> <p>“The student will be able to facilitate advocacy for, and management of, resources and information”</p>	<p>The LGID knows institutions, associations and people. The way he introduces and knowledge of methodology is basic. Starts to know what the LGID does. He/she identifies problems in the community and family and imagines (crafts) possible interventions.</p>	<p>The LGID knows more institutions and identifies projects. He/she knows the basic components of a project.</p>	<p>The LGID designs and carries out his/her project, and evaluates its impact. He/she knows the foundations of NGOs and institutions and their resources, and connects all this with his/her project. He/she is able to advocate.</p>	<p>The LGID works with financial and human resources. He/she designs projects from the “bottom up” and not “top down.” He/she not only advocates but empowers others in the community, teaching them to advocate. He/she intervenes in returning the empowerment to the people.</p>

				He sees projects through and serves as an intermediary in conjunction with others. He/she practices in the real world.
Competency 3: “The student will be able to strengthen and design intercultural initiatives”	The LGID identifies and knows his/her community and their problems/issues. He/she knows how to be observant and how to evaluate problematic situations (does “diagnostics”). He/she identifies initiatives based on the problematic situation. He/she does not fully strengthen initiatives because he does not know how to implement them yet. He/she is able to see areas of lacking and weaknesses.	The LGID is able to identify with more depth his/her community and its problems/issues. He/she identifies focal groups that can support him/her as he/she carries out initiatives. He/she designs his/her initiatives and identifies tools to carry them out.	The LGID inquires sharply and deeply what the needs/issues of the community are, and implements his/her initiatives with the community group.	The LGID promotes and facilitates solutions, not only understands problems/issues. He/she inserts him/herself and feels totally part of the community. He/she is able to insert him/herself in the job market.
Competency 4: “The student will be able to highlight regional/indigenous knowledge and initiatives”	The LGID starts to visualize and compare regions. He/she recognizes potentials in a community. He/she does not know and therefore judges from his/her cultural point of view the different types of knowledge (indigenous, academic). He/she begins with talks about how to disseminate information.	The LGID has demonstrated awareness of the problems/issues of the community and the different types of knowledge (indigenous/academic) of the people. He/she is more analytical and does not judge. He/she looks for different initiatives to be part of.	The LGID does not judge but rather values the different types of knowledge of the community. He/she criticizes and analyzes in proposal-like way and with theoretical and practical foundations that are reflected in practice. He/she produces materials such as brochures or videos	The LGID sees preliminary results and attempts to extend towards new project proposals. He/she is able to re-design plans of action of the community. He/she is based on actions. He/she is able to go beyond of what was learned in the classroom and is open to new knowledge(s). He/she sees his/her work totally

			to disseminate information. He/she is able to design workshops.	“horizontal” among advocates/intercultural managers and the community.
<p>Competency 5:</p> <p>“The student will be able to generate diagnostic and proposal-oriented knowledge”</p>	<p>The LGID begins to know the cosmo-vision of the community- this includes the “what” and the “how” they think. He/she knows the concepts of their myths, cultural practices and ways of living. Based on that he is able to create community diagnostics on cultural, social, economic, and political issues. He/she is able to identify problems.</p>	<p>The LGID applies methodology/techniques, valuing the existing types of knowledge. He/she fosters valuing those types of knowledge. The LGID is able to generate new knowledge without forgetting or considering of less value the knowledge that had already been established.</p>	<p>The LGID justifies and searches for the arguments on the why it is necessary to strengthen the initiatives with which he/she is involved. As he/she gains more knowledge, he/she values more his/her own culture and that of others. The student knows and disseminates the different types of knowledge, making visible the interconnectedness and the value of those types of knowledge. The LGID practices genuine interculturalism.</p>	<p>The LGID is constantly advocating for knowledge; he/she knows that the communities are the owners of the knowledge(s) and does everything with them as a group (the student is included). The LGID systematizes knowledge and disseminates it (makes it “visible”). He/she proposes solutions or searches for the most viable ones. He/she identifies the most feasible projects to carry out in the practice. The LGID is able to develop public policies. The LGID is able to advise/serve as a consultant to whoever comes to him/her for services.</p>

* LGID are the Spanish initials for the student who is an IMD major:

“Licenciado en Gestión Intercultural para el Desarrollo”

Appendix K

Letter to Scorers and Scoring Sheet

Appendix K

Letter to Scorers and Scoring Sheet

Dear :

Thank you very much for agreeing to participate as a scorer in my study. These are the guidelines of your participation and the materials attached:

A) **My Dissertation Summary**, to help you contextualize my work.

B) **A rubric**, developed by two different groups of seniors from the Universidad Intercultural (IU) in Mexico. The first group of seniors contributed by collectively indicating what each level of attainment look should like. I gave them a blank rubric and I had one too to compile their input. They filled out their rubrics, discussing among themselves what should go on mine (the collective one). I typed all up and sent it to them for member-checking purposes. Once they approved the first draft, I met a month later with another group of seniors who served as “expert judges.” I gave them a copy of the rubric their peers developed, a Table of Specifications (to estimate content validity based on consensus), and a copy of what the competencies of their program of study where. I then incorporated the second group’s revisions to the rubric and sent it to them for member-checking.

C) The desired **competencies** that graduates should attain at the end of their program of Study

D) **3 documentos recepcionales (first round)**, which are projects that seniors submit to graduate from the Universidad Intercultural. These students graduated with a BA in Intercultural Management for Development, and their performance-based projects range from ethnographies, catalogues, workshops, and manuals based on needs/problems/potential projects they observed in their communities. I have removed the pages where the student(s) names appear for confidentiality purposes. The documents are in Spanish and you might find a few words in other languages such as Náhuatl or Popoluca, but the meaning should be clear from the context.

E) **Your Scores Report**: Please use this form to report your results and comments. Your Scorer ID is " .”

The projects you will score do not follow the traditional thesis model in terms of chapters, research questions, and other components. Although your scores do not affect the graduates in any way or form (this is post-facto scoring), your input and effort will help me estimate the inter- and intra-rater reliability of the rubric the students developed. Your creativity, flexibility, and global perspective when scoring these projects are

pivotal. The questions to keep in mind are: **Is this project reflecting the attainment of the competencies of the Intercultural Management for Development's major? If so, at what level: basic (1 point), developing (2 points), satisfactory (3 points), or exemplary (4 points)? ¿Usted opine que a la comunidad les sirva este proyecto? ¿Podrían vender/desarrollar sus productos así? ¿O ponerlos en internet? Muchachos jóvenes aprendiendo el oficio, ¿podrían usar este documento? ¿Puede diseminarse el conocimiento de ellos con este documento?**

Deadlines:

JULY 23, 2012: Submit scores of your first three papers and your comments/feedback about the process

- **AUGUST 6, 2012:** You will receive three papers again. Two will be new and one will not. The repeated one is to estimate intra-rater reliability. Please score it again as if new.

- **AUGUST 13, 2012:** Submit scores of your three papers and any comments about the process.

- If the two of the scores I receive from different raters are not discrepant, the final score will be an average of the two raters' scores. In case there are discrepancies of more than one point in scoring among raters, there will be a third round of scoring. I will notify you, if this occurs.

Finally, to show that I selected highly qualified scorers, I would ask that you please send me your résumé (your scores will be confidential).

Feel free to contact me should you have any questions or comments. Your feedback to me as a junior researcher is very valuable.

With much gratitude,

Fernanda

Estimado :

Quisiera saber si les gustaría pilotear la matriz de descripciones que hicieron los muchachos en la intercultural que visité al principio del año, como parte de mi tesis doctoral. Como sabe, los documentos recepcionales de los chicos deben reflejar el logro de las competencias de su programa de estudio y quiero saber si la matriz (rúbrica) nos ayuda a detectar esto. Como usted sabe, todo instrumento requiere tener estimados de

validez y confiabilidad aceptables. Ya los muchachos hicieron la validez de contenido y ahora, como parte de mi estudio, quisiera calcular la confiabilidad de la matriz, con inter-e intra-evaluadores. Esto no va a tomar mucho de su tiempo. Su participación sería primero “calificar” 3 documentos recepcionales cortos usando la matriz (en julio) y luego, 2 semanas después, otros 3 documentos (uno va a ser el mismo, pero se califica como nuevo). Todas estas “calificaciones” son post-facto y no afecta a los autores de ninguna manera. Después ustedes me mandan sus resultados y su retroalimentación sobre la matriz y sobre el proceso en general. El objetivo y el enfoque de este ejercicio es pilotear el instrumento nada más. Espero que les interese participar y de ser así, que me dejen saber su respuesta lo más pronto que pueda para poder comenzar.

De antemano muchas gracias,

Fernanda

Reporting Form

SCORING SESSION #:

SCORING ID:

DR:	SCORES					TOTAL	COMMENTS
	Compet. 1	Compet. 2	Compet.3	Compet. 4	Compet. 5		

Feedback:

APPENDIX L

Scorers' Résumés

Azucena Caro

I. GENERAL INFORMATION

NAME: Caro Dueñas María Azucena

II. WORK EXPERIENCE

Area: Social Sciences and Humanities

Discipline: Law and Education

Present Appointment: Educational Facilitator and Director of Tutoring and Solidarity Work

III.- PROFESSIONAL MEMBERSHIPS:

- a) Founding and active member of Los Mochis' Board of Lawyers, "Dr. Diego Valadés A.C."
- b) Active member of the Confederation of Colleges and Associations of Lawyers and the Lawyers' Federation of Sinaloa.

IV.- PREVIOUS EMPLOYMENT

Year	Position
2011-2012	Educational Facilitator and Director of Tutoring and Solidarity Work
2001-2005	General Lawyer; Instructor of the following subjects: Constitutional and Indigenous People's Law, Tourism Law, IT Law, Selected Topics of IT Law, Goods Civic Law, Contracts Civic Law.
1999-2001	Instructor of the following subjects: development of thought abilities, Constitutional and Indigenous People's Law, and Agrarian Law.

V. EDUCATION

Level	Program	Institution	Year Completed
Bachelor's	Law	Universidad Autónoma de Sinaloa	1999
Master's	Social Education	Universidad Autónoma Indígena de México	2010
Master's	Constitutional Law and Protection	Universidad Autónoma de Sinaloa	

UNDERGRADUATE THESES CHAIRED

Name of the Theses	Institution	Year
Glosary of Selected Judicial Terms Spanish-ch'ol to facilitate the impartation of Justice for the Indigenous Groups of Tila Chiapas.	Universidad Autónoma Indígena de México	2010
Analysis of the Award of Damages for the Victims of a Crime in a Trial in Ahome, Sinaloa	Universidad Autónoma Indígena de México	2010
Analysis of Sexual Assault Crimes in Mexico	Universidad Autónoma Indígena de México	2010
The Impact of intestate succession on the Family	Universidad Autónoma Indígena de México	2009
Analysis of the Agrarian Conflict of Delimiting Land between Nayeri (Coras) of Santa Cruz de Guaybel and the Presidio de los Reyes Mpio. de Ruiz, in the State of Nayarit	Universidad Autónoma Indígena de México	2010

Daniel Sergio Friedrich
Assistant Professor, Dept. of Curriculum & Teaching
Teachers College, Columbia University
Friedrich@tc.edu

Education

PhD in Curriculum and Instruction, minor in Latin American, Caribbean and Iberian Studies, 2010.

University of Wisconsin-Madison, WI

Dissertation: *Creating a Past in the Present. Memory, Identity and Teaching in Post-Dictatorship Argentina.*

Advisor: Dr. Thomas S. Popkewitz

GPA: 4.0

B.A. [Licenciatura] in Educational Sciences, minor in Philosophy, 2003.

Universidad Nacional de Buenos Aires, Buenos Aires, Argentina

Graduated with Honors

GPA: 3.50

Teaching Certification for Primary Schooling (first through seventh grades), 2000.

Escuela Normal Superior Nro. 10, Buenos Aires, Argentina

GPA: 3.70

Research and Teaching Interests

I am interested in epistemological and philosophical questions, as they relate to the politics of schooling and of teacher education. My background allows me to engage in interdisciplinary, comparative and international perspectives, particularly with a focus in Latin America.

Conference Papers / Publications

Publications

Peer-reviewed journals

Friedrich, D. (Under Review). Discourse On (Teaching) Method: Challenging the Reason of Contemporary Teacher Education. *Education et Sociétés*.

Friedrich, D. (Under Review). "We brought it upon ourselves": University-based Teacher Education and the Emergence of Boot-Camp-Style Routes to Teacher Certification. *Teachers College Record*.

Friedrich, D. (2011). The Memoryscape in Buenos Aires: Re-presentation, Pedagogy and Memory. *Journal of Curriculum Theorizing*, 27(3), 171–189.

Friedrich, D. (2010). An Atheist('s) Manifesto. Commentary on Marc Dapaepe's "The Ten Commandments of Good Practices in History of Education Research." *Journal for the Historiography of Education*, (1).

Friedrich, D. (2010). Historical Consciousness as a Pedagogical Device in the Production of the Responsible Citizen. *Discourse: Studies in the Cultural Politics of Education*, 31(5), 649-664.

Friedrich, D., Jaastad, B., & Popkewitz, T. S. (2010). Democratic Education: An (Im)Possibility That Yet Remains To Come. *Educational Philosophy and Theory*, 42(5-6), 571-587.

Popkewitz, T. S., & Friedrich, D. (2008). Escuelas de Desarrollo Profesional: narrativas de democracia, tesis de redención, ¿y negación de la política? *Revista Educación y Pedagogía*, XX(50), 157-176.

Books and chapters

Friedrich, D. (Under Review). *The Limits of Democratic Education as a Curricular Problem*. Routledge.

Friedrich, D. (Forthcoming). The Mobilization Of Historical Consciousness In The Narratives About The Last Argentine Dictatorship. In J. Williams (Ed.), *(Re)Building Memory: School Textbooks, Identity, and the Pedagogies and Politics of Imagining Community*. Sense Publishers.

Friedrich, D. (forthcoming). The Unfinished Cosmopolitan as the Embodiment of the Paradoxes and Promises of Democratic Education. In B. Franklin & M. Pereyra (Eds.), *Systems of Reason and the Politics of Schooling: Alternatives Studies on School Reforms and Sciences of Education in tradition of Thomas S. Popkewitz*. Routledge.

Friedrich, D., Jaastad, B., & Popkewitz, T. S. (2011). Democratic Education: An (Im)Possibility That Yet Remains To Come. In M. Simons & J. Masschelein (Eds.), *Rancière, Public Education and the Taming of Democracy* (pp. 60–75). Wiley-Blackwell.

Conference Papers

““Enseña por ...” (complete el espacio en blanco). Formación docente en la época de su reproductibilidad técnica,” Latin American Studies Association Meeting, San Francisco, 05/2012.

“Teach for... (Fill in the Blank). Teacher Education in the Age of Technical Reproduction,” Annual Conference of the CIES, Puerto Rico, 04/2012.

“Pedagogies of Torture, Pedagogies as Torture: A Curricular Palimpsest,” AERA Annual Meeting, Vancouver, BC, 04/2012. (With C. Goulding and M. Walters)

“The Mobilization Of Historical Consciousness In The Narratives About The Last Argentine Dictatorship,” 55th Annual Conference of the Comparative and International Educational Society, Montreal, ON, 05/2011.

“Resisting Resistance,” AERA Annual Meeting, New Orleans, LA, 04/2011

“Educating a historical consciousness as a strategy to produce the Argentine citizen,” AAACS Annual Meeting, New Orleans, LA, 04/2011

“Nueva crisis, viejos problemas: aproximaciones y resignificaciones desde el campo educativo en la Argentina” (Symposium, president and participant), Latin American Studies Association Meeting, Montreal, 10/2010.

“Legislating Progress: Argentine Education Laws and the Framing of Darkness”, 54th Annual Conference of the Comparative and International Educational Society, Chicago, IL, 03/2010.

“The Paradoxes of a Democratic Education: Rancière and the Question of Equality.” AERA Annual Meeting, Denver, CO, 05/2010 (with Bryn Jaastad).

“Democratic Education: An (Im)Possibility That Yet Remains To Come.” 4th Biennial Provoking Curriculum Conference, Ottawa, Canada, 05/2009 (with Bryn Jaastad).

“Creating a Past in the Present. Memory, Identity and Teaching in Post-Dictatorship Argentina.” AERA Annual Meeting, San Diego, CA, 04/2009.

“Argentinity, the Mobilization of the Recent Past and the Responsible Citizen.” 29th Annual Bergamo Conference on Curriculum Theory and Classroom Practice, Dayton, OH, 10/2008.

“De Víctimas y Resistencias. La Dictadura en las Aulas” [Victims and resistances. The dictatorship in the classroom]. VII Congreso Iberoamericano de historia de la educación latinoamericana, Buenos Aires, Argentina, 10/2007.

“De Víctimas y Resistencias. El Pasado Reciente en las Escuelas Argentinas” *Victims and resistances. The recent past in Argentine schools]. Mid-America Conference on Hispanic Literature, Madison, WI, 10/ 2007.

“The Construction of Argentinity Post-2001 and the Debates for a New National Education Law.” World University Network (WUN) Conference. Bergen, Norway, 09/2006.

Honors and Awards

Provost's Investment Fund, "A 'Course Staff' Model for Teaching Assistants", AY 2012-2013. (Faculty group led by D. Hansen).

Dean's grant for Pre-Tenured and Non-Tenure Track Faculty, Teachers College, Summer 2012.

Vilas Travel Grant, Graduate Student Collaborative, University of Wisconsin – Madison, 11/2009.

Dissertator Travel Funds, Dept. of Curriculum and Instruction, University of Wisconsin – Madison, 10/2008.

Tinker-Nave Field Research Grant, 06/2008.

Teaching Experience

Teachers College, Columbia University, New York, NY.

Assistant Professor, Dept. of Curriculum and Teaching, 2010-present

Curriculum Policy

Curriculum Theory and History

Doctoral Seminar: Memory, History and Curriculum

Seminar: Globalization, Democracy and Curriculum

Critical Perspectives in Elementary Education

Masters Action Research Project

University of Wisconsin, Madison, WI.

Graduate Teaching Assistant, 2006 – 2010.

Spanish Education Supervisor for C&I 442, 443: Student Teaching in World Languages (K-12), Spring 2006 – Spring 2009.

Instructor for C&I 340: Elementary Education Practicum One: Practicum in Community and Neighborhood Settings, Fall 2006.

Universidad Nacional de Lomas de Zamora, Buenos Aires, Argentina

Teaching Assistant, 2004.

Study Methodology course for incoming law students

Family College, Buenos Aires, Argentina

Professor, 2003 – 2004.

Learning Theories

School Organization and Administration

Universidad Nacional de Lomas de Zamora, Buenos Aires, Argentina

Teaching Assistant, 2003.

Seminar: Post-Structuralist Perspectives in Education

Nuestro Lugar Primary School, Buenos Aires, Argentina,

Teacher, 2002 – 2005.

Social Sciences, Science, Spanish and Math Teacher for 5th, 6th and 7th grades

English Teacher for 2nd grade

Related Work Experience

Pestalozzi High School, Buenos Aires, Argentina

Advisor, 2004 – 2005.

Coordinator of the Program for Solidarity Actions in rural schools

Service

Co-Chair, Latin America SIG, CIES, 2011-2013.

Chair for the session “Democratization of, Equity in, and Access to Education Around the World.” 2010 AERA Annual Meeting.

Reviewer, Education Policy Analysis Archives

Reviewer, *Polity* journal.

Reviewer, *Journal of Curriculum Theorizing*.

Reviewer, 2010-2012 AERA Annual Meetings.

CURRICULUM VITAE

EDUCATION

Posdoctorate	Fractal Geometry in the College of Architecture, Autonomous University of Nuevo Leon, 2008.
Graduate Degrees	Doctorate in Higher Education Center for Research and Teaching Humanities in the State of Morelos, 2000-2006. Graduated with honors.
	Master's in Industrial Economics School of Economics of the Autonomous University of Nuevo Leon, 1991-1993.
Profesional	Bachelor of Science in Physics College of Physics and Mathematics, Autonomous University of Nuevo Leon, 1980-1984.

EXPERIENCE

Since 1980 I have taught physics, mathematics, statistics, economics, and education to undergraduate and graduate students at numerous institutions. Among these institutions are Technological Institute of Higher Education of Monterrey (1980-1988), College of Physics and Mathematics (1991-1992) and the School of Public Accounting and Administration (1993-1994) of the Autonomous University of Nuevo Leon; Universidad of Occidente (since 1997) and the Autonomous Indigenous University of Mexico since 2000.

ACADEMIC PRODUCTIVITY

I have directed more than 50 bachelor's and master's theses and I have published over ten scientific articles in magazines and edited books. I am currently the ruling officer (since 2005) of the *Convergencia*, produced by the School of Political Sciences and Public Administration of the Autonomous University of the State of Mexico.

PROFESSIONAL EXPERIENCE

As Educational *Administrator*:

October 1992 to October 1993	Statistics Project Director in the Statistics Division of the Coordination and Program Development Department of the Ministry of Education of the State of Nuevo León.
1993-1994	<i>Member of the Consulting Board</i> in the Committee for Planning and Program Development for Academic Excellence in the Faculty of Public Accounting and Administration in the University of Nuevo Leon.
1994	Representing Member of the Consulting Board in the Committee for Planning and Program Development for Academic Excellence in the Faculty of Public Accounting and Administration in the University of Nuevo Leon at the <i>Texas International Education Consortium</i> , Austin Texas.
October 1995 to October 1998	<i>Coordinator of the Educational Development Center of the Colegio de Bachilleres</i> in the State of Sinaloa (North Zone).
September 1999 to June 2001	Academic Coordinator (designing and implementing) of the Masters in Business Administration of the Universidad of Occidente, Unit Guasave Sinaloa.
October 1998 to December 2001.	Academic Development Director of the Universidad of Occidente
September 1999 to December 2007.	Helped design and create the IU, and served as the first Coordinator.
January 2008 - present	General Coordinator of Research at the IU

FLAVIA E. IUSPA

OVERVIEW

Highly effective **Educational Program Executive** with extensive experience in international programs and multi-cultural overseas environments. Over 5 years of experience managing graduate programs to increase program offerings, community partnerships, and budget compliance. Doctoral degree with specialization in the internationalization of HEIs. Leverages thorough knowledge of higher education system. International background with full fluency in English and Spanish plus proficiency in Italian and French.

- Program Management
- Recruitment & Admissions
- Partnership Development
- Staff Management
- Budget Management
- Accreditation

SELECTED ACCOMPLISHMENTS

- Coordinated logistics of Masters' programs in Jamaica, Bolivia, and the Dominican Republic, as well as new programs.
- Launched Advisory Board in Jamaica by tapping key Jamaican businesspeople to contribute networking, marketing, and fundraising support.
- Promoted university goal to reach out to international community by establishing partnership in Jamaica with University College of the Caribbean (UCC) and in the Dominican Republic with UNIBE; leveraged diplomacy and mediation in establishing each party's goals.
- Launched MBA program for Public Managers; managed program approval, recruitment, admissions, and registration; coordinated outreach with key demographic of Miami-Dade County employees.
- Consistently attained targets through effective preparation and management of annual budget of over \$200,000.
- Prepared prospectus for Southern Association of Colleges and Schools (SACS) to support launch of new offshore Master of Science in Human Resource program.
- Increased profile and awareness of offshore programs by maintaining constant dialogue with university faculty and administrators as well as community and industry stakeholders.

EDUCATION

FLORIDA INTERNATIONAL UNIVERSITY, Miami, FL

Doctor of Education, 2010

- Specialization in International and Intercultural Development Education
- Dissertation: Assessing the Internationalization Process of a Higher Education Institution (HEI): A Case Study of Florida International University

International Masters in Business Administration, 2000

Bachelor of Arts in International Relations, 1997

- Study Abroad, Nanterre Université, France

Professional Development:

- Human Resource Certificate, Center for Management Development, Florida International University, Miami, FL.

PROFESSIONAL EXPERIENCE

FLORIDA INTERNATIONAL UNIVERSITY, Miami, FL 2000-2007, 2009-Present

Director (2012-Present)

Assistant/Acting Director (2011- 2012)

Program Coordinator, College of Education (2011-08/2011)

Graduate Assistant, College of Education (2009-2010)

- Manage Master of Science in Curriculum and Instruction (MSCI) program in Jamaica. Program management and administrative tasks includes (but not limited to): Recruiting, scheduling, admission, budgeting, accreditation, and partners' relationships, among others.

Program Manager, Global Programs Office, Chapman Graduate School of Business (2002-2007)

- Promoted to manage all aspects of program management and operations for overseas graduate program.
- Oversaw admissions decisions for over 125 Jamaican students, 41 Bolivian students, and 24 Dominican Republic students.
- Served on e-learning panel at Effective Internet Strategies Conference organized by EISalvador.com.
- Planned and coordinated all program events, such as orientation, graduations, and networking sessions.

- Participated in developing program-marketing strategies in collaboration with school marketing staff.
- Prepared Substantive Change Prospectuses for offshore programs.

**Assistant Director, Global Programs Office & Office of Professional Education,
College of Business Administration (2000-2002)**

- Coordinated logistics of masters' programs in Jamaica, Bolivia, and France and managed 191 international students.
- Hired and supervised office staff, including part-time assistants, and liaised between faculty and office staff.
- Planned and executed marketing of new executive programs, including creating marketing collateral, making presentations to recruit students, and managing program web sites.

FLORIDA INTERNATIONAL UNIVERSITY, Miami, FL

Academic

Adjunct Faculty, Department of Teaching and Learning May 2011- August 2011
Taught Curriculum and Instruction master and doctoral courses.

Instructor, Department of Teaching and Learning August 2011- Present
Taught Curriculum and Instruction undergraduate, master and doctoral courses.

PUBLICATIONS

- Iuspa, F. (April 2005). Japan's Educational System. Proceeding of The Fourth Annual College of Education Research Conference (pp 54-58). Miami, FL: Florida International University.
http://education.fiu.edu/research_conference/docs/proceedings/2005_COERC_Proceedings.pdf
- Iuspa, F. (2004). "E-learning in Universities" panel. Effective Internet Strategies Conference, organized by ElSalvador.com. El Salvador, El Salvador.

Alejandra Leon, Ph. D.

EXPERIENCE

Spanish Teacher

Welsh Hills School, Granville, OH Sep '2010-Present

Teaching Spanish throughout every grade from Toddler to Middle School.

Organization of Spanish and multicultural lessons.

Spanish and French Tutor

Central Ohio Technical College, OH Mar 2009-Sep 2009

Tutoring Spanish and French to students of Spanish 101, 102, and 103; and French 101, 102, and 103; in cooperation with Professors from the Foreign Language Department.

Graduate Research Assistant

The Ohio State University, Columbus, OH Aug '07 – July 2010

Database processing with Ethnograph V.5 and V.6. Review of book translation from Spanish into English.

Conference presentation at AAA Annual Meeting, 2009.

Writing of scholarly articles. Linguistics and discourse analysis applied to transnational communities.

Graduate Teaching Assistant

The Ohio State University, Columbus, OH Mar '04 - Jul '07

Instructor of School and Society, 306 in the School of Educational Policy and Leadership. Continuous

adaptation of course plan to each cohort interests and needs. Application of knowledge acquired in parallel with Graduate coursework and while attending to different conferences on Comparative Education.

EDUCATION

Doctor of Philosophy (Comparative Educational Policy)

The Ohio State University, Columbus, OH Mar '04 – Mar '11

Mexican Educational Policy Implementation: A Study on Outward Migration as a Social Influence in the Primary School Classroom. Dissertation in partial fulfillment for Doctorate Degree. Columbus, Ohio, United States: School of Educational Policy and Leadership, CEHE, The Ohio State University.

Master's Degree Received Aug '02 (Korean Studies)

Seoul National University, Seoul, South Korea

Changes on Korean Women's Status after the Establishment of the Ten Policy Priorities. Thesis for Master's Degree. Seoul, South Korea: SNU (Written in Korean). Institute of International and Area Studies.

Licenciatura Received Feb '98 (International Affairs)

Universidad de Guadalajara, Guadalajara, México

Efectos Comerciales del Embargo Atunero en la Comercialización Interna y Externa de este Producto (Commercial Effects of the Tuna Fish Embargo on the Mexican Market) Thesis for Licenciatura. Guadalajara, Mexico: U de G. Centro de Estudios Internacionales

LANGUAGE SKILLS

Native fluency in Spanish, French and English. Very high proficiency in Korean.

SCHOLARLY WORK

- Farr, M. & Leon Garcia, M. (December 2009) "Uchepos and Hamburgers: Food and Talk in Mexican Transnational Families. 2009 AAA Annual Conference. Philadelphia:AAA
- Leon-Garcia, M. A. (2002). El Confucianismo en Corea (Confucianism in Korea) México y la Cuenca del Pacífico, 5(17)P?. 47-52.
- Leon-Garcia, M. A. (2004). Excelencia Educativa Coreana: Un Modelo en Transición (Excellence in Korean Education: A Model in Transition). México y la Cuenca del Pacífico, 7 (22), pp. 33-39
- Leon-Garcia, M. A. (March, 2009). Globalized Education in President Calderón's Administration and its projected goals: Reality or Official Rhetoric? 2009 Annual CIES Conference. Charleston: CIES

- Leon-Garcia, M. A. (March, 2010). How do Mexican Teachers see their Educational Roles within US bound migrant communities? 2010 Annual CIES Conference. Chicago: CIES
- Leon-Garcia, M. A. (2000). La Mujer en la Sociedad Coreana (Women in Korean Society). México y la Cuenca del Pacífico (Mexico and the Pacific Rim),3(11), pp. 58-63.
- Leon-Garcia, M. A. (2008). Mexican Educational Policy Implementation: A Focus on Outward Migration as a Social Influence in the Primary School Classroom. First Conference on Childhood and Migration. Philadelphia: University of Pennsylvania.
- Leon-Garcia, M. A. (In progress). Mexican Educational Policy Implementation: A Study on Outward Migration as a Social Influence in the Primary School Classroom. Dissertation for Doctorate Degree . Columbus, Ohio, United States: School of Educational Policy and Leadership, CEHE, The Ohio State University.

Carlos Ornelas

Curriculum vitae

Academic Degrees. Ph. D. in education from Stanford University in 1980; MA in international education from Stanford University (1978).

Teaching Experience. Presently: professor of Education and Communications and coordinator of the PhD. Program in the Social Sciences at the Universidad Autónoma Metropolitana (UAM). Also taught at the Graduate Division of Political and Social Sciences Faculty of the National University of Mexico (1981-1990). Associate Professor of logic at the Universidad Juárez del Estado de Durango (1972-1975). Visiting positions: Visiting Professor of Comparative and Transcultural Studies at Teacher College Columbia University (2008-2009); Professor at the Virtual University of the Instituto Tecnológico de Monterrey (1998); Guest Professor of Political Studies at the Centro de Investigación y Docencia Económicas (CIDE), in Mexico City (1994-1995); Guest Professor at the Department of Education of Universidad Iberoamericana (1984-1985); and Visiting Lecturer and Fulbright Scholar at Harvard University Graduate School of Education (1986-1987).

Publications: Author of **Educación, colonización y rebeldía: la herencia del pacto Calderón-Gordillo** (México City: Siglo XXI Editores, 2012); **Política, poder y pupitres** (México City: Siglo XXI Editores, 2008, second edition 2010); **De la crisis a la reconstrucción: El cambio institucional en la Benemérita Universidad Autónoma de Puebla** (México City: ANUIES, 1999) and of **El Sistema Educativo Mexicano: La Transición de Fin de Siglo** (Mexico City: Fondo de Cultura Económica, 1995; twelfth reprint 2011). Editor of **Democracia y educación cívica** (México City: Instituto Electoral del Distrito Federal, 2007); **Buenas prácticas de educación básica en América Latina** (México City: Instituto Latinoamericano de la Comunicación Educativa, 2005); and of **Investigación y política educativas: Ensayos en honor de Pablo Latapí** (México City: Santillana, 2001). Coauthor of **La Metafísica de la eficiencia: Ensayos sobre la planeación universitaria en Mexico** (Mexico City: CISE-UNAM, 1991), and of **Superación académica y reforma universitaria** (Mexico City: UNAM, 1989). He also edited other four books.

Contributed chapters to **Diversidad cultural, desigualdad social y estrategias de políticas educativas** (Buenos Aires: Instituto Internacional de Planeamiento de la Educación-UNESCO, 2010); **El nuevo milenio mexicano** (México City: UAM-Azcapotzalco, 2004); **NAFTA's Impact on North America: The First Decade** (Washington, DC: The Center for Strategic and International Studies, 2004); **Carreras magisteriales, desempeño educativo y sindicatos de maestros en América Latina** (Buenos Aires: Facultad Latinoamericana de Ciencias Sociales-Argentina, 2002); **México en la globalización: condiciones y requisitos de un desarrollo sustentable y equitativo** (Mexico City: Fondo de Cultura Económica, 1996); **Universidad nacional y economía** (Mexico City: UNAM and Miguel Ángel Porrúa, 1990); **La ideología**

educativa de la Revolución mexicana (Mexico City: Universidad Autónoma Metropolitana, 1984); **Sociología de la educación: corrientes contemporáneas** (Mexico City: Centro de Estudios Educativos, 1981); as well as in other 27 other books.

Author of 30 articles published in professional journals (in Spanish, English and French; one of his papers was translated into Arabic, Chinese and Russian). In addition, 72 reviews, extended essays reviews and other professional articles. He is also author of 22 additional unpublished policy and research reports. He has weekly column in **Excélsior**, a Mexican national newspaper.

Professional Experience. He was advisor to the Secretary of Public Education (1984-1985 and 1996-1998); Founding Executive Director of the US-Mexico Commission for Educational and Cultural Exchange (the Fulbright Commission): 1991-1993. He was appointed as special advisor to the Secretary of Public Education (1989-1991). His first administrative post was as Founding Director of the Colegio de Ciencias y Humanidades of the Universidad Juárez del Estado de Durango (1973-1975).

He received the "W. G. Walker Best Paper Award" for his article: *Politics of educational decentralization in Mexico*, published in the **Journal of Educational Administration** (Vol. 38, No. 5, 2000) granted by the MCB University Press, in England. The Universidad Autónoma Metropolitana has honored him with the accolade for research and teaching, at the highest level in eleven different years.

Other Activities. Presented talks and papers in over 200 academic, professional, and policy-making gatherings worldwide. He was the director of a collection of books, **AulaXXI/Mexico**, for teachers; 20 titles were published in 2001-2004.

Carlos Ornelas is a member of the Comparative and International Education Society (since 1982). He also is affiliated to the National System of Researchers of Mexico; the Consejo Mexicano de Investigación Educativa, and the Latin American Studies Association. He was an associate fellow of the Club of Rome.

March 2012

Jacqueline Peña

EDUCATION

Ph.D., Curriculum & Instruction: Language, Literacy, and Culture, 2010

Florida International University, Miami, FL

Dissertation: Engagement Experiences of Generation 1.5 English Language Learners at a Massachusetts Community College

Advisor: Dr. Eric Dwyer

Graduate Certificate, Institutional Research, 2010

Florida State University, Tallahassee, FL

MA, English Literature, 2002

Boston College, Boston, MA

BA, English Literature, Summa Cum Laude, 2000

Herbert H. Lehman College, New York

Minors: Secondary Education and Italian

ADMINISTRATIVE EXPERIENCE

COLLEGE PREP DEPARTMENT CHAIRPERSON, AUGUST 2010—PRESENT

Miami Dade College, Wolfson Campus, Miami, FL

Hire, supervise, and evaluate full-time and part-time faculty, lab instructors, and tutors as well as support staff for the School of Arts and Letters and the IDEA Center (academic support lab).

Establish and evaluate departmental policies and procedures. Provide professional development opportunities to all employees. Implement new programs such as cohorts for developmental education students and mandatory orientations. Resolve issues between faculty, staff, and students. Participate on multiple campus and college-wide committees and projects. Assist with the development of campus and college programs including the American Dream Scholars cohorts for college ready students, the Pathway to Academic Excellence and Confidence cohort program for developmental reading and writing students, and the IDEA (Inter-disciplinary Educational Achievement) Center for Wolfson students. Teach EAP and developmental writing courses. Collaborate with constituents in the external educational, professional, and business communities, including the feeder high schools, local universities, and community colleges throughout the state.

ASSESSMENT COORDINATOR, OCTOBER 2008—AUGUST 2010

**Florida International University, Office of Academic Planning and Accountability,
Miami, FL**

Developed and implemented institutional assessment processes for the university. Assisted academic and student services units in developing effective assessment plans, reports, and improvement strategies. Assisted with the university's regional accreditation process. Completed annual Board of Governor's Academic Learning Compacts reports. Provided professional development to the faculty, staff, and administrators on developing effective assessment processes, valid and reliable assessment tools, and sustainable improvement strategies.

INTERIM DEPARTMENT CHAIR, 2004—2005

**Northern Essex Community College, Department of Developmental Studies, Haverhill,
MA**

Worked with the department's four student support centers (Writing Center, Reading Center, Math Lab, and English Language Center), the Adult Basic Education area, and the ESL, Developmental Writing, Reading, and Developmental Math programs. Disseminated institutional information to faculty and staff. Assisted with the department's documentation of student learning goals and objectives. Provided feedback on department reports for the four subject areas. Collaborated with non-profit literacy organizations for the smooth transition of their students to the college. Assisted with professional development for faculty.

CURRICULUM AND ASSESSMENT DEVELOPER, 2003

Centro Educacional Hermanas Mirabal, funded through a HUD grant, Lawrence, MA

Developed a two-level ESL curriculum for female childcare providers in the Merrimack Valley. Developed placement exams, course-level assessments, bilingual informational booklets, registration processes, and supplemental course materials. Developed processes to bridge this grant-funded community program with the community college academic ESL program.

TEACHING EXPERIENCE (HIGHLIGHTS)

WRITING SEMINAR INSTRUCTOR, 2007—2012

**Florida International University, College of Education, Office of Field Experience,
Miami, FL**

Assisted with the development of a new teaching internship application process. Currently teach

and assess a required two-hour writing seminar for all teacher preparation students applying for student teaching (approximately 400 students a year).

ADJUNCT INSTRUCTOR, EDUCATION, 2006—PRESENT

**Florida International University, College of Education, Curriculum and Instruction,
Miami, FL**

Teach graduate and undergraduate education courses. Complete student learning outcomes assessments via TaskStream (online student learning assessment tool), and develop internet-based course supplements (e.g., Facebook pages, BlackBoard course modules).

ESL AND BASIC WRITING PROFESSOR (TENURE TRACK), 2001—2005

**Northern Essex Community College, Department of Developmental Studies, Haverhill,
MA**

Taught developmental writing, freshmen composition, and ESL courses. Served on college committees, completed institutional research, and assisted with curriculum revision. Mentored minority students. Advised and registered a minimum of 50 students a semester. Administered ESL assessments. Assisted with the revision of institutional placement tests.

TEACHING FELLOW, 2001—2002

Boston College, English Department, Boston, MA

Taught freshmen composition, which included developing and implementing the course syllabus and assessment processes. Assisted Dr. Henry Blackwell with the instruction and assessment of students in an undergraduate African-American literature course.

ESL ADJUNCT INSTRUCTOR, 2000—2001

**Northern Essex Community College, Department of Developmental Studies, Haverhill,
MA**

Taught ESL communication and grammar to day and evening students on the main campus and the extension campus.

WRITING CENTER TUTOR, 2000—2001

**Northern Essex Community College, Department of Developmental Studies, Haverhill,
MA**

Oversaw the Writing Center on the college's extension campus and tutored 40-100 students each

*week. Maintained center and student records. Revised outdated resources and developed new ones, such as writing manuals and handouts. Designed and edited **The Elliott Street Journal**, a faculty and staff journal. Edited **Writers in Progress**, the annually published book of student papers used as the Basic Writing text.*

WRITING TUTOR, 2000—2001

Boston College, Student Support Center, Boston, MA

Assisted undergraduate students with their reading and writing skills. Assisted graduate students with their research papers, master theses, and doctoral dissertations.

RESEARCH EXPERIENCE

RESEARCH COORDINATOR, EARLY READING FIRST GRANT, APRIL 2007—SEPTEMBER 2008

Florida International University, College of Education, Miami, FL

Title: Learning Educational Approaches to Reading Now

Principle Investigators: Dr. Laura Dinehart, Florida International University, Miami, FL

Dr. Sylvia Palenzuela, Early Learning Coalition of Miami-Dade/Monroe County

Assisted with the research design of the multi-site assessment project of a pre-k 4 literacy curriculum in 5 early childhood learning centers. Coordinated data collection, entry, analysis, and reporting. Administered pretests and posttests (Peabody Picture Vocabulary Test, Expressive One-word Vocabulary Test, Receptive One-word Vocabulary Test, Test of Preschool Early Literacy, and the Phonological Awareness Literacy Screening Test). Supervised students and employees working on the project. Coded and analyzed all qualitative data. Assisted with the analysis and reporting of all quantitative data.

RESEARCH ASSISTANT, 2007

Florida International University, College of Education, Miami, FL

Title: Using Rap to Increase Reading Fluency in African-American Students

Title: Principle Investigator: Cynthia Janeska

Assisted the principal investigator with the development of a second-grade reading intervention and implemented the intervention in four of the eight research classrooms in two of Miami's

public schools.

RESEARCH ASSISTANT, EVEN START GRANT, 2005—2007

Florida International University, College of Education, Miami, FL

Title: Educating Mothers and Babies in Early Reading Strategies (EMBERS)

Principle Investigator: Dr. Lisbeth Krauss

Assessed the students participating in the grant-funded project using standardized language and literacy tools such as the E-LAP and LAPR. Assessed the parents' English language skills and developed an ESL curriculum to meet their language learning needs. Assessed the effectiveness of the teachers administering the early childhood curriculum. Entered and analyzed assessment and research data.

RESEARCH ASSISTANT, MURMSI GRANT, 2005—2007

Florida International University, College of Education, Miami, FL

Title: Science Academic Vocabulary for English Language Learners Project

Principle Investigators: Dr. Eric Dwyer and Dr. Lisbeth Krauss

Developed, implemented, and assessed a curriculum intervention for fifth grade Spanish-speaking English Language Learners (ELLs) taking science in a bilingual public school. Administered the APRENDA and C-SAVE close tests. Entered, analyzed and reported on assessment/research findings.

CLASS-ACTION RESEARCHER, CARNEGIE ACADEMY FOR THE SCHOLARSHIP OF TEACHING AND LEARNING, 2004—2005

Northern Essex Community College, Haverhill, MA

Title: Service-learning for ESL Students at a Community College

Developed and carried out an action research project on using service-learning for ESL students in order to enhance English language skills while helping them gain experience in areas related to their desired academic majors. Reported on the study to the college community and to the Carnegie cluster.

CO-RESEARCHER, PROCESS MANAGEMENT TEAM, 2004—2005

Northern Essex Community College, Haverhill, MA

Focus: ESL Student Retention at Northern Essex Community College

Completed institutional research with a faculty/staff team on ESL student retention through various data collection methods (e.g., student and faculty focus groups, phone interviews, and surveys). Assisted with data analysis and reporting to the college community and to the Board of Trustees as well as the development of action plans for ESL student retention improvement during and after ESL coursework

PUBLICATIONS AND REPORTS

Dinehart, L., Peña, J., & Larralde, A. (2008). *Environmental assessment report: 2007-2008*. Miami, FL: Early Reading First, Early Learning Coalition of Miami-Dade/Monroe.

Dinehart, L., Chae, C., Peña, J., Britt, Y., & Lora, A. (2007). *Child language and literacy assessment report*. Miami, FL: Early Reading First, Early Learning Coalition of Miami-Dade/Monroe.

Dinehart, L., Chae, C., Peña, J., Britt, Y., & Lora, A. (2007). *Environmental assessment report*. Miami, FL: Early Reading First, Early Learning Coalition of Miami-Dade/Monroe.

Dixon-Krauss, L., Dwyer, E., Peña, J., Chae, C., & Grigorescu, C. (submitted 2012). An inquiry into academic vocabulary learning in elementary science classrooms and bilingual literacy development. *TESOL Journal*.

Peña, J. (2008). Moving to the front of the classroom: English graduate students as composition instructors. In M. S. Plakhotnik & S. M. Nielsen (Eds.), *Proceedings of the Seventh Annual College of Education Research Conference, Urban and International Education Section* (pp. 111-116). Miami, FL: Florida International University. http://coeweb.fiu.edu/research_conference/

Peña, J. (2005). *Scholarship of Teaching and Learning Report: Service-learning with my ESL students*. Haverhill, MA: Northern Essex Community College.

Peña, J. (2004, Fall). The power of collaborative writing. *Learning Assistance Association of New England Newsletter*, 1-3.

Peña, J. (2002). Plugging holes: Lucy and Sue as desiring subjects. *SOLAS—A Boston College Magazine*. Boston: Boston College.

Peña, J., Britt, Y., Collado, G., Lora, A., Dinehart, L., & Chae, C. (2007). *Teachers and Early Reading First: Commentary on the curriculum and professional development*. Miami, FL: Early Reading First, Early Learning Coalition of Miami-Dade/Monroe.

- Peña, J., Britt, Y., Collado, G., Lora, A., Dinehart, L., & Chae, C. (2007). *Teachers' perceptions of factors that affect teaching and learning*. Miami, FL: Early Reading First, Early Learning Coalition of Miami-Dade/Monroe.
- Peña, J., Britt, Y., Collado, G., Lora, A., Dinehart, L., & Chae, C. (2007). *Implementation of Early Reading First: Direct instructional effects*. Miami, FL: Early Reading First, Early Learning Coalition of Miami-Dade/Monroe.
- Peña, J., Chae, C., Dinehart, L., Wasserberg, M., Lora, A., & Britt, Y. (2008). *Child language and literacy assessment report: 2007-2008*. Miami, FL: Early Reading First, Early Learning Coalition of Miami-Dade/Monroe.

VITA

MARTHA FERNANDA PINEDA CASTILLEJA

EDUCATION

2007-present Florida International University Miami, FL
EdD., Curriculum and Instruction with specialization in International and Intercultural Education

2005-2007 Florida International University Miami, FL
M.S., International and Intercultural Education, GPA: 3.93

2001–2005 Florida International University Miami, FL
B.S., Early Childhood Education.

EXPERIENCE

Sept 2010 – Aug. 2012 Office of Global Learning Initiatives Miami, FL
Graduate Assistant

2005-2010 Comparative and International Education Society Miami, FL
Graduate Assistant

2008, Summer World Health Organization Manila, Philippines
Volunteer-Research Assistant at WPRO

2006, Summer Transnational and Comparative Studies Miami, FL
Graduate Assistant

2006, Summer La Paz Institute Acapulco, MEX
Summer Camp Coordinator & Reading/Writing tutor

2005, Summer La Paz Institute Acapulco, MEX
Summer Camp Instructor

2003, Spring La Paz Institute Acapulco, MEX
Pre-Kindergarten Teacher & Science and Music Appreciation coordinator

SELECTED PUBLICATIONS AND CONFERENCE PRESENTATIONS

- Book Chapter: Pineda, M. F. (2012). *A Mexican educator on teaching diverse populations*. In Wynne, J., Delpit, L., and R. Milles (Eds). *Confessions of a White Educator: Stories in Search of Justice and Diversity*. Dubuque, PA: Kendall Hunt.
- Book Chapter: Pineda, M. F. (2010). Standardized Tests in an Era of International Competition and Accountability. In Wiseman, A. (Ed). *The Impact of International Achievement Studies on National Education Policymaking*, Volume 13, International Perspectives on Education and Society Series. (pp. 331-353). UK: Emerald Publishing.
- Book Chapter: Pineda, F. and Landorf, H. (2010). Historical and Philosophical Foundations of Intercultural Education in Mexico. In C. Grant, *Intercultural and Multicultural Education: Enhancing Global Interconnectedness*. (pp. 138-153). New York, NY: Routledge.
- Article: Schratz, A., Pineda, M. F., Reforma, L., Fox, N., Le Anh, T. Cavalli-Sforza, T., Henderson, M., Mendoza, R., Utzinger, J., Ehrenberg, J., and A. Tee (2010). Neglected Diseases and Ethnic Minorities in the Western Pacific Region: Exploring the Links. In *Advances in Parasitology*, Vol. 72 (pp. 80-108), 2010.
- Conference Presentation: American Evaluation Association Annual Meeting, Anaheim, California, November 2011. Paper presented: “Content Validity using Mixed Methods Approach: Its application and development through the use of a Table of Specifications Methodology” (Coauthors: Dr. Isadore Newman, Dr. Janine Lim and Fernanda Pineda)
- Conference Presentation: World Council of Comparative Education Societies, XIV Congress, Istanbul, Turkey, June 2010. Paper presented: “Global Learning and the Human Capability Approach” Coauthors: Hilary Landorf and Fernanda Pineda
- Conference Presentation: IAIE (International Association for Intercultural Education) in Athens, Greece, 2009 – Paper presented: “Intercultural Universities in Mexico: A Curriculum Perspective”
- Conference Presentation/Proceedings Publication: Pineda, M. F. (2007). “Unity and Diversity: Exploring High School Students Perceptions about Multiculturalism and the Intercultural University of Guerrero, Mexico” (different section than CIES 2007’s presentation), México. In S. M. Nielsen & M. S. Plakhotnik (Eds.), *Proceedings of the Sixth Annual Conference of the College of Education: Urban and International Education Section* (p. 90-95). Miami: Florida International University.
URL://coeweb.fiu.edu/research_conference/