Intercultural Competence: Development and Validation of a Theoretical Framework, a Cross-Cultural Multimethod Test, and a Collaborative Assessment Intervention

Dissertation

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

vorgelegt von

M.A. Deborah Schnabel

aus Frankfurt am Main

Tübingen

2015

Tag der mündlichen Prüfung: 22.05.2015

Dekan: Professor Dr. rer. soc. Josef Schmid

1. Gutachter: Professor Dr. phil. nat. Augustin Kelava

2. Gutachter: Professor Dr. Fons Van de Vijver

“To spend too much time in *Studies* is sloth; to use them too much for Ornament is affectation; to make Judgment wholly by their Rules, is the humor of a Scholar. They perfect Nature, and are perfected by experience: for Natural Abilities are like Natural Plants, that need Proyning by *Study*; and *Studies* themselves do give forth Directions too much at large, except they be bounded in by experience. Crafty Men contemn *Studies*, Simple Men admire them, and Wise Men use them: For they teach not their own use; but that is a Wisdom without them, and above them, won by Observation. Read not to Contradict and Confute, nor to believe and take for granted, nor to find Talk and Discourse, but to weigh and consider.“ (Bacon, 1696, p. 135)

Summary

Globalization, migration, and international traffic have substantially raised the relevance of intercultural competence (ICC) in today's world of work (Fantini, 2009; Leung, Ang, & Tan, 2014; Spitzberg & Changnon, 2009). This has also increased the demand for assessing and training individuals to enable them to handle their intercultural encounters while studying or working in an international context. Despite the existence of many definitions, models, and approaches to the measurement of ICC, the circular process of understanding, assessing, and training ICC has been understudied (Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). The interconnection between assessment and training calls for behavior-oriented characteristics that can actually be trained, whereas most of the ICC concepts are trait- and/or attitude-based (e.g., Multicultural Personality Model: Van der Zee & Van Oudenhoven, 2000; Developmental Model of Intercultural Sensitivity: Bennett, 1993). This monograph presents a novel ICC framework that was created by using a phenomenological expert-based strategy. The framework includes an ICC onion model, which organizes the various aspects of ICC. Malleable ability-based intercultural competences are located in the heart of this model, whereas traits, attitudes, or (cultural) intelligence are situated in an outer layer of the onion model. As competence is directly linked to performance (Erpenbeck, 2010), ICC is subsequently defined as a global behavioral orientation, which enables individuals to handle intercultural situations (Schnabel et al., 2015). Moreover, ICC is understood as a multidimensional construct in which multiple competences, together or separately, influence the quality of intercultural interaction. These competences were collected through exploratory expert interviews and literature research and empirically investigated in a pretest. The results are outlined in Chapter 3.

Chapter 4 shows how the Test to Measure Intercultural Competence (TMIC; Schnabel et al., 2015) was developed on the basis of the preliminary model presented in Chapter 3. Even though there is a strong need for researchers to use more than one method to assess a complex construct such as ICC, there was no validated multimethod instrument that was available in the ICC field (Leung et al., 2014). To fill this gap, Schnabel et al. (2015) created a situational judgment test (TMIC-SJT) that complements a Likert-scale-based self-appraisal scale (TMIC-SA). Chapter 4 illustrates the design

and results of two studies, one conducted with German students and the other one with German professionals. I show the satisfactory fit statistics of a first-order factor model with 17 factors, calculated with exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009), and of a second-order factor model with six factors, calculated with confirmatory factor analysis. Evidence for construct and criterion validity as well as for the incremental validity of the TMIC-SJT is given.

Chapter 5 reviews a newly designed brief training intervention that was based on collaborative assessment (Fischer, 1994, 2000). The benefits of taking part in the TMIC and receiving written feedback as well as a 1-hr collaborative test-feedback intervention for students going abroad are illustrated with the results of an experimental study (Schnabel, Kelava, & Van de Vijver, in press). There is a focus on self-reported ICC development, the stages of change (DiClemente & Prochaska, 1998; Prochaska, DiClemente, & Norcross, 1992), intercultural self-confidence, and intercultural self-understanding.

Chapter 6 shows how a short version of the TMIC (TMIC-S; Schnabel, Kelava, Van de Vijver, & Seifert, 2014), particularly useful for employee selection purposes, was developed and tested in Germany and Brazil. Through the analysis of measurement invariance (Meredith, 1993), a cross-culturally valid model of ICC was supported.

The overall discussion in Chapter 7 integrates the aforementioned subtopics of this monograph and describes the major implications and limitations. Additionally, I give a global outlook concerning future research topics in the ICC domain.

Keywords: intercultural competence, assessment, training, situational judgment test, multimethod, collaborative assessment, feedback, measurement invariance, cross-cultural validity, employee development, employee selection

Zusammenfassung (German Summary)

Aktuelle Entwicklungen hinsichtlich der Globalisierung, der Migration und des internationalen Handelsverkehrs wirken sich unmittelbar auf die Bedeutung interkultureller Kompetenz in der heutigen Arbeitswelt aus (Fantini, 2009; Leung, Ang, & Tan, 2014; Spitzberg & Changnon, 2009). Der Bedarf an Mess- und Trainingsinstrumenten, die ein Individuum dabei unterstützen, interkulturelle Herausforderungen während des Studiums oder der Arbeit im internationalen Kontext zu meistern, wächst entsprechend. Zwar existieren bereits Definitionen, Modelle und Messansätze zur interkulturellen Kompetenz, jedoch wurde der zirkuläre Prozess aus Definition, Messung und Training von interkultureller Kompetenz in der Wissenschaft bislang vernachlässigt (Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). Die Vernetzung zwischen Messung und Training verlangt nach verhaltensorientierten Eigenschaften, die auch tatsächlich trainiert werden können. Die meisten bestehenden Konzepte basieren allerdings auf stabilen Persönlichkeitsmerkmalen und/oder Einstellungen (z. B. Multicultural Personality Model: Van der Zee & Van Oudenhoven, 2000; Developmental Model of Intercultural Sensitivity: Bennett, 1993).

In dieser Monographie wird ein Rahmenkonzept interkultureller Kompetenz präsentiert, das mithilfe einer phänomenologisch-expertenbasierten Strategie entwickelt wurde. Das Rahmenkonzept enthält ein Zwiebelmodell interkultureller Kompetenz, das die unterschiedlichen Aspekte interkultureller Kompetenz organisiert. Veränderbare fähigkeitsbasierte interkulturelle Kompetenzen sind im Herzen des Modells angesiedelt, wohingegen Persönlichkeitsmerkmale, Einstellungen oder auch (kulturelle) Intelligenz in der äußeren Schicht lokalisiert sind. Da Kompetenz direkt an das Verhalten gekoppelt ist (Erpenbeck, 2010), wird interkulturelle Kompetenz folglich als globale Verhaltensorientierung definiert, die ein Individuum dazu befähigt, interkulturelle Situationen zu meistern (Schnabel et al., 2015). Zudem wird interkulturelle Kompetenz als multidimensionales Konstrukt verstanden, in dem multiple Kompetenzen, zusammen oder einzeln, die Qualität interkultureller Interaktion beeinflussen. Diese Kompetenzen wurden durch explorative Experteninterviews sowie durch eine ausführliche Literaturrecherche gesammelt und empirisch in einer Vorstudie untersucht. Die Ergebnisse sind in Kapitel 3 aufgeführt.

Kapitel 4 beschreibt, wie der Test zur Messung Interkultureller Kompetenz (TMIK; Schnabel et al., 2015) auf Basis des Modells aus Kapitel 3 entwickelt wurde. Obwohl, in Anbetracht der Komplexität des interkulturellen Kompetenzkonstrukts, eine große Notwendigkeit besteht, mehr als eine Methode zur Messung heranzuziehen, war bislang kein validiertes multimethodales Instrument dafür verfügbar (Leung et al., 2014). Um diese Lücke zu schließen, entwickelten Schnabel et al. (2015) einen Situationsbeurteilungstest (TMIK-SJT), der eine auf Selbsteinschätzung basierende Likert-Skala (TMIK-SA) ergänzt. Kapitel 4 gibt einen Überblick über das Design und die Ergebnisse zweier Studien, von denen eine mit deutschen Studenten und die andere mit deutschen Berufstätigen durchgeführt wurde. Ich zeige, dass sowohl die durch Exploratory Structural Equation Modeling (ESEM; Asparouhov & Muthén, 2009) ermittelten Fit-Indizes für ein 17-faktorielles Modell erster Ordnung als auch jene, die durch Konfirmatorische Faktorenanalyse für das sechs-faktorielle Modell zweiter Ordnung berechnet wurden, insgesamt zufriedenstellend waren. Ergebnisse zur Konstrukt- und Kriteriumsvalidität sowie zur inkrementellen Validität des TMIK-SJT werden zudem präsentiert.

Kapitel 5 stellt eine eigens entwickelte Kurzintervention vor, die dem Ansatz des Kollaborativen Assessments folgt (Fischer, 1994, 2000). Die Ergebnisse aus einer experimentellen Studie (Schnabel, Kelava, & Van de Vijver, in press) zeigen den Nutzen, der sich für ein Individuum aus der Teilnahme am TMIK mit einer anschließenden einstündigen kollaborativen Testfeedback-Intervention ergibt. Dieser Nutzen bezieht sich auf die durch Selbsteinschätzung ermittelte interkulturelle Kompetenzentwicklung, die individuelle Stufe im Transtheoretischen Modell (DiClemente & Prochaska, 1998; Prochaska, DiClemente, & Norcross, 1992), das interkulturelle Selbstvertrauen und das interkulturelle Selbstverständnis.

In Kapitel 6 fasse ich zusammen, wie die, vor allem für die Personalauswahl entwickelte, Kurzversion des TMIK (TMIK-S; Schnabel, Kelava, Van de Vijver, & Seifert, 2014) in Deutschland und Brasilien validiert wurde. Die Ergebnisse der Analyse der Messmodellinvarianz (Meredith, 1993) stützen das Vorliegen eines kulturübergreifend validen Modells interkultureller Kompetenz.

In Kapitel 7 diskutiere ich die zuvor genannten Themen dieser Monographie umfassend und beschreibe die zentralen Implikationen und Limitationen. Zusätzlich gebe ich einen Ausblick auf zukünftige Forschungsfragen zum Thema interkulturelle Kompetenz.

Schlüsselwörter: Interkulturelle Kompetenz, Assessment, Training, Situational Judgement Test, multimethodal, Kollaboratives Assessment, Feedback, Messmodellinvarianz, kulturübergreifende Validität, Personalentwicklung, Personalauswahl

Contents

Chapter 1: Introduction

Without a doubt, globalization is one of the 21st century's magic words. Globalization means an increase in the quality and quantity of transnational and transcultural interaction with an immediate and profound effect on society, economics, politics, and culture (Al-Rodhan & Stoudmann, 2006). Globalization also dramatically motivates global migration. This requires a rethinking of traditional and mainly one-dimensional diversity concepts that differentiated individuals only on the basis of their country of origin or their ethnicity (Fantini, 2009; Vertovec, 2007). Such a diversified diversity concept is Vertovec's (2007) *super-diversity*, which he derived from his observations of the transformation in British society. Super-diversity acknowledges that multiple demographic and social variables—including country of origin and ethnicity on the one hand and languages, subcultures, modes of migration, and legal status on the other hand—dynamically interact and together form an extremely manifold reality. In this reality, an individual might have multiple origins, cross-cultural connections, and social-economic backgrounds. Modern employers—whether located in Britain or in another country—have to face this reality and adapt their corporate strategies as well as their human-resource practices accordingly.

Over the past 10 years, an increase of 25% was found for the number of employees who were sent abroad (so-called *expatriates*). If globalization proceeds and the lack of qualified labor in certain parts of the world remains, this trend will continue (ECA International, 2012). The wish to work and live abroad for a defined, usually limited period of time may well become an important characteristic of evolving generations. A great number of individuals who were born in the 1980s and thus belong to the age cohort called *Generation Y* (Sheahan, 2005) claim that working in a foreign country is one of their main professional goals (The Economist Intelligence Unit Limited, 2010).

However, expatriates are not the only group of individuals who have to deal directly with multicultural matters. Nowadays, the global mobility of students is a common practice in most parts of the world. In the academic year 2012/2013, 35,000 German students took part in the so-called Erasmus Mundus Programme (European Union, 2014) to study in another European country. This was an all-time record (Deutscher Akademischer Austauschdienst, 2014). Moreover, the importance

of traveling has surged. In 2012, Germans spent 63.9 Million Euro on traveling, an amount that is higher than ever before (Deutscher Reiseverband, 2013).

Dealing with super-diversity is also required in one's home country, particularly because of significant increases in migration rates. For example, in 2011, the ratio of individuals with an migration background was 20% in Germany (Statistisches Bundesamt, 2013); the skills shortage in Germany will attract even more foreign labor in the future (Bundesinstitut für Berufsbildung, 2013).

Globalization, super-diversity, and internationalization are all developments on a macro level. They call for a closer look at the microlevel to answer the question of what it takes for an individual to benefit from living and working under such circumstances. Clearly, interacting with individuals of various cultural backgrounds domestically and abroad is not an exception anymore (Fantini, 2009; Lustig, 2005). Thus, it is insufficient to acquire knowledge about one specific culture. Rather, individuals have to develop competences that allow them to quickly (a) understand themselves and another person in the context of culture (cf. Chen & Starosta, 1998; Thomas, Kammhuber, & Layes, 1997; Triandis, 1977), (b) switch between different behavioral actions and communication styles (cf. Chen & Starosta, 1998; Fantini & Tirmizi, 2006; Kelley & Meyers, 1995), and (c) shift between different cultural frames of reference (i.e., integration; cf. Hammer, Bennett, & Wiseman, 2003; Moosmüller, 2007). Intercultural competence (ICC), which incorporates all of these aspects, has become a major 21st century skill (Bremer, 2006; Deardorff & Hunter, 2006; Hulstrand, 2008; Institute for the Future for University of Phoenix Research Institute, 2011).

This monograph addresses the nature, assessment, and training of ICC. Hereby, it focuses on students and employees in Germany and Brazil. It is structured as follows: In Chapter 2, existing definitions, models, and approaches to the assessment of ICC are reviewed and the role of feedback in the assessment process is outlined. Chapter 2 provides an overview of the four main research questions, which are answered in the three chapters that follow. Accordingly, Chapter 3 focuses on the ICC framework, which was derived from the literature research and from exploratory expert interviews. Chapter 4 describes the development and validation of the Test to Measure Intercultural Competence (TMIC; Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). Chapter 5 introduces the newly

developed collaborative test-feedback intervention, which is aimed at fostering ICC (Schnabel, Kelava, & Van de Vijver, in press). Chapter 6 presents a short version and the cross-cultural validation of the TMIC (TMIC-S; Schnabel, Kelava, Van de Vijver, & Seifert, 2014), which was designed for employee selection purposes. Chapter 7 contains an overarching discussion of this monograph including major results, study limitations, and implications for research and practice.

Chapter 2: Theoretical Background

Despite the rather long research tradition concerning the nature of ICC, the actual importance and relevance of this topic is greater than ever (Fantini, 2009; Leung, Ang, & Tan, 2014; Spitzberg & Changnon, 2009). Altogether, this might explain the vast number of ICC definitions and models rooted in various disciplines such as management (cf. Bücker & Poutsma, 2010; Lloyd & Härtel, 2010), psychology (cf. Chiu & Hong, 2005; LaFromboise, Coleman, & Gerton, 1993), education (e.g., Anderson, Lawton, Rexeisen, & Hubbard, 2006; Deardorff & Hunter, 2006), and so forth. Leung et al. (2014) observed the existence of 30 ICC models and more than 300 related facets. Not surprisingly, this abundance has also led to a lack of conceptual clarity (Ang et al., 2007; Deardorff, 2004). In addition, various terms have been used simultaneously to describe ICC; for example, intercultural competence, cross-cultural competence, intercultural communication competence, intercultural sensitivity, global competence, cross-cultural adaptation, international competence, transcultural competence, and so forth (Fantini, 2009). Some of these terms mean the same de facto. Others mean something different, despite being used as if they were synonyms (e.g., intercultural competence vs. intercultural sensitivity). In the present work, I use the term intercultural competence as it is the most widely accepted one (Fantini, 2009). From a linguistic point of view, most terms have a shared inherent meaning of ICC. They describe a concept that addresses certain aspects of an individual and that applies when two or more cultures collide (Sinicrope, Norris, & Watanabe, 2007). The meaning of culture is up for discussion. In line with a super-diverse worldview (Vertovec, 2007), recent approaches have supported definitions of culture that go beyond the accumulation of similarities (cf. Thomas, 2003) or differences (cf. Auernheimer, 1996) in national or regional matters. Rather, culture has been transferred to any kind of interconnectedness of a group of people who share some specific characteristics (so-called human collectives; Hansen, 2000). This enlarges the horizon of ICC as it consequently relates to more than just international aspects (e.g., corporate culture).

2.1 Conceptualizing Intercultural Competence

In a broader sense, ICC is often understood as an ability or a set of abilities that enable an individual to function effectively and appropriately across cultures (Ang et al., 2007; Fantini &

Tirmizi, 2006; Hammer et al., 2003; Whaley & Davis, 2007). Johnson, Lenartowicz, and Apud (2006) defined ICC more specifically as "an individual's effectiveness in drawing upon a set of knowledge, skills, and personal attributes in order to work successfully with people from different national cultural backgrounds at home or abroad" (p. 530). As an answer to the ongoing debate about the nature of ICC in Germany, Rathje (2007) proposed a definition that includes information about the goal, scope, application, and foundation of ICC: "Given that culture is understood as existing within human groups, characterized by cohesion that is due to familiarity with inherent differences between them the intercultural competence can be defined as a culturegeneric skill, which is required in interactions between individuals from different human groups who are experiencing foreignness as a consequence of their mutual ignorance of the spectra of differences between them with a view to producing culture by creating familiarity and thus cohesion amongst the individuals involved, allowing them to pursue their interactional goals" (p. 264). Rathje's (2007) definition serves as a fruitful basis for the ICC framework presented later in this monograph for the following reasons: (a) ICC is understood as a skill, (b) Hansen's (2000) concept of culture is used, (c) Schönhuth's (2005) and Thomas' (2003) perception of ICC as a key ability to attain goals is integrated, and (d) Wierlacher's (2003) culture-generic approach to ICC is applied.

2.1.1 Intercultural competence models. Although most ICC models are multidimensional (Deardorff, 2006), they differ greatly concerning the nature, relations, and consequences of those dimensions. In the following, I provide an overview of existing ICC models. To distinguish them, I use the three most recent and established classification approaches: (a) Bolten's (2007) threefold taxonomy, (b) Spitzberg and Changnon's (2009) fivefold taxonomy, and (c) a threefold taxonomy based on a trait-competence-attitude distinction adapted from Leung et al. (2014).

In order to structure the various definitions, Bolten (2007) developed a threefold taxonomy. Listing models, structural models, and procedural models are distinguished as a function of the degree of complexity with which facets are organized in the models. Listing models encompass the early research approaches to ICC. The aim of Brislin's (1981) or Ruben's (1976) work was to collect ICC-related characteristics. Finding a higher order of those characteristics was executed with the help of

structural models, which assigned ICC characteristics mostly to affective, cognitive, and behavioral categories (e.g., Dauner, 2011; Gertsen, 1990; Ting-Toomey, 1993). Bolten's (2007) relatively recent definition of ICC exemplifies the essence of procedural models; that is, the interconnection of ICC with other core (social, functional, and strategic) characteristics of an individual as well as with context variables. Bolten (2007) consequently defined ICC as a context-specific competence to act.

Spitzberg and Changnon (2009) distinguished five types of competence models: Compositional models are analogous to the aforementioned listing models (cf. Bolten, 2007). Co-orientational models (e.g., Byram, 1997; Fantini, 1995) focus on the communication component of ICC and herewith on an individual's abilities that enhance intercultural interaction and mutual understanding (e.g., empathy, clarity, and perspective-taking). Developmental models (e.g., Bennett, 1986, 1993; Gullahorn & Gullahorn, 1962) refer to an individual's progress in intercultural matters. Such models define different stages that are often placed along a two-poled continuum and that describe how an individual deals with the cultural differences that he or she is exposed to. In adaptational models (e.g., Berry, Kim, Power, Young, & Bujaki, 1989; Kim, 1988; Navas et al., 2005), adjustment to a different culture is seen as (a) a process that occurs while interacting with individuals from another culture under the influence of situational factors and (b) a criterion of ICC. On the basis of empirical results, causal path models (e.g., Arasaratnam, 2006; Griffith & Harvey, 2000) specifically define the interrelations between ICC facets and outcome variables. In contrast to adaptational models, which conceptualize ICC as a process, path models define ICC as a linear system (Spitzberg & Changnon, 2009).

Leung et al.'s (2014) threefold taxonomy is not as elaborated as the one by Spitzberg and Changnon (2009). However, it stands out due to its clear focus on the type of characteristics (e.g., traits, abilities, attitudes), a focus that is especially important in the course of this monograph.

Traits refer to stable characteristics that are independent of situational determinants (Costa & McCrae, 1992). ICC models (e.g., the Multicultural Personality Model; Van der Zee & Van Oudenhoven, 2000), which are often rooted in the field of personality research, consist of such stable traits. These traits are assumed to have a significant influence on intercultural interactions and to be

reliable predictors of performance. Examples include emotional resilience (Kelley & Meyers, 1995), empathy (Fantini & Tirmizi, 2006; Koester & Olebe, 1988; Ruben, 1976), flexibility (Fantini & Tirmizi, 2006; Kelley & Meyers, 1995; Van der Zee & Van Oudenhoven, 2000), openness to experience (Chen & Starosta, 1998; Kelley & Meyers, 1995; Van der Zee & Van Oudenhoven, 2000), and tolerance of ambiguity (Deardorff, 2006; Fantini & Tirmizi, 2006; Koester & Olebe, 1988; Ruben 1976).

Competences—also referred to as abilities, capabilities, or skills—are malleable characteristics of a person and can therefore be learned (Erpenbeck, 2012; Weinert, 2001). In contrast to traits, competences are directly tied to context-specific performance, which implies that they become manifest variables only through behavior (Erpenbeck, 2010). They enable a person to successfully master challenges in unknown or unexpected situations (Erpenbeck & von Rosenstiel, 2007; Weinert, 2001). There are only a few ICC models that deal with malleable characteristics. One example is Bolten's (2007) ICC model in which established competences to act are generalized to the intercultural context. The concept of cultural intelligence (CQ; Earley & Ang, 2003) is also positioned within the competences framework of ICC (cf. Leung et al., 2014). Clearly, intelligence cannot be equated with competence, especially because of the assumed genetic determination of intelligence and its trait component (Sternberg & Detterman, 1986). However, CQ is defined as a malleable ability to function successfully in different cultural environments (Ang & Van Dyne, 2008; Earley & Ang, 2003). Additional example competences are related to communication (Gudykunst, 1994; Lloyd & Härtel, 2010), social interaction, self-management (Bird, Mendenhall, Stevens, & Oddou, 2010), collaboration, and learning (Leung & Cheng, 2014).

Intercultural attitudes and worldviews fall under the umbrella of intercultural sensitivity and are most often captured in the aforementioned developmental models (see the classification taxonomy; Spitzberg & Changnon, 2009). Whereas intercultural competences or traits determine whether an individual can handle intercultural differences effectively, intercultural sensitivity describes whether and how intercultural differences are perceived in the first place. Related models such as the Developmental Model of Intercultural Sensitivity (Bennett, 1993) distinguish between an

ethnocentric (i.e., reality is evaluated only on the basis of one's own cultural mindset) and a polycentric (i.e., the world is perceived with regard to multiple existing cultural mindsets) worldview of a person (cf. Bennett, 1993; Chen & Starosta, 2000; Hammer, 2011).

Several models integrate traits, competences, and attitudes. Examples are the Global Leadership Competency Model (GCI; Bird et al., 2010), which combines traits, competences, and attitudes; the Intercultural Competence Assessment Model (INCA; Fantini & Tirmizi, 2006), which includes competences and traits; and the Intercultural Sensitivity Model (Bhawuk & Brislin, 1992), which focuses on attitudes and traits.

2.1.2 Discussion of existing approaches to intercultural competence. The number of ICC models covering stable traits or attitudes is extensive. By contrast, very few models focus on malleable abilities that are related to ICC. Nevertheless, competence is assumed to be a more reliable predictor of performance in a cross-cultural setting than traits or attitudes (Ang, Rockstuhl, & Tan, in press). As is the case with comparable constructs such as work competence (e.g., Spencer & Spencer 1993), leadership (e.g., Stogdill, 1948), or social competence (e.g., Sarason, 1981), trait approaches have been historical pioneers in our understanding of all of these phenomena. However, recent research has instead focused on the malleable, behavioral, and situational aspects of these domains (cf. Hoffman, Woehr, Maldagen-Youngjohn, & Lyons, 2011; Kanning, 2002; Sternberg, 2005). A similar development in ICC research would be fruitful, as it would extend our understanding of what would be required and possible with training.

Moreover, the majority of ICC models were developed in an Anglo-American context, and hence, whether or not they also apply to other cultures is an open question (Deardorff, 2006; Martin, 1993). However, there is a strong need for universal approaches to ICC as cross-cultural experiences are often not restricted to one other culture (Arasaratam, 2007; Rathje, 2007). The differences in perspectives on ICC, ICC facets, and the content and wording of ICC items moreover hinder the establishment of a generalizable and universally acceptable ICC model on the one hand (Bolten, 2007; Deardorff, 2006; Rathje, 2006). On the other hand, the diversity of approaches to ICC is an important achievement and would be valued accordingly if different ICC phenomena were no longer

squeezed into one concept. As the scientific community that is addressing ICC enlarges, it might be useful to distinguish between subdomains of Intercultural Studies in the future such as research concerning intercultural personality (i.e., intercultural potential), ICC, intercultural mindset, CQ, or acculturation. For example, Schnabel et al.'s (2015) ICC concept sets a clear focus on malleable competences and the corresponding behavior that facilitates intercultural interaction. Simultaneously, the onion model, which organizes various subdomains of ICC, is part of that concept. For the full definition and a detailed explanation of Schnabel et al.'s (2015) model, which was developed as part of this monograph, see Chapter 3.

2.2 Assessing Intercultural Competence: An Overview of the Instruments used in the Field

Analogous to the quantity and diversity of ICC models, a vast number of instruments can be found in the intercultural research area. According to Fantini (2009), 44 different instruments currently exist. The measured variables, the mode of assessment, as well as the psychometric properties of these instruments differ greatly (Fantini, 2009; Gabrenya, Griffith, Moukarzel, Pomerance, & Reid, 2012; Gabrenya, Moukarzel, Pomerance, Griffith, & Deaton, 2011).

ICC instruments can apply to (a) one specific culture such as placement tests or attitude tests, (b) the field of linguistics such as the bilingual and culture-language dominance test or aptitudes tests, (c) the education sector such as formative tests, achievement tests, or criterion-referenced and norm-referenced tests, or (d) further development and employee selection of adults such as readiness tests, diagnostic tests, or proficiency, communication, and competency-based tests (Fantini, 2009). In line with the scope of this work, I will consider only instruments from the last category.

In addition, ICC instruments can be categorized into direct and indirect measurement procedures (Bolten, 2007; Fantini, 2009; Sinicrope et al., 2007). In the following, I will describe the functionality as well as the major instruments used in the two procedures separately.

2.2.1 Direct measures of intercultural competence. In direct assessment procedures, the variable of interest is measured over the course of performance. Hence, the goal is to observe whether an individual behaves in an interculturally competent manner and whether he or she is able to apply an acquired skill (Bolten, 2007; Fantini, 2009; Sinicrope et al., 2007). Direct methods include, for

example, intercultural assessment centers (e.g., Bolten, 2001; Lievens, Harris, Van Keer, & Bisqueret, 2003; Müller-Neumann, 2005), portfolio assessments (e.g., Byram, 1997; Jacobson, Schleicher, & Maureen, 1999; Pruegger & Rogers, 1994), and interviews (e.g., Byram, 1997; Fantini & Tirmizi, 2006; Straffon, 2003).

A recent approach that has been used to assess ICC performance involves situational judgment tests (cf. Ascalon, Schleicher, & Born, 2008; Rockstuhl, Ang, Ng, Lievens, & Van Dyne, 2013a; Rockstuhl, Presbitero, Ng, & Ang, 2013b). A situational judgment test (SJT; McDaniel, Hartman, Whetzel, & Grubb, 2006) typically consists of work-related critical incident scenarios that are briefly described to the test taker, who in turn selects an answer from a set of predefined ordinally ranked alternatives. SJTs follow the logic of an assessment center as they aim to collect information about the quality of an individual's behavior. Moreover, the criterion-related validity of the two methods is similar (cf. SJT: $\rho = .34$; McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001. Assessment center: $\rho = .37$; Gaugler, Rosenthal, Thornton, & Benson, 1987). However, in comparison with an assessment center, SJTs are highly economical, objective, and robust against biases (e.g., Hooper, Cullen, & Sackett, 2006; Lievens, Peeters, & Schollaert, 2008; McDaniel et al., 2006; McDaniel & Nguyen, 2001; Weekly & Ployhart, 2006). Besides the SJT that will be presented later in this monograph, the two further SJT approaches available in the intercultural domain are tied to the CQ context. The Cross-Cultural Social Intelligence SJT (CCSI SJT; Ascalon et al., 2008) measures two dimensions, namely empathy and ethnocentrism, with 14 scenarios. The concept of *cross-cultural social intelligence* is restricted to a person's ability to successfully manage social interactions that involve more than one culture. In a study with $N = 74$ international students, Ascalon et al. (2008) found an acceptable overall reliability of $\alpha = .68$. Also, the CCSI SJT showed weak average correlations ($r = .20$) with personality factors, derived from the International Personality Item Pool (IPIP; Goldberg, 1999) and an ethnocentrism scale by Aiello and Areni (1998). On the basis of initial studies by Rockstuhl et al. (2013a, b), Rockstuhl, Ang, Ng, Lievens, and Van Dyne (in press) developed an SJT for measuring intercultural interpersonal situations. It consists of seven multimedia vignettes. Participants are asked for their immediate behavioral response. Interrater reliability was between $\alpha = .79$ and .92. The seven items loaded on two factors (situational judgment and response

judgment), $\chi^2(69, N = 132) = 88.73$, *ns*; $\chi^2/df = 1.29$; IFI = .98; RMSEA = .05. Both factors were significantly correlated with task performance, $r(132) = .40$ and .44, ps < .01, and interpersonal citizenship, $r(132) = .28$ and .38, ps < .01, supporting their criterion validity.

2.2.2 Indirect measures of intercultural competence. Indirect self-report measures are typically implemented by administering standardized tests in which an individual is asked to indicate his or her level of agreement with a certain statement about ICC using a predefined Likert scale (Likert, 1932). Hereby, ICC is assessed from the test taker's point of view (Leung et al., 2014). The result is an individual's self-efficacy (Bandura, 1994) with regard to ICC; that is, one's perceived ability, which is known to be a direct indicator of performance (Stajkovic & Luthans, 1998). Semistructured interviews are another, although qualitative, self-report method (Barker, Pistrang, & Elliott, 2005) that is used less often to assess ICC (Leung et al., 2014). Self-report ICC questionnaires are high in economy, objectivity, and external validity (Bolten, 2007; Sebald, 2008; Sinicrope et al., 2007). Nevertheless, they also have some disadvantages such as method invariance (Campbell & Fiske, 1959) and response biases (e.g., social desirability or acquiescence; Barker et al., 2005). The latter especially apply to the employee selection context in which individuals might feel pressured to make a good impression (Leung et al., 2014).

Despite such disadvantages, self-report measures continue to be the most important instrument in ICC assessment (Leung et al., 2014). The latest ICC research articles call for an extension of self-report procedures to increase their incremental validity (cf. Bolten, 2007; Deardorff, 2006; Fantini, 2009; Leung et al., 2014). Thus, it is not a matter of "either-or" anymore but of "and." For example, although Rockstuhl et al. (2013a) found that metacognitive CQ measured with the Cultural Intelligence Scale (CQS; Van Dyne, Ang, & Koh, 2008) predicted task performance on multicultural teams better than an SJT, Schnabel et al. (2015) reported that the combination of an SJT and a self-report scale explained the variance in success-relevant criteria of ICC better than a self-report scale alone. As Bledow and Frese (2009) suggested, this finding can be best explained by the different aspects measured by self-report scales and SJTs, that is, self-concept and behavioral

preferences, respectively. Unfortunately, there is still a lack of use of the combined instruments (Leung et al., 2014).

For an overview of the most established and profound self-report questionnaires that measure either intercultural traits, competences, or attitudes, I selected instruments according to suggestions made in recent reviews. Gabrenya et al. (2011) analyzed 34 ICC instruments according to their validity. From those 34, they identified the following seven instruments for which validation data were available and that had been already used in previous studies: the Cross-Cultural Adaptability Inventory (CCAI; Kelley & Meyers, 1995), CQS (Van Dyne et al., 2008), Intercultural Adjustment Potential Scale (ICAPS; Matsumoto et al., 2001), Intercultural Development Inventory (IDI; Hammer, 2008, 2011; Hammer et al., 2003), Intercultural Sensitivity Scale (ISS; Chen & Starosta, 2000), Multicultural Personality Questionnaire (MPQ; Van der Zee & Van Oudenhoven, 2000, 2001), and Scale of Ethnocultural Empathy (SEE; Wang et al., 2001). Gabrenya et al. (2011) categorized the degrees of face, construct, and criterion validity of those seven instruments as good, moderate, and poor. The cut-off criterion for the relation between the full scale of an instrument and a comparable construct or criterion was $r = .30$. If more supportive validity outcomes than unsupportive outcomes were reported, Gabrenya et al. (2011) rated the validity as good. If the ratio between supportive and unsupportive outcomes was balanced, validity was considered to be moderate. If more unsupportive results were reported, validity was evaluated as poor. The ISS (Chen & Starosta, 2000) and the ICAPS (Matsumoto et al., 2001) each had one or more poor ratings, whereas the MPQ (Van Oudenhoven & Van der Zee, 2000, 2001) and the IDI (Hammer, 2008, 2011; Hammer et al., 2003) were rated good on all criteria. For all other instruments, mixed support was found for their validity (Gabrenya et al., 2011).

Similarly, Matsumoto and Hwang (2013) recently identified 10 tests that fulfilled the following criteria: (a) the predictive validity of outcomes such as living, working, or successfully adjusting to life abroad was reported, (b) psychometric properties were assessed in more than one study, (c) the approach was culture-general instead of culture-specific, and (d) English language papers were published in peer-reviewed journals. Matsumoto and Hwang (2013) considered all of the

instruments identified by Gabrenya et al. (2011) except for the Scale of Ethnocultural Empathy (SEE; Wang et al., 2001). In addition, they analyzed the Cross-Cultural Sensitivity Scale (CCSS; Pruegger & Rogers, 1993), the Behavioral Assessment Scale for Intercultural Communication Effectiveness (BASIC; Koester & Olebe, 1988), the Intercultural Communication Competence scale (ICC; Arasaratnam, 2009), and the ICSI (Bhawuk & Brislin, 1992). Matsumoto and Hwang (2013) concluded that the CQS (Van Dyne et al., 2008), the MPQ, and the ICAPS (Matsumoto et al., 2001) were the most promising instruments in the field.

In the following, I will review three of the most established instruments (cf. Gabrenya et al., 2011; Matsumoto & Hwang, 2013) in detail, each representing one of the three possible subconstructs of ICC: (a) stable traits: MPQ (Van Oudenhoven & Van der Zee, 2000, 2001); (b) malleable competence: CQS (Van Dyne et al., 2008), and (c) malleable attitudes: IDI (Hammer, 2008, 2011; Hammer et al., 2003). Table 1 provides an overview of the dimensions assessed by 11 other prominent instruments (cf. Gabrenya et al., 2011; Leung et al., 2014; Matsumoto & Hwang, 2013; Sinicrope et al., 2007). As can be seen in Table 1, none of these instruments sets a focus on malleable competences.

The Intercultural Development Inventory (IDI; Hammer, 2008, 2011; Hammer et al., 2003) is based on Bennett's (1986, 1993) Developmental Model of Intercultural Sensitivity (DMIS), which consists of six stages (denial, defence renewal, minimization, acceptance, adaptation, and integration) that reflect the development of an individual from ethnocentrism to ethnorelativism. For the first version of the IDI, Hammer et al. (2003) developed 145 items that were intended to assess the aforementioned six stages using a 5-point Likert scale ranging from *disagree* to *agree*. Data from 226 students were analyzed. On the basis of the results of six exploratory factor analyses (one for each of the six scales), several items were rejected, which led to a 60-item version with six scales that nevertheless did not match the theoretical stages of the DMIS. Thus, Hammer et al. (2003) made some revisions and tested the 145 items in a second sample of 591 college students. This resulted in a five-factor model and 50 items with reliabilities ranging from $\alpha = .80$ to .85.

Table 1

Scales, Content Domains, and Methods used in Established Intercultural Competence Instruments

Scale	Operationalization of intercultural competence	Content domain	Type of method
BASIC	Display of respect, interaction posture, orientation to knowledge, empathy, individualistic roles, relational role orientation, task-related role orientation, interaction management, tolerance of ambiguity	Trait-competence-mix	External evaluation
CCAI	Emotional resilience, flexibility and openness, perceptual acuity, personal autonomy	Stable traits	Self-report
CCSS	Cultural knowledge, cultural attitudes, cultural beliefs, cultural lifestyles	Attitude-competence-mix	Self-report
GCI	Nonjudgmentalness, inquisitiveness, tolerance of ambiguity, cosmopolitanism, interest flexibility, relationship interest, interpersonal engagement, emotional sensitivity, self-awareness, social flexibility, optimism, self-confidence, self-identity, emotional resilience, non-stress tendency, stress management	Trait-competence-mix	Self-report
GMI	Global business savvy, cognitive complexity, cosmopolitan outlook, passion for diversity, quest for adventure, self-assurance, intercultural empathy, interpersonal impact, diplomacy	Trait-competence-attitude-mix	Self-report
ICAPS	Emotion regulation, critical thinking, openness, flexibility, interpersonal security, emotional commitment to traditional ways of thinking, tolerance of ambiguity, empathy	Stable traits	Self-report
ICC	Cognitive, affective, and behavioural communication competence	Trait-competence-mix	Self-report
ICSI	Individualism, collectivism, flexibility	Attitude-competence-mix	Self-report
INCA	Tolerance of ambiguity, flexibility, empathy	Trait-competence-mix	Biographic questionnaire, role play, scenario
ISS	Interaction attentiveness, impression rewarding, self-esteem, self-monitoring, perspective taking	Trait-competence-mix	Self-report
SEE	Empathic feeling and expression, empathic perspective-taking, acceptance of cultural differences, empathic awareness	Trait-competence-attitude-mix	Self-report

Note. BASIC = Behavioral Assessment Scale for Intercultural Communication (Koester & Olebe, 1988); CCAI = Cross-Cultural Adaptability Inventory (Kelley & Meyers, 1995); CCSS = Cross-Cultural Sensitivity Scale (Pruegger & Rogers, 1993); GCI = Global Competencies Inventory (Bird et al., 2010); GMI = Global Mindset Inventory (Javidan & Teagarden, 2011); ICAPS = Intercultural Adjustment Potential Scale (Matsumoto et al., 2001); ICC = Intercultural Communication Competence Scale (Arasaratnam, 2009); ICSI = Intercultural Sensitivity Index (Bhawuk & Brislin, 1992); INCA = Intercultural Competence Assessment (Fantini & Tirmizi, 2006); ISS = Intercultural Sensitivity Scale (Chen & Starosta, 2000); SEE = Scale of Ethnocultural Empathy (Wang et al., 2001).

Of the five dimensions (denial/defense, reversal, minimization, acceptance/adaptation, and encapsulated marginality), three were significantly correlated with the Cross-Cultural World-Mindedness Scale (Der-Karabetian, 1992): $r(537) = -.29$, $p = .01$, for denial/defense; $r(523) = .29$, $p = .01$, for acceptance/adaptation; $r(544) = .12$, $p = .01$, for cultural marginality; and with an adapted version of the Social Anxiety Scale (Stephen & Stephen, 1985): $r(543) = .16$, $p = .01$, for denial/defense; $r(527) = -.13$, $p = .01$, for acceptance/adaptation; $r(555) = .14$, $p = .01$, for cultural marginality; thus partially confirming the convergent validity (Hammer et al., 2003). Several other studies have examined the factor structure of the IDI without finding a consistent pattern. Hammer (2011) examined the psychometric properties of the IDI in a cross-cultural sample consisting of 4,763 participants from 11 countries and found support for a seven-factor solution, which Greenholtz (2005) had already proposed in an earlier study. Hammer had investigated the factor structure separately for each culture, however, without testing for measurement invariance. The validity of the IDI had also been examined in several other studies. Paige, Jacobs-Cassuto, Yershova, and DeJaeghere (2003) assessed 353 high school and college students and showed that intercultural sensitivity, measured with the IDI, was higher in students with prior intercultural experience, prior studies of language and culture, friends from other cultures, and socializing experience with peers from other cultures (for the ANOVA results, see Paige et al., 2003). Students from international schools were mostly located at the acceptance and adaptation stages. Also, intercultural sensitivity was positively related to the length of time high school students had spent attending an international school, $r(336) = .19$, $p < .001$ (Straffon, 2003). Altshuler, Sussman, and Kachur (2003) investigated whether the intercultural sensitivity of 24 pediatric resident trainees would change over the course of an intercultural training intervention and found no significant difference between the pretest and the posttest. Moreover, Hammer (2011) reported a strong negative correlation between intercultural sensitivity on teams and their failure to complete diversity-related tasks, $r(6) = -.96$, $p < .001$.

The CQS (Van Dyne et al., 2008) is based on the aforementioned concept of CQ. As it claims to measure malleable abilities in the intercultural context, it falls by definition under the competence category of ICC. The four CQ dimensions—metacognitive CQ, cognitive CQ, behavioral CQ, and motivational CQ—are assessed with 20 self-report items on a 7-point Likert scale. During the first

study, Van Dyne et al. (2008) reported acceptable fit statistics of $\chi^2(164, N = 576) = 822.26$; RMSEA = .08, $p < .05$; SRMR = .06; CFI = .92. A second study revealed a considerably better fit: $\chi^2(164, N = 447) = 381.26$; RMSEA = .01, $p < .05$; SRMR = .04; CFI = .96. Both studies were conducted with Singaporean undergraduate students. The four-factor model was replicated over several periods of time and in the U.S. The validation sample was made up of $N = 794$ American and Singaporean students and working people. In the following, the major validation results for the four CQS dimensions are summarized (for an extensive review, see Ang et al., 2007): (a) weak correlations with openness to experience (Costa & McCrae, 1992), $r(235)$ ranging from .23 to .28, ps < .01, (b) weak to moderate correlations with emotional intelligence (Schutte et al., 1998), $r(235)$ ranging from .18 to .41, ps < .01, (c) nonsignificant to significant moderate correlations with the ability to adapt abroad (Kelley & Meyers, 1995), $r(358)$ ranging from .07 to .48; $p > .05$ to $p < .01$, (d) nonsignificant to significant moderate correlations with task performance (Ang et al., 2007), $r(358)$ ranging from .08 to .46, $p > .05$ to < .01, and (5) nonsignificant to significant weak correlations with cultural decision making (Ang et al., 2007), $r(358)$ ranging from .18 to .27, $p > .05$ to < .01.

The Multicultural Personality Questionnaire (MPQ; Van der Zee & Van Oudenhoven, 2000, 2001) assesses stable personality traits that cover motivational, professional, and occupational aspects in cross-cultural and international contexts. The first study on 257 Dutch college students revealed a model with seven factors, which were measured with 91 items that were rated on a 5-point Likert scale ranging from *not at all applicable* to *totally applicable.* The seven factors explained in total 30.6% of the variance (Van der Zee & Van Oudenhoven, 2000). A second study on 210 college students in the Netherlands examined the factor structure and revealed a solution with five instead of seven factors and 78 instead of 91 items. This was then considered to be the final MPQ model, which covers the following dimensions: cultural empathy, open-mindedness, social initiative, emotional stability, and flexibility. The internal consistency of the five scales ranged from $\alpha = .80$ to .91. Moreover, with one exception, the MPQ self-ratings were moderately correlated, $r(210)$ ranging from .47 to .54, ps < .01, with the MPQ other-ratings. Also, students in international studies who expected to go abroad scored higher in cultural empathy, $F(1, 201) = 5.69$, $p < .05$, $\eta^2 = .03$, open-mindedness, $F(1, 201) = 40.33$, $p < .001$, $\eta^2 = .17$, emotional stability, $F(1, 201) = 7.26$, $p < .001$, $\eta^2 = .04$, social

initiative, $F(1, 201) = 25.81$, $p < .001$, $\eta^2 = .11$, and flexibility, $F(1, 201) = 28.11$, $p < .001$, $\eta^2 = .12$, than psychology students with no international study aspirations (Van der Zee & Van Oudenhoven, 2001). In 2002 an English version of the original Dutch MPQ was administered to students from an international business school for the first time; however, neither the factor structure nor the measurement invariance of the English version was assessed (Van Oudenhoven & Van der Zee, 2002). Nevertheless, other studies have supported the predictive validity of the MPQ. Van Oudenhoven, Mol, and Van der Zee (2003), for example, showed that, when biographical variables were controlled for, cultural empathy and emotional stability predicted the life satisfaction of 102 expatriates in Taiwan, $\beta = .25$, $p < .05$ and $\beta = .19$, $p < .05$, respectively. Further, emotional stability predicted their physical well-being, $\beta = .39$, $p < .001$. Moreover, social initiative and emotional stability predicted their psychological well-being, $\beta = .32$, $p < .01$ and $\beta = .56$, $p < .001$, respectively. In addition, in a sample of 264 job applicants in Belgium and the Netherlands, the Big Five factor model (Goldberg, 1990) and the MPQ were related to each other. Specifically, extraversion was weakly, $r = .14$, $p < .05$, to moderately, $r = .36$, $p < .01$, correlated with the MPQ scale; intellectual autonomy was also weakly, $r = .17$, $p < .05$, to moderately, $r = .49$, $p < .01$, related to the MPQ (Van der Zee, Zaal, & Piekstra, 2003). Leone, Van der Zee, Van Oudenhoven, Perugini, and Ercolani (2005) showed that the MPQ even predicted international orientation over the Big Five dimensions, need for cognition, and learning orientation, $\Delta R^2 = .09$; $F(5, 141) = 4.26$, $p < .01$. Cross-cultural generalizability of the scale was assessed for a group of 421 Italians and 419 Dutchmen and Dutchwomen using multigroup confirmatory factor analysis with aggregated items. Thereby, configural invariance, $\chi^2(160, N = 840) = 528.55$, $p < .001$; NNFI = .93; CFI = .91; RMSEA = .074, and partial metric invariance, $\Delta\chi^2(8) = 1.22$, $p > .15$, were attained (Leone et al., 2005).

2.2.3 Limitations of existing instruments. Overall, my conclusion is that only a small percentage (about 20 to 25%) of the large number of instruments in the field of cross-cultural studies is able to fulfill the relevant psychometric requirements such as reliability and validity (see section 2.2.2). The CQS (Van Dyne et al., 2008), the IDI (Hammer, 2008, 2011; Hammer et al., 2003), and the MPQ (Van der Zee & Van Oudenhoven, 2000, 2001) are the three major exceptions as they provide sufficient information about internal consistencies as well as criterion, convergent, predictive,

and incremental validity. Multiple samples consisting of individuals from various age groups, professions, and cultures were assessed with all three instruments. Although the other 11 instruments (see Table 1) that were selected from previous reviews do not provide such a sound basis, they still fulfill basic quality criteria such as reliability and a minimum of one validity indicator. The major limitations that apply to all existing instruments in the field can be summarized as follows:

1. Content domain/operationalization: The definition and operationalization of ICC differs greatly within and between instruments. Specifically, most instruments measure a combination of traits, attitudes, and competences (see Table 1). No reasons for choosing these combinations are provided by the authors. To some extent, the selection of scales even appeared to be random.
2. Assessment utility: To evaluate the utility of the existing instruments, one major question has to be answered: Which kind of goal does an ICC instrument have (cf. Deardorff, 2004; Fantini, 2009; Rathje, 2007)? Instruments that assess stable characteristics such as personality traits might add value to the selection process as they show the potential that a candidate already has. However, in employee development settings, such as training and coaching, malleable characteristics are central as they are the ones that need to be developed. Thus, if the goal is to picture training needs, there is no sense in assessing traits that cannot be trained. However, only one out of all of the instruments, the CQS (Van Dyne et al., 2008), aims to measure a construct that can be trained. In addition, most ICC assessment approaches primarily follow a summative assessment approach, which is used to gather information about a person (Crooks, 2001). To support individuals in mastering intercultural challenges, formative (Sadler, 1989) or collaborative assessment (Fischer, 1994, 2000) would be needed instead. This would increase the utility of an instrument as it could serve as a (learning) intervention per se.
3. Self-report/method-mix: Although several authors (e.g., Deardorff, 2006; Gelfand, Imai, & Fehr, 2008; Leung et al., 2014; Rathje, 2007; Sinicrope et al., 2007) call for instruments that combine different methods to master the complexity of ICC, most of the established instruments are based solely on self-report. The only instrument that combines

different methods is the INCA (Fantini & Tirmizi, 2009), but support for its validity is still missing.

4. Factor structure: The replicability of the factor structure is challenging for most instruments in the field (e.g., Hammer, 2011; Van der Zee & Van Oudenhoven, 2000). Matsumoto and Hwang (2013) argue that it is possible that a general factor underlies ICC scales such as the CQS (Van Dyne et al., 2008) or the ICAPS (Matsumoto et al., 2001). If this is so, it would call into question the multidimensional nature of ICC (Deardorff, 2004; Spitzberg & Changnon, 2009).
5. Cross-cultural generalizability: Overall, most researchers in the field agree that ICC is not tied to one specific culture (Deardorff, 2004; Rathje, 2007). If the construct is assumed to be universal, the instruments measuring that construct have to be applicable across cultures as well (Arasaratnam, 2007). This would require an analysis of measurement invariance (Meredith, 1993) of the different language and cultural versions of ICC instruments. Although most instruments such as the IDI (Hammer, 2008, 2011; Hammer et al., 2003) are available in various languages and have often been applied in multiple cultures, measurement invariance was almost never tested. For instruments for which measurement invariance has been investigated, the results point to partial invariance at best (cf. MPQ; Leone et al., 2005).
6. Anglo-American Perspective: Most instruments were developed by Anglo-American researchers and address the English-speaking world (Deardorff, 2009; Martin 1993). Thus, there is a need for instruments originating in cultures other than the Anglo-American culture to examine whether or not the country of origin influences the construct that was developed and measured.

2.3 Assessment and Feedback

Feedback is defined as information about an individual's performance, behavior, or characteristics that is provided by one or more agents (Hattie & Timperley, 2007; Kluger & DeNisi, 1996). The mechanisms of feedback are based on the cybernetics in which an active system impacts

its environment. The corresponding consequences of this action are sent back to the system with the goal of initiating an adaption of the system to make it more suitable for its environment and therefore open to learning and development (Birkhan & Ringelband, 2013).

In educational and psychological testing, assessment feedback is a critical component of an individual's learning and development process (Hattie & Timperley, 2007). In a broader sense, summative and information-gathering assessments can be distinguished from formative and collaborative assessments. Summative and information-gathering approaches to assessment have the purpose of enabling assessors to picture the current profile of competence, state, or personality. Formative and collaborative assessments go beyond this descriptive nature as they include feedback for combining assessments with learning or development, respectively (Crooks, 2001; Finn & Tonsager, 1997; Sadler, 1989).

Collaborative assessment (Fischer, 1994, 2000), also called therapeutic assessment (Finn, 1996, 2007), is a highly individualized process of psychological testing and corresponding feedback that functions as a brief intervention. It is primarily used in clinical settings. In an open, trusting, and empathic dialog, test results are discussed with the goal of creating understanding and increasing the motivation to change. Central to collaborative assessment is the unique situation of the assessee and his or her relation to the test results.

In the field of applied psychology, especially in occupational settings, feedback is usually given as a consequence of an analysis of performance, competence, or developmental potential in relation to current and future tasks and challenges (Gunkel, 2014). Assessment feedback is given for several reasons. First, feedback increases the value of assessment. As London and McFarland (2010) put it: "Feedback is a way to maximize value from assessment dollars" (p. 426). Second, the reactions of the assessee concerning his or her test results might serve as an additional validation of the instrument. Third, the feedback session is used to set clear development goals together with the assessee. Fourth, in the employee selection context, feedback sessions offer the opportunity for decisions to be explained (Birkhan & Ringelband, 2013; Gunkel, 2014).

As is the case with feedback in educational settings, the impact of psychological test feedback can be explained with the help of feedback intervention theory (Kluger & DeNisi, 1996), goal setting theory (Locke & Latham, 1984), and control theory (Carver & Scheier, 1982). That is, feedback effectiveness is related to (a) a central focus on the task, job, or assignment (b) the integration of goal setting, and (c) the motivation to reduce the discrepancy between the current and the desired state (Gunkel, 2014; Hattie & Timperley, 2007; Miller & Rollnick, 2002; Sadler, 1989).

Recently, Gunkle (2014) published the first empirically investigated model for explaining the impact of assessment feedback in an occupational context. Her model builds strongly on previous feedback models (cf. Gilliland, 1993; Ilgen, Fisher, & Tylor, 1979) that explain how the perception and acceptance of feedback influences the assessee's motivation and goal setting. Feedback acceptance is then attained when the recipient feels that his or her performance is well reflected by the feedback.

Birkhan and Ringelband (2013) developed a similar model; however, they did not investigate it further. An adapted version of Gunkel's (2014) and Birkhan and Ringelband's (2013) model is displayed in Figure 1. The model shows that whether or not an individual benefits from the assessment feedback depends on the individual's acceptance of the feedback. Acceptance, in turn, is related to several contextual variables as well as to the characteristics of the assessment instrument, the feedback conversation, the feedback giver, and the recipient. If acceptance is established, understanding, self-awareness, and the motivation to change will increase. This finally results in a change action. The studies that have empirically tested the effects of feedback have come from different disciplines. In the educational field, a reasonable number of meta-analyses (e.g., Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Moin, 1986) have supported the positive effects of feedback. For example, a meta-analysis by Kluger and DeNisi (1996) revealed a moderate effect, $d = 0.41$, of feedback on students' performance in various areas. Moreover, Poston and Hanson's (2010) meta-analysis, which included 17 therapeutic assessment studies in the clinical setting, showed that treatment group means were significantly higher than control group means in 66% of the comparisons, $d = 0.423$, 95% CI [0.321, 0.525].

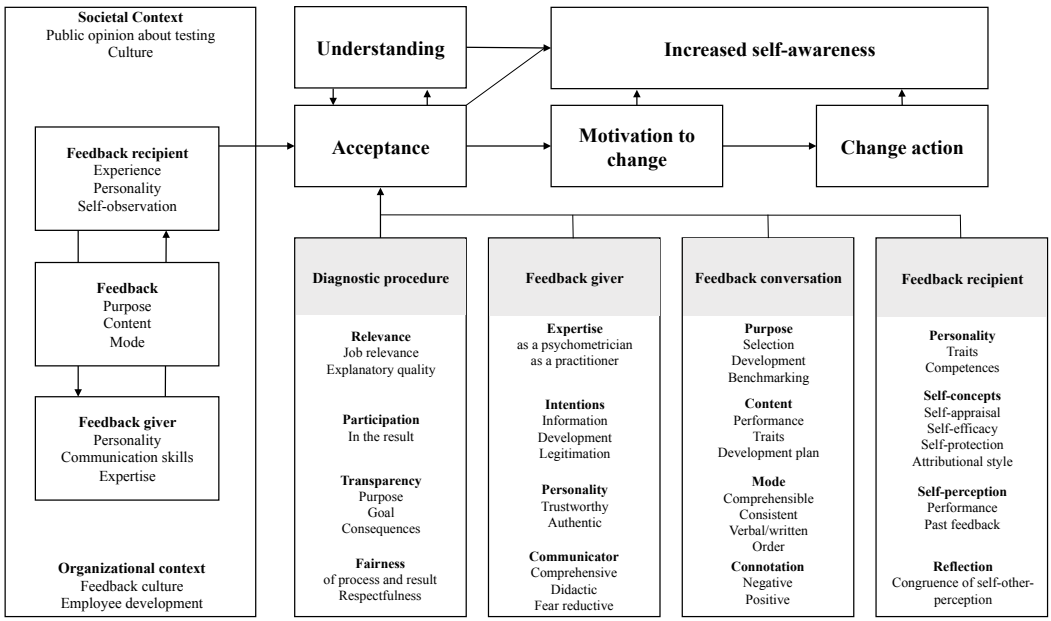

Figure 1. Schematic of the relevant variables and the process of feedback in the context of assessment. Hereby included are the characteristics of the instrument, message, feedback giver, and feedback receiver as well as the consequences of the feedback. Adapted from "*Akzeptanz und Wirkung von Feedback in Potenzialanalysen* [Acceptance and effects of feedback in potential analyses]," by L. Gunkel, 2014, p. 124. Copyright 2014 by Springer VS; and from "Rückmeldung der Eignungs-/Potenzialbeurteilung an den Kandidaten [Feedback for the candidate in aptitude-/potential analysis]," by G. Birkhan and O. J. Ringelband, 2013, *Management-Diagnostik* [Management diagnostics], p. 936. Copyright 2013 by Hogrefe.

In addition, the effects of therapeutic assessment on therapy outcomes such as self-esteem, self-awareness, self-understanding, motivation to seek mental health therapy, and satisfaction with the feedback were found to be moderate, $d = 0.367$, 95% CI [0.256, 0.478] (Allen, Montgomery, Tubman, Frazier, & Escovar, 2003; Finn & Tonsager, 1992; Newman & Greenway, 1997). In the occupational area, only a few studies have dealt with the consequences of feedback on performance and development. Furthermore, most studies refer to the effects of feedback concerning results in assessment and development centers instead of from standardized tests. For example, Byham (2005) showed that when participants accepted feedback after participating in a development center, their motivation to engage in follow-up activities was higher, $r(63) = .29$, $p < .05$. Gunkel (2014), moreover, found that well-accepted feedback concerning the results of a psychological test led to a

better self-awareness, $r(19) = .61$, $p < .01$, and an increase in motivation to initiate development activities, $r(19) = .66$, $p < .01$.

2.4 Overview of this Work and Major Research Goals

ICC is a highly complex construct that is challenging to assess (Arasaratnam, 2009; Deardorff, 2004; Fantini, 2009). There have been some advances in the assessment of ICC such as the INCA multimethod assessment (Fantini & Tirmizi, 2006) or SJTs in the field of CQ (e.g., Rockstuhl et al., in press); however, several fundamental challenges remain. First, the operational definition of ICC and the various related dimensions lack sufficient conceptual clarity (Ang et al., 2007; Deardorff, 2004). In addition, assessment instruments tend to randomly incorporate different content domains for measuring ICC. Second, self-reported Likert-type measures are usually used to describe individuals' stable traits (Leung et al., 2014; Sinicrope et al., 2007). Thereby, they fail to capture behavior-related malleable aspects and the corresponding training needs. Third, the cross-cultural generalizability of instruments is in question as measurement invariance testing has usually been insufficiently applied in the ICC domain (Rathje, 2007).

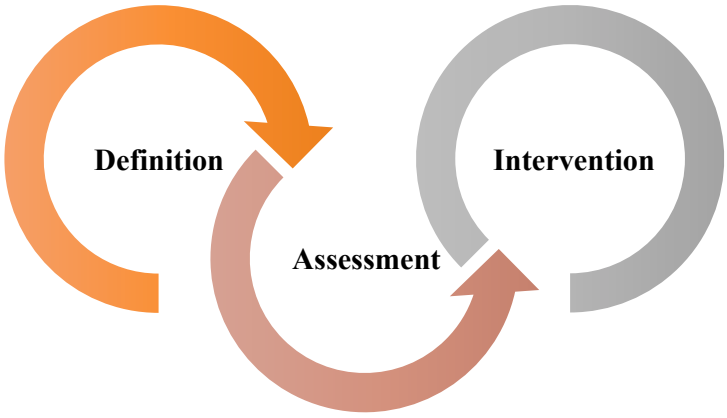

Figure 2. Schematic of the components of this monograph's overarching research topic. The process begins with the definition of intercultural competence, which in turn serves as the basis for the operationalization of the instrument that will be used to assess intercultural competence. This instrument is then used to create and apply an intervention that aims to train the competences that are measured by the instrument."

Consequently, this monograph makes a significant contribution to the creation of an integrative concept of ICC in which the theoretical definition fits the operationalization of ICC for the multimethod cross-cultural assessment of it, and this in turn directly connects to an intervention that is designed to foster ICC (see Figure 2). Specifically, four major research goals will be addressed in the following four chapters:

Chapter 3: Development of a framework and a corresponding set of competencies that were qualitatively and quantitatively evaluated to be (a) crucial in an intercultural context and (b) sensitive to training and coaching (Based on: **Schnabel** et al., 2015; Paper A).

Chapter 4: Development and validation of a behavior-oriented Test to Measure Intercultural Competence (TMIC), which assesses those competences with an innovative multimethod approach (Based on: **Schnabel** et al., 2015; Paper A).

Chapter 5: Development and evaluation of a brief intervention that is aimed at training ICC and related phenomena (Based on: **Schnabel**, Kelava, & Van de Vijver, in press; Paper B).

Chapter 6: Creation of a short version (TMIC-S) that can be used across cultures in addition to other instruments during the employee selection process. Simultaneously, the potential of a universal core set of intercultural competences is exemplarily investigated using German and Brazilian data (Based on: **Schnabel**, Kelava, Van de Vijver, & Seifert, 2014; Paper C).

Pre-print versions of Paper A, B, and C can be found in the Appendix.

Chapter 3: A New Framework of Intercultural Competence

This chapter addresses the genesis of the ICC definition and framework that serves as the basis for the TMIC. In addition, it summarizes the key findings of an exploratory expert interview study and of a quantitative pretest.

The very first step for exploring the ICC framework was to conduct exploratory expert interviews in order to thematically prestructure the object of research. Nonstandardized interview guidelines allow an open dialogue to unfold between the expert and the interviewer. Whatever the interviewer evaluates as important is noted. Neither an exact transcription of the material nor an establishment of comparability is required for these exploratory expert interviews (Bogner, Littig, & Menz, 2005; Lamnek, 1995). The procedure that I applied in this exploratory interview study was as follows: First, I created an interview guideline (see Appendix A), which consisted of an introduction and a closing section as well as interview questions that concerned the following five topics: (a) intercultural development, (b) definition of ICC, (c) ICC facets, (d) the potential for the ICC facets to be trained, and (e) the required characteristics of an ICC instrument. Second, I selected and invited the interview partners. The experts were active in the field of intercultural training, coaching, and mediation for various regions such as Germany, China, Sweden, the U.S., Middle-East, Russia, Eastern Europe, and South America. Third, I interviewed nine experts who had up to 9 years of intercultural training experience and recorded the interviews. Fourth, I converted the audio files into text by writing down everything that was said during the interviews. However, as I conducted the exploratory expert interviews, I did not follow transcription rules such as indicating pauses and highlighting statements. Fifth, I gathered statements from all interviews for the five topics mentioned before. They are displayed in Appendix A.

To develop a framework for the TMIC, I reviewed the ICC definitions proposed by the intercultural experts. The following key components were inferred:

- ICC consists of facets that enable a person to understand and adequately behave in different cultures;
- ICC means adapting one's behavior to the situation;

- ICC is based on the competence to act, which is flexibly used in various cultural contexts;
- ICC is acquired in the process of intercultural development.

I then combined the above elements with (a) Erpenbeck and von Rosenstiel's (2007) and Weinert's (2001) definitions of competence as a malleable construct, (b) Bolten's (2007) ICC approach, which is based on the competence to act (see section 2.1), and (c) Schuler and Prochaska's (2000) onion model, which they applied to achievement motivation. As is the case for achievement motivation, ICC is understood as a multidimensional phenomenon that is closely related to different constructs (Deardorff, 2006). The onion model organizes all relevant key facets, theoretically associated characteristics, and background characteristics in different layers (see Figure 3). The key facets include all of the competences that have an immediate influence on the intercultural experience and behavior of an individual. These include the first-order factors that are described in the following sections. Theoretically related constructs were identified via an analysis of a second-order factor model. A reference to the background characteristics was also useful for clarifying the understanding of ICC. Background characteristics are defined as associated constructs such as CQ, personality traits (e.g., openness to experience), or intercultural sensitivity. It can be assumed that background characteristics facilitate the acquisition of ICC: That is, if an individual realizes that intercultural differences exist and knowledge about and the motivation to explore another culture are present, there is a fruitful basis for ICC development.

Schnabel et al. (2015) assumed that ICC leads to a behavioral orientation that constantly takes cultural diversity into account. ICC is a context-specific competence to act (cf. Bolten, 2007). Accordingly, Schnabel et al. (2015) stated that ICC comprises a multitude of competences from the competence groups "social competence" (e.g., communication competence), "personal competence" (e.g., learning competence), and "methodological expertise" (e.g., problem-solving competence). These competences are supposed to be malleable; hence they can be learned by an individual and can directly influence behavior. They enable actors, individually or in combination, to flexibly master any kind of known, unknown, and/or challenging situation that is connected to the self, to others, or to a

specific task in the intercultural context. An intercultural context exists when more than one culture is (personally or virtually) involved. This is independent from the actual location of the actor (i.e., whether he or she is located in his or her home country or in a foreign country).

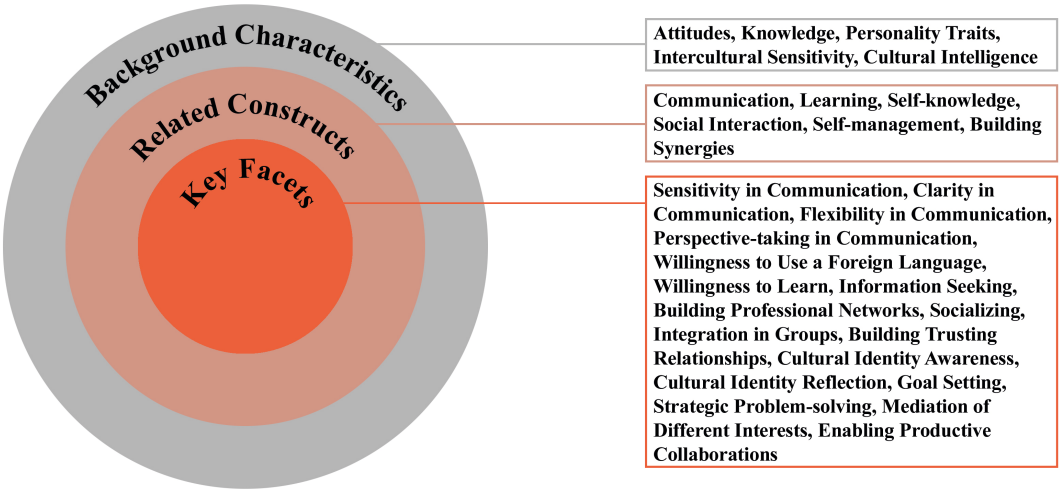

Figure 3. Schematic of the onion model of intercultural competence. The key facets refer to the 17 first-order factors of the Test to Measure Intercultural Competence (TMIC; Schnabel, Kelava, Seifert, & Kuhlbrodt, 2015). The related constructs are the TMIC's six second-order factors. The background characteristics represent the other important characteristics of an individual in the intercultural context.

Furthermore, the key deductions from the exploratory expert interviews were combined with the main findings from the literature search and used as the basis for the first exploratory pretest.

The goal was to define facets that were as selective as possible. This procedure led to 61 ICC facets. For each facet, a minimum of three items were formulated and reviewed by two independent psychologists. A complete list of all facets and items can be found in the supplemental material. The online questionnaire for the pretest consisted of the 238 self-appraisal items that measured the 61 competence facets with a 6-point Likert scale (*does not apply at all* to *fully applies*) and some further questions that were related to the sociodemographic characteristics and the intercultural experiences of the participants. A sample of $N = 150$ employees from an intercultural training company were assessed. The 106 women (71%) and 44 men (29%) were 33.23 years old ($SD = 10.61$) on average. The main goal was to evaluate the comprehensibility of the items and to conduct some exploratory analyses concerning the item difficulties and item discriminations as well as the scales' reliabilities,

intercorrelations, and relations to the intercultural experience factors. Results were interpreted cautiously as the ratio of items to study participants clearly resulted in weak power.

Finally, a total of 79 items belonging to 17 scales met the standard psychometric requirements (cf. Eignor, 2013) and were used in Study 1 (see section 4.2). Table 2 further explains the 17 first-order factors.

During the exploratory expert interviews, the experts also specified their requirements for a new ICC instrument, which they said should (a) focus on skills, (b) integrate items that are formulated as situations and behavioral alternatives, (c) be able to assess the ICC development across different points in time, (d) be independent from a specific cultural context, (e) be less prone to social desirability bias than existing instruments, and (f) fulfill psychometric standards. These aspects matched the previously reviewed requirements that were covered in the latest ICC research articles and thus provided a strong foundation on which to build a multimethod test for measuring malleable intercultural competences. This also led to the decision to integrate situational judgment items to complement the self-report items. The development of the situational judgment scale of the TMIC (TMIC-SJT) is described in the following section.

Table 2

Descriptions of the Factors with the Number of Items and Item Examples in English

Second-order factors	First-order factors	Description	Number of items (SA +SJT)	Example items SA
Communication	FC	Situation-specific adaptability of communication behavior	4 SA + 1 SJT	The way I address something depends on the person I am talking to.
	SC*	Ability to put yourself in the position of another person including sensitivity to nonverbal and paraverbal components of communication	7 SA + 1 SJT	I know how other people feel without them having to tell me.
	CC	Effective articulation of messages	4 SA + 1 SJT	I find it easy to express my thoughts in words.
	PC	Ability to understand another person's thinking and be able to observe circumstances from their point of view	5(3) SA + 1 SJT	I find it easy to view my behavior from other people's points of view.
Learning	WFL	Actual use of a foreign language that was previously learned—independent of how well it is spoken	4 SA + 1 SJT	I communicate in a foreign language even if I do not have a good command of the language.
	IS*	Purposeful collection of information about a foreign country or another culture that leads to valuable practical knowledge	3 SA + 1 SJT	When planning a trip abroad, I use various sources of information.
	WL	Openness of a person to gain new insights and willingness to invest time in learning new things	5 SA + 1 SJT	I spend a large part of my free time learning new things.
Social interaction	BTR	Ability to quickly build a trusting relationship with unknown people from different cultures	(3 SA) + (1 SJT)	When I join a group for the first time, I quickly build relationships with the other group members.
	IG	Ease with which a person can integrate themselves in an existing group and operate successfully in different systems	6(3) SA + 1 SJT	I find it easy to position myself within a group.
	BPN	Conscious creation of a network of people who can provide support to help a person satisfy his/her needs and achieve aims	3(4) SA + 1 SJT	I have a large network of professional contacts.
	S*	Ability and willingness to quickly establish contact with people from other cultures and maintain these contacts	5 SA + 1 SJT	I spend a large part of my free time cultivating contacts.
Self-management	SPS	Recognition and solution of problems in an international context	5(4) SA + 1 SJT	In order to solve a problem I weigh up differing alternative solutions.
	GS*	A person has clear aims and is able to implement them consistently	4(5) SA + 1 SJT	When I plan something, I usually then go on to achieve my aim.
Creating synergies	MI*	Mediating between interested parties in order to achieve the greatest possible benefit from different approaches	5 SA + 1 SJT	I am good at mediating between people who have conflicting interests.
	EPC	Uncovering and solving misunderstandings as well as being able to deal with different approaches during team work	5 SA + 1 SJT	When working in a team I try to highlight the mutual benefits to others.
Self-knowledge	CIR*	A person intensively and constantly considers their own cultural character	4(5) SA + 1 SJT	I make an effort to understand the extent to which my behavior is shaped by culture.
	CIA	Knowledge of one's own cultural values and norms	6 SA + 1 SJT	I am aware of the cultural values and norms that influence my behavior.

Note. SA = Self-appraisal; SJT = Situational Judgment Test; FC = Flexibility in Communication, SC = Sensitivity in Communication, CC = Clarity in Communication, PC = Perspective-taking in Communication, WFL = Willingness to Use a Foreign Language, IS = Information Seeking, WL = Willingness to Learn, BTR = Building Trusting Relationships, IG = Integration in Groups, BPN = Building Professional Networks, S = Socializing, SPS = Strategic Problem-solving, GS = Goal Setting, MI = Mediation of Different Interests, EPC = Enabling Productive Collaborations, CIR = Cultural Identity Reflection, CIA = Cultural Identity Awareness; the content in brackets refers to the final version of the TMIC; * indicates the factors on the short version (TMIC-S), although with a varying number of items.

Chapter 4: Development and Validation of the Test to Measure Intercultural Competence (TMIC)

Chapter 4 describes how the TMIC-SJT was developed and summarizes the method and results of two studies that were conducted to validate the TMIC. The content is based on Schnabel et al.'s (2015) research article, which contains additional details of both studies.

4.1 Development of the TMIC-SJT

First, I took the 17 self-appraisal scales from the TMIC (TMIC-SA) as inferred from the pretest and used them as a starting point. Then, I developed the situational judgment items to complement the TMIC-SA.

The construction of the TMIC-SJT was based on the approach recommended by McDaniel and Whetzel (2005). Twenty critical incidents (Flanagan, 1954) developed by expat managers, expatriates, and intercultural trainers served as the basis for the items. They included a typical international cooperation situation as well as statements about how an individual would ideally behave in this situation and which behaviors would possibly make the situation worse. The critical incidents were job- or country-specific in part and therefore had to be generalized. They were reformulated accordingly and converted into a suitable format. The situational judgment items that were designed were then assigned to the 17 factors in a double-blind procedure with six intercultural experts. Text-based situational judgments and behavior-related answer alternatives were chosen ("How are you most likely to behave in this situation?") as opposed to knowledge-related ("What is the best answer?") answer alternatives, as the former have a higher correlation with behavioral variables (McDaniel, Hartman, & Grubb, 2003; McDaniel et al., 2006), are more culturally independent (Nguyen, McDaniel, & Whetzel, 2005), and correspond more with the specific aim of showing the behavioral aspect of ICC. An example of a situational judgment item measuring flexibility in communication is given in the following; all situational judgment items are available as supplemental material:

You are to pass an important message on to a foreign colleague. However, the person's reaction shows that he or she has clearly not understood it. How are you most likely to behave?

a) I will leave it to someone else to pass on the message.

b) I will repeat the message using the same words but will speak more slowly and louder so that the person I am speaking to can understand me better.

c) I will repeat the message and pay close attention to my choice of words.

d) I will ask my colleague what he or she did not understand and reformulate the statement accordingly.

The TMIC-SJT instructions explicitly invite the test participants to choose one answer out of four. This procedure is commonly found on situational judgment tests (McDaniel & Whetzel, 2005).

4.2 Study 1: Factor Structure, Psychometric Properties, and Criterion Validity of the TMIC

The first study involved $N = 641$ German students. Seventy percent were female (488) and 30% (193) were male. The study participants were 28.25 years old ($SD = 9.29$) on average.

In order to investigate whether the 17 scales that were inferred from the pretest fit the data well, exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) in *Mplus* (Muthén & Muthén, 1998−2012) was used. The following changes were made on the basis of the ESEM results: Four items were excluded due to loadings < .20 (Bortz, 2006). Nine items were assigned to factors other than those that were originally anticipated. One factor (specific use of diversity) was omitted and another factor was split into two factors (integration in groups). The final structural equation model consisted of 75 self-appraisal items and was tested for its goodness of fit, which was found to be highly satisfactory, $\chi^2(1636, N = 641) = 2579.85$, $p < .001$; $\chi^2/df = 1.58$; RMSEA = .031, 90% CI [= .029, .033]; SRMR = .017; CFI = .96; TLI = .93. The numbers and examples of items on each factor are summarized in Table 2.

In a second step, a model containing the situational judgment items was tested. As a result of the aforementioned departures from the pretest model, situational judgment items were available for only 16 factors at this stage. Thus, the model fit was assessed for 17 first-order factors with 75 self-appraisal items and 16 situational judgment items. Furthermore, six second-order factors were included to explain the 17 first-order factors. As second-order models could not be computed with ESEM in Mplus (Asparouhov & Muthén, 2009), a confirmatory factor analysis (CFA) was applied.

This led to an acceptable model fit, $\chi^2(3987, N = 641) = 8280.09$, $p < .001$; $\chi^2/df = 2.08$; RMSEA = .040, 90% CI [.040, .046]; SRMR = .076; CFI = .82; TLI = .81. The internal consistency (Cronbach's α) of the 17 factors ranged from .69 to .90. The overall reliability of the TMIC-SA was excellent (α = .96).

Furthermore, criterion validity was established with the help of three external criteria: (a) duration of previous stays abroad, (b) previous participation in an intercultural training, and (c) private or professional intercultural involvement. Because a latent general factor was theoretically and statistically rejected, the variation in the overall TMIC value was analyzed on a manifest level. Two analyses of variance each revealed a significant effect of the influence of the length of a stay abroad on the overall value on the TMIC-SA, $F(3, 637) = 36.51$, $p < .001$, $\eta^2 = .15$, and on the TMIC-SJT, $F(3, 637) = 9.45$, $p < .001$, $\eta^2 = .10$ (for the related contrasts, see Schnabel et al., 2015). A *t* test showed that participants who had already participated in an intercultural training scored, on the whole, higher on the TMIC-SA, $M = 4.37$, $SD = 0.45$, as well as on the TMIC-SJT, $M = 3.07$, $SD = 0.28$, than people who had never participated in an intercultural training, $M = 4.14$, $SD = 0.48$; $t(639) = 7.87$, $p < .001$, $d = 0.69$, and, $M = 2.94$, $SD = 0.31$; $t(639) = 5.11$, $p < .001$, $d = 0.40$, respectively. Another *t* test revealed that professionally or privately interculturally involved individuals achieved a higher overall value on the TMIC-SA, $M = 4.31$, $SD = 0.47$, than participants who said they were not interculturally involved, $M = 3.90$, $SD = 0.44$; $t(639) = 8.69$, $p < .001$, $d = 0.70$. This also applied to the TMIC-SJT, $t(639) = 5.27$, $p < .001$, $d = 0.52$, such that individuals who had affirmed their intercultural involvement had greater TMIC-SJT scores, $M = 3.01$, $SD = 0.30$, than those who said they were not interculturally involved, $M = 2.85$, $SD = 0.33$. In addition, multiple group comparisons with a Bonferroni-Holm-corrected significance level of $p < .001$ (Holm, 1979) were conducted with sequential equation modeling (SEM) for the 17 latent factors and the external criteria (a) previous participation in intercultural training and (b) private or professional intercultural involvement. Previously interculturally trained and interculturally involved individuals scored higher on 13 of the 17 latent factors. All results are extensively summarized in Paper A (Schnabel et al., 2015).

To further validate the TMIC, correlations between the TMIC-SA and TMIC-SJT, as two different methods, were computed for all first-order factors. This led to one nonsignificant correlation, $r(641) = .05$, *ns* (flexibility in communication), one significant yet very low correlation, $r(641) = .08$, $p < .05$, and 15 other significant correlations that ranged from $r(641) = .11$, $p < .05$ (sensitivity in communication) to $r(641) = .61$, $p < .001$ (willingness to use a foreign language). The overall TMIC-SA was moderately correlated with the overall TMIC-SJT, $r(641) = .49$, $p < .001$. This corresponds to previous findings (cf. Bledow & Frese, 2009) that showed that although the two methods measure the same construct, the SJT measures behavioral preferences, whereas the self-report scale captures self-concept. In addition, the incremental validity of the TMIC-SJT was investigated via regression analyses. Specifically, I examined whether the TMIC would explain the external criteria better when the TMIC-SJT total score was added to the regression model in comparison with the TMIC-SA total score alone. The results showed that adding the TMIC-SJT explained more variance in (a) duration of previous stays abroad, $\Delta R^2 = .001$; $\chi^2(1) = 22.98$, $p < .001$, (b) previous participation in an intercultural training, $\Delta R^2 = .009$; $\chi^2(1) = 19.10$, $p < .001$, and (c) private or professional intercultural involvement, $\Delta R^2 = .001$; $\chi^2(1) = 20.38$, $p < .001$, than the TMIC-SA alone. However, even if the ΔR^2-values were significant, their size was very modest.

4.3 Study 2: Replication of the Factor Structure with Professionals and Construct Validation

A second study with $N = 313$ participants was conducted to replicate the factor structure in the German working population and to examine the construct validity of the TMIC. The average age of the 165 women (53%) and 148 men (47%) was 38 years ($SD = 16.26$). On average, the participants spent 120 weeks ($SD = 214.97$) abroad and had already lived in another country 2.22 times ($SD = 2.69$).

As in Study 1, two different analyses were applied to calculate the fit statistics due to Mplus software restrictions (Muthén & Muthén, 1998−2012). ESEM (Asparouhov & Muthén, 2009) was again used to test for the model fit at the first-order level. Results pointed to similarly satisfactory fit indices for the 17-factor model in the sample of working people, $\chi^2(1636, N = 313) = 2431.67$, $p <$

.001; χ^2/df = 1.49; RMSEA = .040, RMSEA 90% CI [.036, .043]; SRMR = .021; CFI = .94; TLI = .90.

The second-order CFA model, including the 17 situational judgment items, yielded mostly acceptable fit indices, $\chi^2(4062, N = 313) = 6672.01$, $p < .001$; χ^2/df = 1.64; RMSEA = .043, RMSEA 90% CI [.041, .044]; SRMR = .069; CFI = .80; TLI = .78.

Construct validity was assessed using three different comparison scales: (a) the CQS (Van Dyne et al., 2008), (b) the global competence scale from the *Intercultural Sensitivity Index* (Olson & Kroeger, 2001), and (c) the openness to experience scale from the NEO-Five-Factor Inventory (NEO-FFI; Borkenau & Ostendorf, 1993). A positive correlation was established between the TMIC-SA and the CQS, $r(297) = .67$, $p < .001$, the global competence scale, $r(184) = .63$, $p < .001$, and the openness to experience scale, $r(129) = .33$, $p < .001$. The TMIC-SJT was related to the CQS, $r(297) = .26$, $p < .001$, and the global competence scale, $r(184) = .17$, $p < .05$; however, its correlation with the openness to experience scale was nonsignificant, $r(129) = .16$, $p = .06$.

4.4 Implications concerning the TMIC

The results of an exploratory expert interview complemented the extensive literature research. Both sources called for a multimethod instrument that is based on a multidimensional malleable ICC construct. Such an instrument, namely the TMIC, was developed in two steps. First, a pretest was administered to preselect the self-appraisal items and scales. Second, a larger quantitative study tested the model fit, the psychometric properties, the relation between the methods, and the criterion validity of the TMIC. Construct validation was implemented in a study with professionals. After a final selection of items, ESEM (Asparouhov & Muthén, 2009) results pointed to a 17-factor solution. A model with six second-order factors, integrating the newly developed situational judgment scale of the TMIC and the self-appraisal scale, fitted the data to an acceptable degree and supported the assumed onion model of ICC. The moderate correlation between the TMIC-SA and the TMIC-SJT as well as the incremental validity of the TMIC-SJT showed that adding a situational judgment test to a self-report scale to measure ICC is not redundant but is rather complementary and thus enriching.

An interesting finding was that the correlation between the TMIC-SA and openness to experience, assumed as a rather stable personality trait, was relatively weak. The relation between openness and the TMIC-SJT was even nonsignificant. On the contrary, the TMIC results were significantly higher for individuals who (a) had higher values on the CQS, with CQ being a malleable construct too (Earley & Ang, 2003), and (b) had participated in an intercultural training in the past. This emphasizes the evidence for the malleable nature of the integrated competences. However, former participation in an intercultural training consists of retrospective information. Therefore, the study described in the next chapter focused on changes in the test results and other related phenomena as a direct consequence of participating in a brief training intervention.

Chapter 5: TMIC as a Training Intervention for Students Who Are Going Abroad

Chapter 5 reviews Schnabel et al.'s (in press) major findings concerning the development and validation of a collaborative test-feedback intervention that was based on the TMIC and the collaborative assessment approach (Fischer, 1994, 2000).

5.1 Study Background and Methodology

This study was conducted for three main reasons. The first goal was to gather more insights into the potential for the TMIC competences to be trained. The second goal was to validate the TMIC as a training intervention. The third goal was to test a newly developed economical intervention for the purpose of student ICC development. Specifically, I was interested in whether participating in a collaborative test-feedback intervention would positively influence an individual's TMIC-SA score, his or her motivation to change as measured with the transtheoretical model of stages of change (TTM; cf. DiClemente & Prochaska, 1998; Prochaska, DiClemente, & Norcross, 1992), and several other therapeutic benefit variables. These included an individual's intercultural self-understanding, intercultural self-confidence, and perceived benefit from test participation (cf. Allen et al., 2003; Finn & Tonsager, 1992; Newman & Greenway, 1997). In total, nine dependent variables were included. Seven were related to the students' development and two to their satisfaction with the feedback.

An experimental randomized pretest-posttest control group design was implemented and applied to $N = 820$ German university students who had been accepted for an ERASMUS semester abroad. The 480 females (58%), 327 males (40%), and 13 (2%) individuals who did not indicate their sex were 23.37 years old ($SD = 3.89$) on average. Between the two measurement points, no feedback or any other treatment was provided in the control group ($n = 351$), a written feedback report was sent to the comparison group ($n = 396$), and a collaborative test-feedback intervention was offered to the treatment group ($n = 73$). The collaborative test-feedback intervention consisted of a combination of a written feedback report, a results graphic, and an oral collaborative test-feedback talk.

This idea for a TMIC intervention was based on assumptions about the effects of assessment feedback on learning and behavioral change as well as on the collaborative assessment approach in clinical psychology, all of which were reviewed in section 2.3. Accordingly, a collaborative test-

feedback intervention was developed and named SHORT. SHORT is an acronym that stands for the key steps (*start and how we will proceed*, *orientation*, *reflection*, and *targets*) of a 1-hr telephone session conducted between a trained assessor and an assessee. Herewith, the existing guidelines for conducting test feedback and brief interventions were respected (cf. American Psychological Association, 1986, 2010; Finn, 1996; Fischer & Finn, 2008; Miller & Rollnick, 2002). The SHORT procedure was as follows:

1. Start and how we will proceed. The assessor had to make sure that the general framework was established for conducting the oral part of the collaborative test feedback. Confidentiality had to be guaranteed to the assessee who sat in a quiet place without the possibility of disturbances. Then, the assessor gave an overview of the session's content and process. The assessor also encouraged any kinds of questions or remarks during the session.
2. Orientation. First, the assessor clarified the roles of the assessor and assessee in the session. This also emphasized the collaborative nature of the feedback. The assessor presented herself as an expert on the TMIC, whereas the assessees were recognized as the experts on themselves. As such, they were motivated to help the assessor to correctly interpret the TMIC results. The assessor created an equal relationship by thanking the assessees for their willingness to share their thoughts and by communicating that the assessees would gain personal knowledge by participating in the session. Second, the assessor asked the assessees to share their previous experiences with psychological tests. Negative incidents were treated with a high degree of empathy. Before filling out the TMIC-SA, participants in the written-plus-oral collaborative test-feedback group had the chance to formulate questions about what they wanted to learn from the TMIC results. The assessor read these questions aloud in the oral feedback session and asked for additional ones. Previous experiences abroad and future international plans were also discussed. Finally, information concerning the purpose and content of the TMIC along with the graphical results were explained to the assessee.
3. Reflection. The assessees were asked to take some time to review their results. They were then prompted to explain whether and how they found themselves fairly well represented in the profile. They were also asked to identify results that were surprising or seemingly

unrelated to their self-image. The assessor took notes. The greatest therapeutic effects are expected when test feedback is ordered according to the client's existing self-concept (Finn, 1996; Schroeder, Hahn, Finn, & Swann, 1993). Therefore, the results were discussed in the following order: first, high values for competences with high acceptance by the assesse; second, low values for competences with high acceptance by the assessee; third, high and low values for competences with low acceptance by the assessee. The assessor had to ask for real-life examples concerning each competence that was discussed and make reference to the individual assessment questions formulated by the assessee as well as his or her background. If an assessee did not accept a result, there were several ways to deal with it (for extensive recommendations, see Finn, 1996). Examples are to ask what the result should look like in the assessee's opinion, to communicate to the assessee that results might be wrong, or to formulate a take-home message.

4. Targets. In the last step, any remaining questions were answered. It has been found that adding goal setting to feedback enhances performance after the feedback (Balcazar, Hopkins, & Suarez, 1986); therefore, the assessees were prompted to create two goals that they wanted to attain in the future. They were told that the goals should be related to the TMIC-SA results that were discussed. To close out the SHORT session, the assessor explained that the assessees should feel free to contact the assessor at any time if questions should arise. Also, the assessor thanked the assessees for their openness and reassured them that their results would be confidential.

5.2 The Most Important Findings

Repeated measures linear mixed effects modeling was conducted seven times to examine whether the change over time (pretest vs. posttest) in the seven training-related dependent variables relied on the type of experimental group (treatment, comparison, or control group) that a student belonged to. Each time, we first investigated whether the three groups differed on the pretest. They did not differ for most dependent variables; there were two exceptions, namely the contemplations stage and the action stage. The reader is referred to Schnabel et al. (in press) for the specific results. In

the following, I will show that (a) ICC as measured with the TMIC changes over time, particularly if it is trained with the collaborative test-feedback intervention that was developed ad hoc (interaction effect), (b) on the posttest, students who took part in the collaborative test-feedback intervention outperformed students who received only a written feedback report or no intervention at all, and (c) the collaborative test-feedback intervention was positively evaluated by its recipients. Additional results can be found in Schnabel et al.'s (in press) publication.

The significant interaction effect, $F(1, 228) = 17.65$, $MSE = 0.49$, $p < .001$, $\eta^2 = .14$, showed that the TMIC-SA total score increased over time for individuals who took part in the collaborative test-feedback intervention (treatment group), $\Delta M = 0.224$, 95% CI [0.215, 0.234], decreased for students who did not receive any feedback (control group), $\Delta M = -0.045$, 95% CI [-0.050, -0.041], and remained the same for students who received written feedback (comparison group), $\Delta M = 0.003$, 95% CI [-0.012, 0.006]. In addition, on the posttest, the TMIC-SA total score was highest in the treatment group, $F(2, 227) = 8.05$, $MSE = 1.89$, $p < .001$, $\eta^2 = .07$; $t(227) = -3.79$, $p < .001$, $d = -0.50$.

Moreover, interaction effects were significant for all therapeutic outcome variables: $F(2, 226) = 8.05$, $MSE = 1.10$, $p < .001$, $\eta^2 = .07$ (intercultural self-confidence); $F(2, 226) = 25.31$, $MSE = 8.73$, $p < .001$, $\eta^2 = .18$ (perceived benefit from test participation); $F(2, 226) = 69.55$, $MSE = 13.24$, $p < .001$, $\eta^2 = .38$ (intercultural self-understanding). The students who were part of the treatment group benefitted the most from the assessment process in terms of their intercultural self-understanding, $\Delta M = 1.324$, 95% CI [1.312, 1.336], when compared with those who received a written feedback report, $\Delta M = 0.286$, 95% CI [0.275, 0.298], and those who received no feedback at all, $\Delta M = -0.066$, 95% CI [-0.072, -0.016]. Intercultural self-confidence beliefs increased with participation in the collaborative test-feedback intervention, $\Delta M = 0.356$, 95% CI [0.351, 0.361], and after receiving a written feedback report, $\Delta M = 0.109$, 95% CI [0.105, 0.114], but decreased when no feedback was provided, $\Delta M = -0.037$, 95% CI [-0.039, -0.035]. The students who did not receive any treatment at all as well as the students who took part in the collaborative test-feedback intervention showed an increase in their perceptions of the benefits of participating in the test from pretest to posttest, $\Delta M = 0.187$, 95% CI [0.164, 0.210] and $\Delta M = 1.079$, 95% CI [1.029, 1.128], respectively. However, students who were

provided with only written feedback found the value of taking part in the study lower than when questioned the first time, ΔM = -0.297, 95% CI [-0.343, -0.252]. The second assessment revealed that students who were part of the collaborative test-feedback intervention group scored higher in intercultural self-understanding, $F(2, 225) = 30.01$, $MSE = 19.28$, $p < .001$, $\eta^2 = .21$; $t(225) = -7.29$, $p < .001$, $d = -0.97$, intercultural self-confidence, $F(2, 225) = 5.28$, $MSE = 2.64$, $p < .01$, $\eta^2 = .05$; $t(225) = -3.25$, $p < .01$, $d = -0.43$, and in perceived benefit from test participation, $F(2, 225) = 23.36$, $MSE = 22.24$, $p < .001$, $\eta^2 = .17$; $t(225) = -6.80$, $p < .001$, $d = -0.91$, than students in the written-feedback group or in the no-feedback group.

Another major result was found with regard to precontemplation, which is the stage individuals are in when they have no intention to change their problematic intercultural behavior. The precontemplation score decreased after students participated in the collaborative test-feedback intervention, $F(2, 224) = 10.77$, $MSE = 7.05$, $p < .001$, $\eta^2 = .09$; $\Delta M = -0.383$, 95% CI [-0.392, -0.374], but increased for students who received only written feedback, $\Delta M = 0.057$, 95% CI [0.050, 0.066] or no feedback at all, $\Delta M = 0.154$, 95% CI [0.150, 0.158]. On the posttest, precontemplation scores were lowest for the collaborative test-feedback intervention group when compared with the groups with written or no feedback, $F(2, 224) = 10.77$, $MSE = 7.05$, $p < .001$, $\eta^2 = .09$; $t(224) = 4.47$, $p < .001$, $d = 0.60$.

Students provided a very positive evaluation of SHORT, $M = 5.24$, $SD = 0.82$. Moreover, SHORT significantly added value to the collaborative test-feedback intervention as evaluations increased from $M = 3.78$ ($SD = 0.57$) after the written feedback to $M = 4.97$ ($SD = 0.70$) after SHORT, $t(32) = -9.97$, $p < .001$, $d = -3.52$.

5.2 Implications concerning the TMIC Training Intervention

This study showed that ICC, as measured with the TMIC, is a malleable construct because it can change over time when it is appropriately trained. Thus, TMIC provides a suitable basis for a training intervention that is aimed toward helping people to develop ICC. Such an intervention was developed for the purpose of training students who are going abroad. Despite the fact that students, in comparison with expatriates, often lack intercultural knowledge and experience, fewer training

options are offered to them, mostly due to economic reasons. The collaborative test-feedback intervention is a time- and cost-saving option that is yet effective and highly accepted by its recipients. Specifically, the following effects can be attained by applying collaborative test feedback: (a) a broadening of students' intercultural self-understanding, (b) a decrease in students' passive acceptance of their intercultural weaknesses, (c) an increase in students' motivation to change, and (d) an increase in students' intercultural self-efficacy.

Chapter 6: Cross-Cultural Validation of a Short Version of the Test to Measure Intercultural Competence (TMIC-S)

This part of the monograph addresses the cross-cultural validation of a short version of the TMIC (TMIC-S) and presents the primary findings from Schnabel et al.'s (2014) article (for the pre-print version of this article see Appendix D).

6.1 Study Background

Despite many advantages of the highly differentiated multidimensional TMIC model, some challenges exist. The TMIC with its 17 factors and two methods is a time consuming instrument to administer. Taking into consideration the fact that ICC is only one component that is assessed in the employee selection process, a shorter version seems to be more suitable when applied for selection rather than for employee development purposes. In addition, replicating a complex 17-factor structure across cultures is challenging. Therefore, a short version was developed and tested across cultures. This cross-cultural generalization study focused on Germany and Brazil. Brazil, as the comparison culture, was chosen for the following reasons. First, the Brazilian culture differs greatly from the German culture. This is especially true for its localization on the cultural dimensions (Hofstede, 2001) and for its history of migration. On the one hand, through colonialization, Brazil was affected by various cultural influences much more than Germany was (Gawora, de Souza Ide, & Barbosa, 2011). On the other hand, Germany looks back on a recent wave of migration that began after the Second World War. Second, Germany's economy is highly established, whereas Brazil is one of the currently fastest growing economies worldwide; it belongs to the so-called BRICS states (O'Neill, 2001). This comes along with an increase in international traffic. Third, whereas expatriation and impatriation have been part of the corporate practice for many years in Germany, sending employees abroad and integrating foreign employees in organizations is quite a new phenomenon in Brazil. This increases the need for appropriate instruments in the field of intercultural employee selection and development (Muritiba, Muritiba, Campanário, & de Albuquerque, 2010).

This study was conducted in the German and in the Brazilian university context. Of the N = 1,037 Germans who participated in the study, 597 were women (58%), 429 were men (41%), and 11

were missing data on sex (1%). The average age was 27.96 years (*SD* = 9.47). In total, *N* = 769 Brazilian participants were included in the study. The 415 women (54%) and 354 men (46%) were 27.38 years old (*SD* = 10.61) on average.

The short version of the TMIC was obtained as follows. From each of the six second-order factors, one first-order factor was selected. The selection was based on (a) the theoretical importance reported in previous research, (b) the hypothesized independency from any cultural preference system, and (c) the statistical fit of the model (cf. Schnabel et al., 2014). TMIC-S includes six factors: sensitivity in communication, information seeking, socializing, cultural identity reflection, goal setting, and mediation of different interests. They are assessed with 25 self-appraisal and six situational judgment items. All Portuguese items, which were generated through a process of translation and back-translation (Brislin, 1970), are available in the supplemental material.

6.2 The Most Important Findings

The hypothesized TMIC-S model with its six factors and two methods was tested for its fit to the data; this led to satisfactory fit indices in the German sample, $\chi^2(419, N = 1{,}037) = 824.10$, $p < .001$; $\chi^2/df = 1.97$; RMSEA = .036, 90% CI [.033, .040]; WRMR = 1.095; CFI = .913; TLI = .904, as well as in the Brazilian sample, $\chi^2(419, N = 769) = 919.35$, $p < .001$; $\chi^2/df = 2.19$; RMSEA = .039, 90% CI [.036, .043]; WRMR = 1.168; CFI = .902; TLI = .892. Internal consistencies ranged from acceptable to good for both the German, $\alpha = .72$ to .86, and the Brazilian scales, $\alpha = .65$ to .77. Measurement invariance (Meredith, 1993) of the TMIC-S was investigated using multigroup confirmatory factor analysis (MG-CFA) in Mplus. The cut-off values for measurement invariance were: $\Delta CFI \leq .010$, $\Delta TLI \leq .010$, $\Delta RMSEA \leq .015$, and a Bonferroni-Holm-corrected $p < .01$ for $\Delta\chi^2$ (Chen, 2007; Holm, 1979). The results for a model in which all parameters were freely estimated in each group clearly pointed to configural invariance, $\chi^2(838, N = 1{,}806) = 1742.01$, $p < .001$; $\chi^2/df = 2.08$; RMSEA = .038, 90% CI [.035, .040]; WRMR = 1.601; CFI = .908; TLI = .898. Then the factor loadings were held equal in both groups, $\chi^2(863, N = 1{,}806) = 1678.93$, $p < .001$; $\chi^2/df = 1.95$; RMSEA = .036, 90% CI [.033, .038]; WRMR = 1.682; CFI = .917; TLI = .910. The relevant delta coefficients supported metric invariance, $\Delta\chi^2(25) = 46.874$, $p = .01$; $\Delta RMSEA = .002$; $\Delta CFI = .009$.

Only ΔTLI = .012 did not support metric invariance. To investigate scalar invariance, the intercepts of the self-report items and thresholds of the situational judgment items were additionally restricted; this led to an acceptable model fit, $\chi^2(882, N = 1{,}806) = 1802.26$, $p < .001$; $\chi^2/df = 2.04$; RMSEA = .037, 90% CI [.035, .040]; WRMR = 1.752; CFI = .906; TLI = .901. The comparison of fit indices between the metric and the scalar invariance models resulted in rather inconsistent results. The chi-square difference was significant, $\Delta\chi^2(19) = 241.12$, $p < .001$; yet, it was difficult to interpret because of the large sample sizes. On the one hand, scalar invariance was supported by the ΔRMSEA = .002 and the ΔTLI = .009. On the other hand, the ΔCFI showed an unexpected pattern in that the ΔCFI = .011 value contradicted scalar invariance, but ΔCFI = .002 evolved when comparing the configural with the scalar invariance model. Given that the CFI values of the configural and scalar invariance models were very similar, we interpreted the ΔCFI values as considerably supportive of scalar invariance.

The construct validity of the TMIC-S was supported by mostly moderate positive correlations between the TMIC-S factors and the CQS scales with a range of $r(313) = .18$ to .78 in the German sample and $r(769) = .25$ to .64 in the Brazilian sample. The comparison of a model in which all parameters were freely estimated and another model in which all 24 pairs of correlations were held equal in both samples revealed a significant chi-square difference, $\chi^2(24) = 45.49$, $p < .01$. Yet, all other delta coefficients, ΔRMSEA = .001; ΔCFI = .007; ΔTLI = .008, showed that the correlations between the TMIC-S and the CQS were equal in the two samples.

Moreover, several multigroup SEM comparisons were conducted to test for criterion validity. In both samples and for all six factors, I investigated the differences in the TMIC-S latent means while taking the following three external criteria into consideration: (a) previously versus never participated in an intercultural training, (b) involved versus not involved in intercultural matters, and (c) stayed abroad for longer than 3 months versus less than 3 months. The criterion validity of the TMIC-S was supported as interculturally experienced individuals scored higher on most factors in both samples; see Schnabel et al. (2014) for the detailed results.

6.3 Implications concerning the TMIC-S

This study is one of the very few attempts in ICC research to support a universal model of ICC with the help of measurement invariance testing. The TMIC-S is a short version of the TMIC (Schnabel et al., 2015) that was specifically developed for settings such as employee selection in which more than one instrument is applied. The TMIC-S avoids cognitively overloading the test taker (Eignor, 2013). The results showed that (a) the six-factor multimethod TMIC-S model fitted the data well and fulfilled reliability as well as validity requirements, (b) metric invariance of the TMIC-S was supported by most and scalar invariance by all relevant delta coefficients, and (c) there was strong evidence for a common model of ICC in two highly diverse cultures. Especially the last implication can be seen as a promising foundation for further cross-cultural comparison studies that are focused on validating a universal ICC model.

Chapter 7: Overarching Discussion

The following chapter provides an overall discussion of this monograph. First, the major findings are summarized with respect to the main research questions. Second, benefits, limitations, and future research topics are reviewed. Third, a conclusion is provided.

7.1 Summary of this Monograph organized with respect to the Main Research Questions

Growing globalization, high migration rates, and immensely diverse immigrant groups have significantly changed the social reality, including modes of living, working, and studying (Fantini, 2009; Vertovec, 2007). Hence, the importance of ICC instruments, which select, guide, and train individuals with regard to mastering their intercultural challenges, is constantly increasing. Over the past 20 years, a large body of research has emerged from various disciplines, resulting in fuzzy models and instruments of ICC.

To overcome the limitations of existing approaches the current research focused on the development and validation of a new theoretical framework, a cross-cultural multimethod test, and a collaborative assessment intervention in the field of ICC. Specifically, the thus-far missing interconnection between defining, assessing, and training ICC (Schnabel et al., 2015) was established as part of this research.

This monograph integrated three different papers (A, B, and C) that covered four major research questions. Chapter 3 built on Paper A and focused on the derivation of the ICC framework. The new definition of ICC as a malleable multidimensional construct as well as the onion model of ICC that organizes the ICC facets and other relevant constructs were conceptualized with the help of an extensive literature review, nine exploratory expert interviews, and a quantitative pretest. Chapter 4 also referred to Paper A and addressed the development and validation of the TMIC. The TMIC matches with the ICC framework and fulfills the relevant requirements that were formulated in previous research and in the exploratory expert interviews. It incorporates 17 first-order and six second-order factors assessed with self-appraisal and situational judgment items. The procedure and the results of two studies, one conducted with students and the other one with professionals, were outlined. Chapter 5 presented a collaborative test-feedback intervention that was developed ad hoc, in

which participants took the TMIC-SA and received written and graphical feedback on their test results as well as a 1-hr telephone test feedback (SHORT). In a randomized pretest-posttest control group design with students awaiting an ERASMUS year, this intervention was compared with taking part in the TMIC only and with receiving written feedback after TMIC participation. Chapter 6 contained the key findings originally presented in Paper C, which was aimed at closing an important gap in the field of ICC assessment, namely the cross-cultural validity of ICC. Chapter 6 also addressed the development of a short version of the TMIC, which is especially applicable in employee selection settings. Findings concerning the model fit, psychometric properties, and measurement invariance of the TMIC-S from a cross-cultural validation study that integrated samples from Germany and Brazil represented the core of Chapter 6.

7.2 Benefits, Limitations, and Outlook of this Monograph

Below, I will present the key aspects that create an added value for the ICC research community together with important limitations of this work. In addition, I will attempt to stimulate ideas for future research in the field.

7.2.1 TMIC framework. Whereas existing approaches merely include personality traits, attitudes, worldviews, or a combination of different individual characteristics, the TMIC framework focuses on malleable aspects that can be directly developed through training and coaching. As competences are manifested through behavior, the TMIC framework defines ICC as a global behavioral orientation toward perceiving and dealing with intercultural differences. Moreover, this research is the first attempt at structuring the various aspects of an individual in the intercultural context through an onion model. In contrast to other ICC approaches, the key features of the TMIC framework, such as the malleable nature of ICC or the existence of the different layers in the onion model, were empirically supported. Relevant indicators in this context were, for example, the moderate relation between the TMIC framework and CQ as well as ICC's responsiveness to training as reflected by the TMIC model. In addition, the expert perspective was considered while developing the TMIC framework, thus building a necessary bridge between research and practice.

However, there are some concerns that should be discussed with respect to the malleable nature of ICC. So far, we still do not know how malleable the various TMIC factors actually are. In the long run, the stable trait concept has to be empirically distinguished from the ability concept. This would require a profound analysis of the discriminant validity to learn more about what ICC is not. The first indicators of this were the relatively weak (TMIC-SA) and nonsignificant (TMIC-SJT) correlations with openness to experience, a rather stable personality trait (Goldberg, 1990). As the multidimensionality and the understanding of ICC in terms of a context-specific competence to act stretch ICC substantially, a clear definition of the scope of ICC will be an important topic for future research. Thereby, it would be interesting to examine the incremental validity of different ICC concepts with respect to an individual's integration, intercultural success, and satisfaction.

7.2.2 Seventeen-factor solution and second-order factor model. The factor structure of many ICC instruments has been explored (e.g., Hammer, 2011; Van der Zee & Van Oudenhoven, 2000), and there is an ongoing debate about the existence of a general factor of ICC, thus calling into question the multidimensional nature of ICC (Matsumoto & Hwang, 2013). Schnabel et al. (2015) made an important contribution to this discussion when they showed that the TMIC second-order model with only one second-order factor was inferior to a model with six second-order factors. In addition, these authors tested several alternative models with varying factor numbers against the 17-factor solution and showed that the 17-factor solution consequently surpassed all other solutions. Consequently, the 17-factor TMIC model was satisfactory not only because of its large number of factors or because the loadings were allowed to be freely estimated by using ESEM (Asparouhov & Muthén, 2009). This allows for a highly differentiated evaluation of an individual's ICC and provides the perfect basis for intercultural training and coaching sessions. The advantage of a highly multidimensional construct is that the specific details of the construct can be measured separately. For example, it can be assumed that a person can change his or her communication style easily even if he or she finds it hard to get to the point. Both aspects belong to communication competence, but one belongs to flexibility in communication and the other one to clarity in communication. However, this might result in cross-loadings, which are not tolerated by traditional CFA, and hence the model fit would suffer. ESEM (Asparouhov & Muthén, 2009), which was used to determine the fit of the 17-

factor model, overcomes this limitation by including cross-loadings while simultaneously offering many of the benefits of CFA (Asparouhov & Muthén, 2009). As second-order models could not be examined with ESEM in Mplus, CFA had to be applied. This resulted in CFIs and TLIs that did not meet the proposed cut-off values (Chen, 2007). As ESEM is developed further, it would be interesting to examine the second-order model again and compare the results.

7.2.3 Multimethod approach. Recent research has called for ICC instruments that apply more than one method (e.g., Deardorff, 2006; Gelfand et al., 2008; Leung et al., 2014; Sinicrope et al., 2007). The TMIC is the first test in the intercultural research area to combine SJT and standard Likert-scale items. The SJT (TMIC-SJT) complements the Likert scale (TMIC-SA), which represents an individual's self-concept, in that it assesses behavioral preferences in typical and critical situations in the intercultural context (Bledow & Frese, 2009; Schnabel et al., 2014, 2015). Moreover SJTs are generally less prone to the influence of social desirability (McDaniel et al., 2006), and this was a major requirement for an ICC instrument as identified in the expert interviews. In the future, it might be possible to implicitly control for under- and overestimations of competence by comparing the results between the TMIC-SJT and the TMIC-SA. That is, we could examine a TMIC profile, and if we find systematic deviations between the TMIC-SJT and the TMIC-SA, we might conclude that a biased response pattern exists. As social desirability was not explicitly measured in this study, evidence for this procedure will need to be collected in future studies. Despite the positive aspects of the TMIC-SJT, it is rather challenging in terms of its psychometric properties. For example, the incremental validity of the TMIC-SJT with respect to external criteria was significant but rather small. Also, most but not all of the situational judgment items were found to be significantly correlated with the appropriate self-appraisal factor. Although the loadings for the SJT on the corresponding factors were significant on the one hand, they were predominantly small on the other hand. Whetzel and McDaniel (2009) argued that SJTs are seldom unidimensional. Unlike with self-report instruments, which measure multiple constructs, SJTs rarely lead to a clearly interpretable factor structure. Thus, it might be worthwhile to compare the current TMIC model with a restructured version in which all situational judgment items are summarized in one additional factor. This 18th factor would then function as the behavioral component of ICC. In terms of my aim to create an instrument that could

function cross-culturally, more research is also needed to understand the unique challenges of applying SJTs in different languages and cultures. Schnabel et al. (2014) already found that the measurement invariance of the TMIC-S model was affected by some situational judgment items. Hence, it would be interesting to further investigate whether the behavioral manifestation, as measured with the TMIC-SJT, of a latent culture-transcending competence actually depends on cultural standards. For example, Kong (2014) applied the TMIC in a Chinese context. In a think-aloud study, Chinese test participants struggled with the behavioral alternatives given on the TMIC-SJT. Therefore, Kong (2014) suggested a new item format for the TMIC-SJT in which test participants were first asked to develop their own answer alternative, and second, to read the proposed answer alternatives and select the one that best resembled their own answer. Clearly, the strengths and weaknesses of this format have to be evaluated through quantitative studies as well.

As mentioned earlier, integrating the SJT as a second method is a clear advantage of the TMIC. However, the TMIC is still not an objective measure of ICC. Conducting a validation study to compare the TMIC test results with the ratings of colleagues, supervisors, customers, professors, and so forth is a fortiori important.

7.2.4 TMIC training intervention. The collaborative test-feedback intervention described in Chapter 4 and based on Paper B is novel in many ways. First, collaborative assessment was previously applied only in clinical settings. Second, until now, no feedback intervention was available in the intercultural field. Third, it is the first intercultural intervention that is directly connected to an assessment instrument. Fourth, the intervention is highly economical because it is time- and cost-saving on the one hand and effective on the other hand. Its effectiveness was supported by the large effect sizes in the treatment condition (cf. Schnabel et al., in press). Thus, it can be offered to groups that are often put at a disadvantage with respect to intercultural training due to concerns with time and cost (e.g., students). The sophisticated experimental design allowed a detailed evaluation of the intervention to be made, including its incremental validity when compared with test participation only or with written test feedback. Thereby, ICC was measured across different points in time, and this was one of the major requirements found in the exploratory expert interviews. In addition, the integration

of a comparison and a control group as well as the satisfactory retest reliability of the instrument provided the first indications that the TMIC could be applied repeatedly to track change. However, I have not yet tested for measurement invariance across multiple points in time.

The results also strongly supported the malleable nature of the TMIC model as participants' competence levels increased significantly as a consequence of the collaborative test-feedback intervention. Nevertheless, their ICC development was measured only via self-report. Different methods (e.g., the TMIC-SJT) or objective measures (e.g., other-ratings) have to be included in the future to support the initial findings that favored the TMIC model's responsiveness to training. Further, more research is needed to validate the intervention in other important samples, for example, those consisting of expatriates.

The collaborative test-feedback intervention is a promising starting point for broadening the TMIC-based training approach in the future. New training concepts that integrate the TMIC factors may emerge. Hereby, the TMIC can function as a measure that represents training needs beforehand and evaluates training success afterwards.

7.2.5 Cross-cultural validity of the TMIC-S. Paper C examined the cross-cultural validity of the TMIC-S using data from Germany and Brazil. Very few studies have looked at the cross-cultural validity of ICC constructs. Thus, the study in Paper C contributed important empirical evidence to fill the gap. The comparison of models with varying parameter restrictions led to promising delta coefficients in terms of their reference to measurement invariance. The results moreover strongly supported the hypothesis that the TMIC framework can function in at least one other culture that is very different from the culture of origin.

The main limitation is the restriction of the sample to participants from Germany and Brazil. Further cross-cultural comparisons are needed to make a clear statement about the universal nature of the TMIC framework. To date, culture has not been explicitly measured. Future cross-cultural comparisons should therefore include a scale to assess individual scores on various cultural dimensions (e.g., Hofstede, 2001). This would lead to a differentiated investigation of measurement invariance that distinguishes between ethnicity and language.

An additional statistical concern is that this study focused on data at the individual level; however, it would be worth examining the reliability at the group level, the variance that can be attributed to individuals from different countries, and also the extent to which ratings are similar in different countries. Moreover, it would be beneficial to study underlying issues such as response patterns within each of the samples. Because the study was focused solely on analyses at the individual level, this might be particularly useful. The reasoning behind this is the following. Similarities in overall scores may hide different response patterns in different samples. Similar means may result from acquiescence, extreme responses, or other biases, which are issues that are worth exploring in a cross-cultural comparison of a new measure (Van de Vijver & Leung, 1997; Van de Vijver & Tanzer, 2004).

7.2.6 Study design. To date, the TMIC was applied at two levels of professional experience, in two cultures, and at two points in time. As the TMIC and TMIC-S models showed some methodological complexity, future studies should extend the application radius and investigate the TMIC model in different age, cultural, and occupational groups. Besides the aforementioned possibilities to realize that, a further option could be to accompany expatriates during their assignment process and thereby to investigate the predictive validity of the TMIC with respect to experiences of integration, success, and satisfaction abroad. Moreover, the TMIC-S was designed to be used in employee selection; however, it was never tested in this setting. Thus, it would be interesting to examine whether response distortions and faking tendencies, which often apply to employee selection situations (Whetzel & McDaniel, 2009), influence the validity of the TMIC-S.

7.3 Conclusion

The TMIC, along with its comprehensive framework, interconnected training opportunities, different language versions, and versions of different lengths is a promising instrument in the intercultural field. The malleable nature of the TMIC construct and the integration of two methods fill an important gap in the research and practice of ICC. Ample investigations of the underlying factor structure and the psychometric properties showed that the TMIC, including all its versions, is reliable and valid at the same time. Thus, it is ready to be applied in intercultural counseling, training, and

coaching settings as well as in employee selection. Furthermore, it creates opportunities for ICC research to make cross-cultural comparisons. Most importantly, it presents an opportunity for rethinking a diverse research area.

References

Aiello, A., & Areni, A. (1998). Un aggiornamento della Scala di De Grada ed altri (1975) per la misura dell'Etnocentrismo [An updating of the attitude scale of De Grada et al. (1975) for the measurement of ethnocentrism], *Rassegna di Psicologia*, *15*, 145−160. Retrieved from http://www.francoangeli.it/riviste/Scheda_Rivista.aspx?IDArticolo=11511&idRivista=39

Allen, A., Montgomery, M., Tubman, J., Frazier, L., & Escovar, L. (2003). The effects of assessment feedback on rapport-building and self-enhancement processes. *Journal of Mental Health Counseling, 25*, 165−181. Retrieved from: http://essential.metapress.com/content/lw7h84q861dw6ytj/

Al-Rodhan, N. R. F., & Stoudmann, G. (2006). *Definitions of globalization: A comprehensive overview and a proposed definition.* Retrieved from Geneva Center for Security Policy website: http://gcsp.ch/content/download/1267/9834/file/Definitions%20of%20Globalization%20-%20A%20Comprehensive%20Overview%20and%20a%20Proposed%20Definition.pdf

Altshuler, L., Sussman, N. M., & Kachur, E. (2003). Assessing changes in intercultural sensitivity among physician trainees using the Intercultural Development Inventory. *International Journal of Intercultural Relations, 27*, 387−401. doi:10.1016/S0147-1767(03)00029-4

American Psychological Association (1986). *Guidelines for computer-based tests and interpretations.* Washington, DC: American Psychological Association.

American Psychological Association (2010). *Ethical principles of psychologists and code of conduct.* Washington, DC: American Psychological Association.

Anderson, P. H., Lawton, L., Rexeisen, R. J., & Hubbard, A. C. (2006). Short-term study abroad and intercultural sensitivity: A pilot study. *International Journal of Intercultural Relations, 30*, 457−469. doi:10.1016/j.ijintrel.2005.10.004

Ang, S., Rockstuhl, T., & Tan, M. L. (in press). Cultural intelligence and competencies. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral*

sciences. Retrieved from http://soonang.com/wp-content/uploads/2014/03/Paper_CQ-and-Competencies-Encyclopedia-CQ.pdf

Ang, S., & Van Dyne, L. (Eds.). (2008). *Handbook of cultural intelligence*. New York, NY: Sharpe.

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural intelligence: Its measure and effects on cultural judgment and decision making, cultural adaptation and task performance. *Management and Organization Review*, *3*, 335–371. doi:10.1111/j.1740-8784.2007.00082.x

Arasaratnam, L. A. (2006). Further testing of a new model of intercultural communication competence. *Communication Research Reports*, *23*, 93–99. doi:10.1080/08824090600668923

Arasaratnam, L. A. (2007). Research in intercultural communication competence. Past perspectives and future directions. *Journal of International Communication, 13*, 66–73. doi:10.1080/13216597.2007.9674715

Arasaratnam, L. A. (2009). The development of a new instrument of intercultural communication competence. *Journal of Intercultural Communication, 20*, 1–11. Retrieved from http://www.immi.se/intercultural/

Ascalon, M., Schleicher, D., & Born, M. (2008). Cross-cultural social intelligence: An assessment for employees working in cross-national contexts. *Cross Cultural Management, 15*, 109–30. doi: 10.1108/13527600810870570

Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 397–438. doi:10.1080/10705510903008204

Auernheimer, G. (1996). Mit kultureller Differenz umgehen lernen [Learn to cope with cultural differences]. *Pädagogik, 11*, 50–52. Retrieved from http://www.redaktion-paedagogik.de/abo-hefte/

Bacon, F. (1696). *The essays, or councils, civil and moral, of Sir. Francis Bacon. With a table of the colours of good and evil. And a discourse of the wisdom of the ancients*. London, United

Kingdom: E. Holt, Henry Herringman, Timothy Childe. Retrieved from https://archive.org/details/essaysorcouncils00baco

Balcazar, F., Hopkins, B. L., & Suarez, Y. (1986). A critical, objective review of performance feedback. *Applied Psychology*, *66*, 314−323. doi:10.1300/J075v07n03_05

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behaviour* (Vol. 4, pp. 71−81). New York, NY: Academic Press.

Bangert-Drowns, R. L., Kulik, C. L., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*, 213–237. doi:10.3102/00346543061002213

Barker, C., Pistrang, N., & Elliott, R. (2005). *Research methods in clinical psychology: An introduction for students and practitioners* (2nd ed.). New York, NY: Wiley.

Bennett, M. J. (1986). A development approach to training for intercultural sensitivity. *International Journal of Intercultural Relations, 10*, 179–196. doi:10.1016/0147-1767(86)90005-2

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Faige (Ed.), *Education for the intercultural experience* (2nd ed., pp. 21–77). Yarmouth, ME: Intercultural Press.

Berry, J. W., Kim, U., Power, S., Young, M., & Bujaki, M. (1989). Acculturation in plural societies. *Applied Psychology: An International Review, 38*, 185–206. doi:10.1111/j.1464-0597.1989.tb01208.x

Bhawuk, D., & Brislin, R. (1992). The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International Journal of Intercultural Relations*, *16,* 413–436. doi:10.1016/0147-1767(92)90031-O

Bird, A., Mendenhall, M., Stevens, M. J., Oddou, G. (2010). Defining the content domain of intercultural competence for global leaders. *Journal of Managerial Psychology, 28,* 810–828. doi:10.1108/02683941011089107

Birkhan, G., & Ringelband, O. J. (2013). Rückmeldung der Eignungs-/Potenzialbeurteilung an den Kandidaten [Feedback for the candidate in aptitude-/potential analysis]. In W. Sarges (Ed.),

Management-Diagnostik [Management Diagnostics] (4th ed., pp. 931–939). Göttingen, Germany: Hogrefe.

Bledow, R., & Frese, M. (2009). A situational judgment test of personal initiative and its relationship to performance. *Personnel Psychology, 62*, 229–258. doi:10.1111/j.1744-6570.2009.01137.x

Bogner, A., Littig, B., & Menz, W. (2005). *Das Experteninterview: Theorie, Methode, Anwendung* [The expert interview: Theory, method, application] (2nd ed.). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Bolten, J. (2001). Interkulturelle Assessment-Center [Intercultural assessment center]. In W. Sarges (Ed.), *Weiterentwicklungen der Assessment Center-Methode* [Further developments of the assessment center method] (2nd ed., pp. 213–218). Göttingen, Germany: Hogrefe.

Bolten, J. (2007). Was heißt Interkulturelle Kompetenz? Perspektiven für die internationale Personalentwicklung [What does intercultural competence mean? Perspectives for international personnel development]. In V. Künzer & J. Berninghausen (Eds.), *Wirtschaft als interkulturelle Herausforderung* [Economy as an intercultural challenge] (pp. 21–42). Frankfurt am Main, Germany: IKO.

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae. Manual* [Neo-Five-Factor Inventory (NEO-FFI) after Costa and McCrae. Manual]. Göttingen, Germany: Hogrefe.

Bortz, J. (2006). *Statistik für Human- und Sozialwissenschaftler* [Statistics for human and social scientists] (6th ed.). Heidelberg, Germany: Springer Medizin.

Bremer, D. (2006). Wanted: Global workers. *International Educator, 15*(3), 40–45. Retrieved from http://www.nafsa.org/_/File/_/global_workers_inted_2006.pdf

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology 1*, 185–216. doi:10.1177/135910457000100301

Brislin, R. W. (1981). *Cross-cultural encounters. Face-to-face encounters*. New York, NY: Pergamon.

Bundesinstitut für Berufsbildung (2013). *Fragen zum Fachkräftemangel in Deutschland* [Questions concerning the skills shortage in Germany]. Retrieved from http://www.bibb.de/dokumente/pdf/a22_qube_FAQ_2Welle.pdf

Bücker, J., & Poutsma, E. (2010). Global management competencies: A theoretical foundation. *Journal of Managerial Psychology, 25*, 29–44. doi:10.1108/02683941011089116

Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence.* Philadelphia, PA: Multilingual Matters.

Byham, T. M. (2005). *Factors affecting the acceptance and application of development feedback from an executive assessment program*. (Unpublished doctoral dissertation). University of Akron, Akron, OH.

Campbell, D. T., & Fiske, D. W. (1959). Convergent discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105. doi:10.1037/h0046016

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin*, *92*, 111–135. doi:10.1037/0033-2909.92.1.111

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. doi:10.1080/10705510701301834

Chen, G.-M., & Starosta, W. J. (1998). *Foundations of intercultural communication*. Boston, MA: Allyn & Bacon.

Chen, G.-M., & Starosta, W. J. (2000). The development and validation of the Intercultural Sensitivity Scale. *Human Communication, 3*, 1–15. Retrieved from http://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1035&context=com_facpubs

Chiu, C. Y., & Hong, Y. Y. (2005). Cultural competence: Dynamic processes. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 489–505). New York, NY: Guilford Publications.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO PI-R) and NEO five factor inventory (NEO FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Crooks, T. (2001, September). *The validity of formative assessments*. Paper presented at the Sixth British Educational Research Association Annual Conference, Leeds, England.

Dauner, A. B. (2011). Die organisierende Funktion von Kompetenz—Entwurf eines zwei-Ebenen-Modells interkultureller Kompetenz [The organizing function of competence—Draft of a two-level-model of intercultural competence]. *Interculture Journal*, *10*(14), 25–46. Retrieved from http://www.interculture-journal.com/index.php/icj/article/view/115/194

Deardorff, D. K. (2004). *The identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States* (Unpublished doctoral dissertation). North Carolina State University, Raleigh, NC.

Deardorff, D. K. (2006). *Policy Paper zur Interkulturellen Kompetenz* [Policy paper about intercultural competence]. Gütersloh, Germany: Bertelsmann Stiftung.

Deardorff, D. K. (Ed.). (2009). *The Sage handbook of intercultural competence.* Thousand Oaks, CA: Sage.

Deardorff, D. K., & Hunter, W. (2006). Educating global-ready graduates. *International Educator, 15*(3), 72–83. Retrieved from http://www.nafsa.org/Find_Resources/Publications/Periodicals/International_Educator/International_Educator_-_May_June_2006/

Der-Karabetian, A. (1992, August). *World-mindedness and the nuclear threat: A four-nation test of the superordinate goal hypothesis*. Paper presented at the annual meeting of the American Psychological Association, Washington, DC.

Deutscher Akademischer Austauschdienst (2014). *Erasmus-Mobilität auf neuem Rekordniveau* [Erasmus mobility reaches all-time record]. Retrieved from https://www.daad.de/portrait/presse/pressemitteilungen/2014/25936.de.html

Deutscher Reiseverband e.V. (2013). Fakten und Zahlen 2012 zum deutschen Reisemarkt [Facts and figures 2012 concerning the German tourism industry]. Retrieved from http://www.drv.de/fileadmin/user_upload/Fachbereiche/Statistik_und_Marktforschung/DRV_Zahlen_Fakten2012.pdf

DiClemente, C. C., & Prochaska, J. O. (1998). Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors* (2nd ed., pp. 3−24). New York, NY: Plenum Press.

Earley, P. C., & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures.* Palo Alto, CA: Stanford University Press.

ECA International (2012). *Managing Mobility Survey.* Retrieved from http://www.eca-international.com/shop/policy_and_salary_benchmarking/managing_mobility_survey

Eignor, D. R. (2013). *The standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Erpenbeck, J. (2010). Vereinfachung durch Komplexität. Persönlichkeitseigenschaften und Kompetenzen [Simplification through complexity. Personality traits and competences]. *Sitzungsberichte der Leibniz-Sozietät der Wissenschaften zu Berlin*, *108*, 79–91. Retrieved from http://leibnizsozietaet.de/wp-content/uploads/2012/11/06_erpenbeck.pdf

Erpenbeck, J. (2012). Siegeszug trotz Kritik: Warum Kompetenzdenken notwendig ist [Triumphal procession despite criticism: Why it is important to focus on competences]. *Wirtschaftspsychologie Aktuell*, *19*, 19–22. Retrieved from http://www.wirtschaftspsychologie-aktuell.de/

Erpenbeck, J., & von Rosenstiel, L. (Eds.). (2007). *Handbuch Kompetenzmessung. Erkennen, Verstehen und Bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis* [Handbook competence assessment. Identifying, understanding, and evaluating competences in the organizational, pedagogical, and psychological practice] (2nd ed.). Stuttgart, Germany: Schäffer-Poeschel.

European Union (2014). *Erasmus Mundus Programme.* Retrieved from http://eacea.ec.europa.eu/erasmus_mundus/

Fantini, A. E. (1995). Language, culture, and world view: Exploring the nexus. *International Journal of Intercultural Relations, 19,* 143–153. doi:10.1016/0147-1767(95)00025-7

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

Fantini, A. E., & Tirmizi, A. (2006). Exploring and assessing intercultural competence. *World Learning Publications, 1,* 1–75. Retrieved from http://digitalcollections.sit.edu/worlslearning_publications/1

Finn, S. E. (1996). *Manual for using the MMPI-2 as a therapeutic intervention*. Minneapolis, MN: University of Minnesota Press.

Finn, S. E. (2007). *In our clients' shoes: Theory and techniques of therapeutic assessment*. Mahwah, NJ: Earlbaum.

Finn, S. E., & Tonsager, M. E. (1992). Therapeutic effects of providing MMPI-2 test feedback to college students awaiting therapy. *Psychological Assessment, 4*, 278–287. doi:10.1037//1040-3590.4.3.278

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment, 9,* 374–385. doi:10.1037/1040-3590 .9.4.374

Fischer, C. T. (1994). *Individualizing psychological assessment.* Mahwah, NJ: Erlbaum.

Fischer, C. T. (2000). Collaborative, individualized assessment. *Journal of Personality Assessment, 74*, 2–14. doi:10.1207/S15327752JPA740102

Fischer, C. T., & Finn, S. E. (2008). Developing the life meaning of psychological test data: Collaborative and therapeutic approaches. In R. P. Archer & S. R. Smith (Eds.), *Personality assessment* (pp. 379–404). New York, NY: Routledge.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*, 327–358. doi:10.1037/h0061470

Gabrenya, W. K., Jr., Griffith, R. L., Moukarzel, R. G., Pomerance, M. H., & Reid, P. A. (2012). Theoretical and practical advances in the assessment of cross-cultural competence. In D. D. Schmorrow & D. M. Nicholson (Eds.), *Advances in design for cross-cultural activities* (pp. 317–332). Boca Raton, FL: Taylor Francis.

Gabrenya, W. K., Jr., Moukarzel, R. G., Pomerance, M. H., Griffith, H., & Deaton, J. (2011). *A validation study of the defense language office framework for cross-cultural competence*. Retrieved from Defense Equal Opportunity Management Institute website: http://www.defenseculture.org/Research/EmergeResearchFiles/AssessmentandMeasurement/ValidationDLOFramework.pdf

Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Benson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, *72*, 493–511. doi:10.1037//0021-9010.72.3.493

Gawora, D., de Souza Ide, M. H., & Barbosa, R. S. (Eds.). (2011). *Traditionelle Völker und Gemeinschaften in Brasilien. Lateinamerika-Dokumentationsstelle* [Traditional peoples and communities in Brazil. Latin-America documentation center]. Kassel, Germany: Kassel University Press.

Gelfand, M. J., Imai, L., & Fehr, R. (2008). Thinking intelligently about cultural intelligence: The road ahead. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 375–387). New York, NY: Sharpe.

Gertsen, M. C. (1990). Intercultural competence and expatriates. *The International Journal of Human Resource Management*, *1*, 341–362. doi:10.1080/09585199000000054

Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *The Academy of Management Review*, *18*, 694–734. doi:10.2307/258595

Goldberg, L. R. (1990). An alternative "description of personality": The Big Five factor structure. *Psychological Assessment, 4,* 26–42. doi:10.1037/0022-3514.59.6.1216

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.

Greenholtz, J. F. (2005). Does intercultural sensitivity cross cultures? Validity issues in porting instruments across languages and cultures. *International Journal of Intercultural Relations*, *29*, 73–89. doi:10.1016/j.ijintrel.2005.04.010

Griffith, D. A., & Harvey, M. G. (2000). An intercultural communication model for use in global interorganizational networks. *Journal of International Marketing, 9*(3), 87–103. doi:10.1509/jimk.9.3.87.19924

Gudykunst, W. B. (1994). *Bridging differences: Effective intergroup communication* (2nd ed.). London, United Kingdom: Sage.

Gullahorn, J. R., & Gullahorn, J. E. (1962). An extension of the U-curve hypothesis. *Journal of Social Issues, 3*, 33–47. doi:10.1111/j.1540-4560.1963.tb00447.x

Gunkel, L. (2014). *Akzeptanz und Wirkung von Feedback in Potenzialanalysen* [Acceptance and effects of feedback in potential analyses]. Wiesbaden, Germany: Springer VS.

Hammer, M. R. (2008). The Intercultural Development Inventory (IDI): An approach for assessing and building intercultural competence. In M. A. Moodian (Ed.), *Contemporary leadership and intercultural competence: Understanding and utilizing cultural diversity to build successful organizations* (pp. 245–259). Thousand Oaks, CA: Sage.

Hammer, M. R. (2011). Additional cross-cultural validity testing of the Intercultural Development Inventory. *International Journal of Intercultural Relations*, *35*, 474–487. doi:10.1016/j.ijintrel.2011.02.014

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The Intercultural Development Inventory. *International Journal of Intercultural Relations, 27*, 421–443. doi:10.1016/S0147-1767(03)00032-4

Hansen, K. P. (2000). *Kultur und Kulturwissenschaft* [Culture and culture studies]. Paderborn, Germany: UTB.

Hattie, J. A., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81–112. doi:10.3102/003465430298487

Hoffman, B. J., Woehr, D. J., Maldagen-Youngjohn, R., & Lyons, B. D. (2011). Great man or great myth? A quantitative review of the relationship between individual differences and leader effectiveness. *Journal of Occupational and Organizational Psychology, 84*, 347–381. doi:10.1348/096317909x485207

Hofstede, G. (2001). Culture's consequences—comparing values, behaviors, institutions and organizations across nations (2nd ed.). Thousand Oaks, CA: Sage.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70. Retrieved from http://www.jstor.org/stable/4615733

Hooper, A. C., Cullen, M. J., & Sackett, P. R. (2006). Operational threats to the use of SJTs: Faking, coaching, and retesting issues. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgement tests: Theory, measurement and application* (pp. 205–232). London, United Kingdom: Lawrence Erlbaum Associates.

Hulstrand, J. (2008). Supply and demand: Building a fluent workforce. *International Educator, 17*(5), 24–31. Retrieved from http://www.nafsa.org/_/File/_/ie_sepoct08_fluentwork.pdf

Ilgen, D. R., Fisher, C. D., Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*, 349–371. doi:10.1037/0021-9010.64.4.349

Institute for the Future for University of Phoenix Research Institute (2011). *Future work skills 2020*. Retrieved from http://www.iftf.org/uploads/media/SR1382A_UPRI_future_work_skills_sm.pdf

Jacobson, W., Schleicher, D., & Maureen, B. (1999). Portfolio assessment of intercultural competence. *International Journal of Intercultural Relations, 23*, 467–492. doi:10.1016/S0147-1767(99)00006-1

Javidan, M., & Teagarden, M. B. (2011). Conceptualizing and measuring global mindset. In W. H. Mobley, M. Li, & Y. Wang (Eds.), *Advances in global leadership* (Vol. 6, pp. 13–39). Bingley, United Kingdom: Emerald.

Johnson, J. P., Lenartowicz, T., & Apud, S. (2006). Cross-cultural competence in international business: Toward a definition and a model. *Journal of International Business Studies, 37*, 525–543. doi:10.1057/palgrave.jibs.8400205

Kanning, U. K. (2002). Soziale Kompetenz—Definition, Strukturen und Prozesse [Social competence—definition, structures, and processes]. *Zeitschrift für Psychologie*, *210*, 154–163. doi:10.1026//0044-3409.210.4.154

Kelley, C., & Meyers, J. (1995). *The Cross-Cultural Adaptability Inventory (manual).* Minneapolis, MN: National Computer Systems.

Kim Y. Y. (1988). *Communication and cross-cultural adaptation: An integrative theory*. Philadelphia, PA: Multilingual Matters.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254–284. doi:10.1037/0033-2909.119.2.254

Koester, J., & Olebe, M. (1988). The behavioral assessment scale for intercultural communication effectiveness. *International Journal of Intercultural Relations*, *12*, 233–246. doi:10.1016/0147-1767(88)90017-x

Kong, W. (2014). *跨文化能力测试的文化特殊性.中德受试者的测试处理过程比较研究* [Culture-specificity of the Test to Measure Intercultural Competence. A cross-cultural study concerning the response processes of Chinese and German test participants] (Unpublished master's thesis). Beijing Foreign Studies University, Beijing, China.

LaFromboise, T., Coleman, H. L., Gerton, J. (1993). Psychological impact of biculturalism: Evidence and theory. *Psychological Bulletin, 114*, 395–412. doi:10.1037//0033-2909.114.3.395

Lamnek, S. (1995). *Qualitative Sozialforschung* [Qualitative social science]. Weinheim, Germany: Beltz.

Leone, L., Van der Zee, K. I., Van Oudenhoven, J. P., Perugini, M., & Ercolani, A. P. (2005). The cross-cultural generalizability and validity of the Multicultural Personality Questionnaire. *Personality and Individual Differences, 38*, 1449–1462. doi:10.1016/j.paid.2004.09.010

Leung, K., Ang, S., & Tan, M. L. (2014). Intercultural competence. *Annual Review of Organizational Psychology and Organizational Behavior, 1*, 489–519. doi:10.1146/annurev-orgpsych-031413-091229

Leung, K., & Cheng, G. H. (2014). Adaptability and intercultural interaction in the work context: A cultural tuning perspective. In D. Chan (Ed.), *Individual adaptability to changes at work: New directions in research* (pp. 156–174). New York, NY: Taylor & Francis.

Lievens, F., Harris, M. M., Van Keer, E., & Bisqueret C. (2003). Predicting intercultural training performance: The validity of personality, cognitive ability, and dimensions measured by an assessment center and a behavior description interview. *Journal of Applied Psychology, 88*, 476–488. doi:10.1037/0021-9010.88.3.476

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*, *37*, 426–441. doi:10.1108/00483480810877598

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *140,* 1–55. Retrieved from http://www.voteview.com/Likert_1932.pdf

Lloyd, S., & Härtel, C. (2010). Intercultural competencies for culturally diverse work teams. *Journal of Managerial Psychology*, *25*, 845–875. doi:10.1108/02683941011089125

Locke, E. A., & Latham, G. P. (Eds.). (1984). *Goal setting: A motivational technique that works!* Englewood Cliffs, NJ: Prentice Hall.

London, M., & McFarland, L. A. (2010). Assessment feedback. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of employee selection* (pp. 417–436). New York, NY: Routledge.

Lustig, M. W. (2005). Toward a well-functioning intercultural nation. *Western Journal of Communication, 69*, 377–379. doi:10.1080/10570310500305612

Olson, C. L., & Kroeger, K. R. (2001). Global competency and intercultural sensitivity. *Journal of Studies in International Education, 5*, 116–137. doi:10.1177/102831530152003

Martin, J. N. (1993). Intercultural communication competence: A review. In R. L. Wiseman & J. Koester (Eds.), *Intercultural communication competence* (pp. 16–29). Newbury Park, CA: Sage.

Matsumoto, D., & Hwang, H. C. (2013). Assessing cross-cultural competence. A review of available tests. *Journal of Cross-Cultural Psychology, 44*, 849–873. doi: 10.1177/0022022113492891

Matsumoto, D., LeRoux, J. A., Ratzlaff, C., Tatani, H., Uchida, H., Kim, C., & Araki, S. (2001). Development and validation of a measure of intercultural adjustment potential in Japanese sojourners: The Intercultural Adjustment Potential Scale (ICPS). *International Journal of Intercultural Relations, 25*, 483–510. doi:10.1016/S0147-1767(01)00019-0

McDaniel, M. A., Hartman, N. S., & Grubb, W. L. (2003, April). *Situational judgment tests, knowledge, behavioral tendency, and validity: A meta-analysis.* Paper presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2006). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, *60*, 63–91. doi:10.1111/j.1744-6570.2007.00065.x

McDaniel, M. A., Morgeson, R. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, *86*, 730–740. doi:10.1037//0021-9010.86.4.730

McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment test: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, *9*, 103–113. doi:10.1111/1468-2389.00167

McDaniel, M. A., & Whetzel, D. L. (2005, June). *Situational Judgment Tests.* Paper presented at the 29th Annual IPMAAC Conference on Personnel Assessment, Orlando, FL.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543. doi:10.1007/BF02294825

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing. Preparing people for change*. New York, NY: Guilford Press.

Moin, A. K. (1986). *Relative effectiveness of various techniques of calculus instruction: A meta-analysis* (Unpublished doctoral dissertation). University of Syracuse, Syracuse, NY.

Moosmüller, A. (2007). Lebenswelten von „Expatriates“ [Lifeworlds of so-called “Expatriates”]. In J. Straub, A. Weidemann, & D. Weidemann (Eds.), *Handbuch interkulturelle Kommunikation und Kompetenz* [Handbook of intercultural communication and competence] (pp. 480–488). Stuttgart, Germany: J. B. Metzler.

Mplus (Version 7.1 for Windows) [Computer software]. Los Angeles, CA: Muthén & Muthén.

Muritiba, P. M., Muritiba, S. N., Campanário, M., & de Albuquerque, L. G. (2010). International HR strategy in Brazilian technology multinationals. *BAR—Brazilian Administration Review*, *7*(4), 325–344. doi:10.1590/S1807-76922010000400002

Müller-Neumann, C. (2005). Assessment-Center interkulturell [Assessment center intercultural]. In V. Fischer, M. Springer, & I. Zacharaki (Eds.), *Interkulturelle Kompetenz. Fortbildung—Transfer—Organisationsentwicklung* [Intercultural competence. Further develoment—transfer—organizational development] (pp. 159–179). Schwalbach, Germany: Wochenschau-Verlag.

Navas, M., García, M. C., Sánchez, J., Rojas, A. J., Pumares, P., & Fernández, J. S. (2005). Relative acculturation extended model (RAEM): New contributions with regard to the study of acculturation. *International Journal of Intercultural Relations, 29*, 21–37. doi:10.1016/j.ijintrel.2005.04.001

Newman, M. L., & Greenway, P. (1997). Therapeutic effects of providing MMPI-2 test feedback to clients at a university counseling service. *Psychological Assessment, 9,* 122–131. doi:10.1037/1040-3590.9.2.122

Nguyen, N., McDaniel, M. A., & Whetzel, D. L. (2005, April). *Subgroup differences in situational judgment test performance: A meta-analysis*. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

O'Neill, J. (2001). *Building better global economic BRICs* (Global Economics Paper No. 66). Retrieved from the Goldman Sachs website: http://www.goldmansachs.com/our-thinking/archive/archive-pdfs/build-better-brics.pdf

Paige, R. M., Jacobs-Cassuto, M., Yershova, Y. A., & DeJaeghere, J. (2003). Assessing intercultural sensitivity: An empirical analysis of the Hammer and Bennett Intercultural Development Inventory. *International Journal of Intercultural Relations*, *27*, 467–486. doi:10.1016/S0147-1767(03)00034-8

Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment, 22*, 203–212. doi:10.1037/a0018679

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change. Applications to addictive behaviors. *American Psychologist, 47*, 1102–1114. doi:10.1037/0003-066X.47.9.1102

Pruegger, V. J., & Rogers, T. B. (1993). Development of a scale to measure cross-cultural sensitivity in the Canadian context. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *25*, 615–621. doi:10.1037/h0078846

Pruegger, V. J., & Rogers, T. B. (1994). Cross-cultural sensitivity training: Methods and assessments. *International Journal of Intercultural Relations*, *18*, 369–387. doi:10.1016/0147-1767(94)90038-8

Rathje, S. (2006). Interkulturelle Kompetenz—Zustand und Zukunft eines umstrittenen Konzepts [Intercultural competence—Status and future of a controversial concept]. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, *11*(3), 1–15. Retrieved from http://zif.spz.tu-darmstadt.de/jg-11-3/docs/Rathje.pdf

Rathje, S. (2007). Intercultural competence: The status and future of a controversial concept. *Language and Intercultural Communication, 7,* 254–266. doi:10.2167/laic285.0

Rockstuhl, T., Ang, S., Ng, K. Y., Lievens, F., Van Dyne, L. (2013a, April). *The incremental value of assessing situational perspective taking in SJTs.* Paper presented at the 28th Annual Conference of the Society for Industrial and Organizational Psychology, Houston, TX.

Rockstuhl, T., Ang, S., Ng, K. Y., Lievens, F., & Van Dyne, L. (in press). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology*. Retrieved from http://soonang.com/wp-content/uploads/2014/09/JAP-Putting-Judging-Situations-into-SJTs_accepted.pdf

Rockstuhl T., Presbitero, A., Ng, K. Y., Ang, S. (2013b, June). *Metacognitive cultural intelligence and offshoring performance: Predictive validity of an intercultural situational judgment test (iSJT).* Paper presented at the Third International Conference on Outsourcing of Information Services, Mannheim, Germany.

Ruben, B. D. (1976). Assessing communication competency for intercultural adaptation. *Group & Organization Studies, 1*, 334–354. doi:10.1177/105960117600100308

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144. doi:10.1007/BF00117714

Sarason, R. B. (1981). *The dimensions of social competence. Contributions from a variety of research areas.* In J. D. Wine & M. D. Smye (Eds.), *Social competence* (pp. 100–122). New York, NY: Guilford Press.

Schnabel, D., Kelava, A., Van de Vijver, F. J. R., & Seifert, L. (2014). *Examining psychometric properties, measurement invariance, and construct validity of a short version of the Test to Measure Intercultural Competence (TMIC-S) in Germany and Brazil.* Manuscript submitted for publication.

Schnabel, D., Kelava, A., Seifert, L., & Kuhlbrodt, B. (2015). Konstruktion und Validierung eines multimethodalen berufsbezogenen Tests zur Messung interkultureller Kompetenz [Development and validation of a job-related multimethod Test to Measure Intercultural Competence]. *Diagnostica*, *61*, 3–21. doi:10.1026/0012-1924/a000110

Schnabel, D., Kelava, A., & Van de Vijver, F. J. R. (in press). The effects of using collaborative assessment with students going abroad: Intercultural competence development, self-understanding, self-confidence, and stages of change. *Journal of College Student Development.*

Schönhuth, M. (2005). *Glossar Kultur und Entwicklung: Ein Vademecum durch den Kulturdschungel* [Glossary culture and development: A vade mecum through the culture jungle]. Trier, Germany: Trierer Materialien zur Ethnologie.

Schroeder, D. G., Hahn, E. D., Finn, S. E., & Swann, W. B., Jr., (1993, June). *Personality feedback has more impact when mildly discrepant form self views.* Paper presented at the Fifth Annual Convention of the American Psychological Society, Chicago, IL.

Schuler, H., & Prochaska, M. (2000). Entwicklung und Konstruktvalidierung eines berufsbezogenen Leistungsmotivationstests [Development and validation of an occupational achievement motivation inventory]. *Diagnostica, 46*, 61–72. doi:10.1026//0012-1924.46.2.61

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, *25*, 167–177. doi:10.1016/S0191-8869(98)00001-4

Sebald, J. A. P. (2008). *Effektivität und Angemessenheit—Messung Interkultureller Kompetenz im Assessment Center* [Effectivity and adequacy—Measuring intercultural competence in an assessment center] (Unpublished dissertation), Friedrich-Schiller-Universität, Jena, Germany.

Sheahan, P. (2005). *Generation Y: Thriving and surviving with Generation Y at work.* Prahran, Australia: Hardie Grant Books.

Sinicrope, C., Norris, J., & Watanabe, Y. (2007). Understanding and assessing intercultural competence: A summary of theory, research, and practice. *Second Language Studies, 26*, 1–58. Retrieved from http://www.hawaii.edu/sls/uhwpesl/26(1)/Norris.pdf

Spencer, L. M., & Spencer, S. M. (1993). *Competence at work: Models for superior performance.* New York, NY: Wiley.

Spitzberg, B. H., & Changnon, G. (2009). Conceptualizing intercultural competence. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 2–52). Los Angeles, CA: Sage Publications.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*, 240–261. doi:10.1037//0033-2909.124.2.240

Statistisches Bundesamt (2013). *Bevölkerung und Erwerbstätigkeit* [Population and labor]. Retrieved from https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegration/MigrationshintergrundErweitert2010220127004.pdf?__blob=publicationFile

Stephen, W. G., & Stephen, C. W. (1985). Intergroup anxiety. *Journal of Social Issues, 41*, 157–176. doi:10.1111/j.1540-4560.1985.tb01134.x

Sternberg, R. J. (2005). WICS: A model of leadership. *The Psychologist-Manager Journal*, *8*, 29–43. doi:10.1207/s15503461tpmj0801_4

Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition.* Norwood, NJ: Ablex.

Stogdill, R. M. (1948). Personal factors associated with leadership: A survey of the literature. *Journal of Psychology, 25*, 35–71. doi:10.1080/00223980.1948.9917362

Straffon, D. A. (2003). Assessing the intercultural sensitivity of high school students attending an international school. *International Journal of Intercultural Relations*, *27*, 487–501. doi:10.1016/S0147-1767(03)00035-X

The Economist Intelligence Unit Limited (2010). *Up or out: Next moves for the modern expatriate*. London, United Kingdom: The Economist Intelligence Unit Limited.

Thomas, A. (2003). Interkulturelle Kompetenz—Grundlagen, Probleme und Konzepte [Intercultural Competence—Basics, problems, and concepts]. *Erwägen, Wissen, Ethik, 14,* 137–221. Retrieved from http://groups.uni-paderborn.de/ewe/index.php

Thomas, A., Kammhuber, S., & Layes, G. (1997). Interkulturelle Kompetenz. Ein Handbuch für internationale Einsätze der Bundeswehr [Intercultural competence. A manual for international missions of the German Federal Armed Forces]. *Untersuchungen des Psychologischen Dienstes der Bundeswehr*, *32*, 7–190.

Ting-Toomey, S. (1993). Communicative resourcefulness: An identity negotiation perspective. In R. L. Wiseman & J. Koester (Eds.), *Intercultural communication theory* (pp. 72–111). Newbury Park, CA: Sage.

Triandis, H. C. (1977). *Interpersonal behavior*. Monterey, CA: Brooks/Cole.

Van de Vijver, F. J. R., & Leung, K. (1997). Methods and data analysis of comparative research. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), *Handbook of cross-cultural psychology* (2nd ed., pp. 257–300). Boston, MA: Allyn & Bacon.

Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue européenne de psychologie appliquée, 54,* 119–135. doi:10.1016/j.erap.2003.12.004

Van Der Zee, K. I., & Van Oudenhoven, J. P. (2000). The Multicultural Personality Questionnaire: A multidimensional instrument of multicultural effectiveness. *European Journal of Personality*, *14*, 291–309. doi:10.1002/1099-0984(200007/08)14:4<291::AID-PER377>3.0.CO;2-6

Van der Zee, K. I., & Van Oudenhoven, J. P. (2001). The Multicultural Personality Questionnaire: Reliability and validity of self- and other ratings of multicultural effectiveness. *Journal of Research in Personality, 35*, 278–288. doi:10.1006/jrpe.2001.2320.

Van der Zee, K. I., Zaal, J. N., & Piekstra, J. (2003). Validation of the Multicultural Personality Questionnaire in the context of personnel selection. *European Journal of Personality*, *17*, 77–100. doi:10.1002/per.483

Van Dyne, L., Ang, S., & Koh, C. (2008). Development and Validation of the CQS. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence: Theory, measurement and application* (pp. 16–38). Armonk, NY: M. E. Sharpe.

Van Oudenhoven, J. P., Mol, S., & Van der Zee, K. I. (2003). Study of the adjustment of Western expatriates in Taiwan ROC with the Multicultural Personality Questionnaire. *Asian Journal of Social Psychology*, *6*, 159–170. doi:10.1111/1467-839X.t01-1-00018

Van Oudenhoven, J. P., & Van der Zee, K. I. (2002). Predicting multicultural effectiveness of international students: The Multicultural Personality Questionnaire. *International Journal of Intercultural Relations, 26,* 679–694. doi:10.1016/S0147-1767(02)00041-X

Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies, 29*, 1024–1054. doi:10.1080/01419870701599465

Wang, Y.-W., Davidson, M., Yakushko, O. F., Savoy, H. B., Tan, J. A., & Bleier, J. K. (2003). The Scale of Ethnocultural Empathy: Development, validation, and reliability. *Journal of Counseling Psychology, 50,* 221–234. doi:10.1037/0022-0167.50.2.221

Weekly, J. A., & Ployhart, R. E. (2006). *Situational Judgment Tests: Theory, measurement and application.* London, United Kingdom: Lawrence Erlbaum Associates.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen—eine umstrittene Selbstverständlichkeit [Comparative performance tests in schools—a controversial matter of course]. In F. E. Weinert (Ed.), *Leistungsmessungen in Schulen* [Performance assessment in schools] (pp. 17–31). Weinheim, Germany: Beltz.

Whaley, A. L., & Davis, K. E. (2007). Cultural competence and evidence-based practice in mental health services: A complementary perspective. *American Psychologist, 62*, 563–574. doi:10.1037/0003-066X.62.6.563

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, *19*, 188–202. doi:10.1016/j.hrmr.2009.03.007

Wierlacher, A. (2003). Das tragfähige Zwischen [The stable in between]. *Erwägen, Wissen, Ethik, 14,* 215–217. Retrieved from http://groups.uni-paderborn.de/ewe/index.php?id=79

Appendix A

Interview Guideline and Summary of Statements from the Exploratory Expert Interviews

Topic	Corresponding answers
Introduction	
• Dank für die Teilnahme • Zweck des Interviews: Entwicklung eines neuen Potenzialanalysetools • Vorgehen • Lockeres Gespräch, kein Wissenstest • Aufforderung, Input über die Fragen hinweg zu geben • Erlaubnis erbitten, Gespräch aufzuzeichnen—Anonymität	-
Content-related questions: Intercultural development	
Was beutet für Sie interkulturelle Entwicklung (IE) einer Person?	• IE entsteht durch Erfahrung, Trainings, Wissensvermittlung?! • IE ist ein sehr langer Prozess, in dem man lernt, das Gelernte auch einzusetzen • IE ist Teil der Persönlichkeitsentwicklung • IE heißt, mit verschiedenen Situationen umzugehen, Signale von Interaktionspartnern wahrzunehmen, verschiedene Deutungsmöglichkeiten und Reaktionsweisen auszuprobieren, den eigenen Ansatz zu hinterfragen und zu relativieren und neue Lösungsmöglichkeiten zu erlernen • IE als Fähigkeit, mit Unterschiedlichkeit umgehen zu können • Entwicklung ist, wenn ein Teilnehmer von einer Stufe auf die andere kommt. Eine hohe Stufe ist, wenn die Leute akzeptieren können, dass es Unterschiedlichkeit gibt. Die höchste Stufe ist Synergien schaffen • Man entwickelt ein bestimmtes Handlungsrepertoire. Wenn man sich in seiner eigenen Kultur befindet, hat man bestimmte Konzepte, denen man sich nicht so bewusst ist und wenn man dann in eine andere Kultur kommt, versteht man zunächst die eigenen Werte und man lernt etwas dazu • 1. Schritt ist Selbsterkenntnis, 2. Schritt ist das Wissen über das Andere, die Wertschätzung, dass man Dinge auch anders sehen kann, die Relativität zur eigenen Position, sich Mühe machen zu verstehen, warum jemand aus einer anderen Kultur an einen Sachverhalt, an ein Problem oder an eine Situation anders heran geht, andere Prioritäten, Gedanken und Handlungsweisen hat. Objektive Distanz zum eigenen Verhalten. 3. Schritt ist die Entwicklung von Handlungsstrategien, wie man mit dem Anderen umgehen kann • Akzeptanz als Anfang jeder kultureller Entwicklung
Wie zeichnet sich eine IE bei Ihren Trainingsteilnehmern ab?	• Durch erworbenes Wissen • Durch Aha-Effekt • Durch Einsichten, die Teilnehmer gewinnen • Dadurch, dass Teilnehmer die Perspektiven der Anderen einnehmen • Wenn Probleme gut analysiert werden können • Durch einen gewissen Denkprozess • Wenn Leute auch andere Möglichkeiten sehen • Eher gar nicht • Durch ein gesteigertes Bewusstsein • Dadurch, dass ein Teilnehmer sich öffnet für die Veränderungssituation, dass er sagt: „So, ich muss jetzt hier etwas ändern, um effektiver zu sein“. Das Ausprobieren neuer Verhaltensweisen, besseres Zuhören, häufigeres Fragenstellen, Ausprobieren anderer Verhaltensweisen
Content-related questions: Definition of intercultural competence	
Wie würden Sie, aus Ihrer beruflichen Erfahrung heraus, interkulturelle Kompetenz definieren?	• All das, was uns dazu befähigt mit Menschen aus anderen Kulturen adäquat umzugehen oder auch deren Kulturen zu verstehen. Verständnis, eine gewisse Offenheit, Respekt den anderen Kulturen gegenüber, aber auch dass ich mich auskenne mit gewissen Dingen • Interkulturelle Kompetenz ist, dass man das, was man in der Entwicklung erlernt, dann auch verfügbar hat • Interkulturelle Kompetenz bedeutet seine Kultur mit der fremden Kultur zu

verbinden und sich situationsadäquat zu verhalten, d.h. sich je nach Kultur flexibel anders zu verhalten. Dafür ist es wichtig, sich darüber bewusst zu sein, in welcher Situation was akzeptabel ist
- Interkulturelle Kompetenz besteht darin, dass man die Handlungskompetenz (Bolten) flexibel im interkulturellen Kontext einsetzen kann
- Interkulturelle Kompetenz bedeutet, sich Ursache-Wirkungs-Zusammenhänge zwischen Verhalten, Wahrnehmen seitens der anderen Person, Interpretation und Bewertung klarmachen zu können
- Eine Person besitzt eine hohe interkulturelle Kompetenz, wenn sie sehr gut spüren kann, was die Umgebung ihr erzählt. Dazu gehört, eine gewisse Wissensbasis, eine Sensibilität für das Umfeld und für das eigene Verhalten sowie Empathie

Content-related questions: Intercultural competence facets

Welche Kompetenzen sind im Rahmen der interkulturellen Zusammenarbeit im Generellen in Ihren Augen wichtig?

Welche Kompetenzen werden benötigt, wenn ein Mitarbeiter in einen anderen, von ihm nicht präferierten, Kulturraum geht?

Kommunikation:
- Verstehen von verschiedenen Kommunikationsstilen
- Den eigenen Kommunikationsstil variabel halten können
- Zwischen verschiedenen Stilen switchen können
- Wahrnehmen können, wann welcher Kommunikationsstil richtig ist
- Fähigkeit, die Wirkung der eigenen Kommunikation auf andere einzuschätzen
- Zuhören
- Botschaften so vermitteln, dass sie ankommen
- Fähigkeit, das kommunikative Ziel zu erreichen, indem man verschiedene Mittel einsetzt
- Zwischen den Zeilen lesen können. Nicht nur darauf achten, was gesagt wird, sondern auch wie es gesagt wird
- Deuten können, wie Dinge gemeint sind
- Sensibilität dafür, wo man Gefühle verletzen kann
- Achtsamkeit im Umgang mit sich und anderen

Zielgerichtetheit:
- Fähigkeit, Ziele zu erreichen
- Fähigkeit, das Wissen über eine andere Kultur zur Erreichung der eigenen Ziele einzusetzen
- Fähigkeit, eine Richtung zu halten und Ziele zu erreichen, auch wenn es turbulent ist, ohne die eigenen Strukturen durchzudrücken, was zu Widerständen führen kann, sondern zu versuchen, neue Wege zu finden, abhängig von den Mitteln, die in der jeweiligen Kultur zur Verfügung stehen

Produktive Zusammenarbeit ermöglichen:
- Fähigkeit das Anderssein einzusetzen und dadurch Mehrwerte in internationalen Teams zu schaffen
- Ergänzungspotenziale in Teams mit unterschiedlichen Fähigkeiten erkennen können
- Fähigkeit, Synergien zu schaffen
- Akzeptanz verschiedener Lösungsansätze und Zielerreichungsstrategien

Modellwissen:
- Modellwissen darüber, welche Werte hinter bestimmten Reaktionen stecken
- Das Denken und Handeln der Leute verstehen können

Offenheit:
- Neugierde: wie handeln, denken, interagieren andere Menschen
- Offenheit für andere Menschen, Meinungen, Ideen
- Interesse an anderen Menschen

Toleranz:
- Wertschätzende Grundhaltung
- Offene Grundhaltung (weg vom Ethnozentrismus)
- Wahrnehmung und Wertschätzung ohne Bewertung
- Vorurteilsfreiheit

Selbstmanagement:
- Selbststeuerung
- Fürsorge für sich und andere
- Umgang mit spontanen Impulsen
- Umgang mit Emotionen

	• Selbstbeherrschung • Stressmanagement • Kriesenmanagement, Fähigkeit mit Belastung und Kulturschock umzugehen • Stressresistenz • Reflektion eigener Erfahrungen und Emotionen zur Zielerreichung Erwartungsmanagement: • Bewusstsein für unterschiedliche Erwartungen haben • Erkennen, wo Erwartungshaltungen bei dem anderen bestehen • Konfliktmanagement • Vermittler, Mediator sein zwischen Stammhaus und lokaler Organisation • Sandwich Position, zwischen den Stühlen sitzen, Interessenkonflikte zwischen Heimatunternehmen und lokaler Organisation managen können Kompetenter Umgang mit Diversity: • Mit Diversity ohne Vorverurteilung und wertschätzend umgehen • Diversity als Wettbewerbsvorteil, als Plus, als Asset verstehen Flexibilität: • Bewusst situationsadäquat agieren können • Die Fähigkeit, das eigene Verhalten flexibel zu verändern • Veränderungsbereitschaft • Anpassungsfähigkeit Fähigkeit zur Selbst- und Fremdwahrnehmung: • Kenntnis unterschiedlicher kultureller Normen und Werte • Verstehen, wie die eigene Nationalität im Ausland gesehen wird • Eigene Präferenzen kennen und das Bewusstsein und die Toleranz für andere Präferenzen haben • Fähigkeit, eine gesunde Distanz zum eigenen Verhalten zu haben • Bewusstsein, wie man selbst rüberkommt • Bewusstsein über die eigene Identität, die eigenen Normen und Werte • Bewusstsein der Unterschiedlichkeiten • Fähigkeit, zwischen verschiedenen kulturellen Identitäten zu wechseln Fremdsprachenkompetenz: • Interesse an der Fremdsprache • Bereitschaft, sie zu erlernen • Bereitschaft, Zeit zu investieren • Keine Angst zu kommunizieren und Fehler zu machen • Fähigkeit, mit dem Sprachniveau, das man hat, Ziele zu erreichen Soziale Interaktion: • Fähigkeit, Kontakte zu knüpfen und aufrechtzuerhalten • Fähigkeit, andere für sich zu gewinnen und Sympathie zu schaffen • Fähigkeit, Vertrauen aufzubauen und anderen Menschen Wärme und Aufmerksamkeit zu geben • Fähigkeit, sich Netzwerke außerhalb der Organisation aufzubauen • Fähigkeit, sich in neuen Organisationen zurechtzufinden, zu verstehen, wer der informelle Entscheider ist Führungskompetenzen: • Fähigkeit, andere Menschen zu motivieren • Delegieren • Feedback geben • Führungsstil als Mischung aus Partizipation und Autorität • Fähigkeit ein Best practice im Team zu finden Lernbereitschaft und Lerntechniken
Content-related questions: The potential for intercultural competence facets to be trained	
Welche Kompetenzen werden in Ihren interkulturellen Trainings am häufigsten entwickelt? Was (Sachverhalte/Kompetenzen/Skills) lässt sich, Ihrer Meinung, am Besten/am Schlechtesten in den Trainings entwickeln?	Am besten zu entwickelnde interkulturelle Kompetenzen • Kommunikation • Aktives Zuhören • Reflektion über Wahrnehmungsmuster: Wie nehme ich mich wahr? Wie nehme ich andere wahr? Wie beeinflusst das mein Handeln? • Scharfsinn • Aufmerksamkeit

	• Bewusstsein, dass Empathie bei der interkulturellen Zusammenarbeit wichtig ist • Handlungsflexibilität • Das Bewusstsein, dass es solche Kompetenzen gibt, vor allem im Kommunikationsumfeld • Das Bewusstsein für Diversität • Bewusstmachen individueller Hygiene-Faktoren und Motivatoren
	Am schlechtesten zu entwickelnde interkulturelle Kompetenzen
	• Stabile Kompetenzen • Empathie • Durch Erziehung und Erfahrung (nicht) entwickelte wertschätzende Grundhaltung, Aufmerksamkeit, Offenheit, Achtsamkeit • Umgang mit Zeit • Vorteile der Beziehungsorientierung • Resilienz • Positive Einstellung zum Unbekannten • Neugier an anderen Personen
Content-related questions: Requirements concerning an intercultural competence instrument	
Was fehlt Ihnen inhaltlich an den bisher von Ihnen einsetzten Tests, das für Ihre interkulturellen Trainingszwecke hilfreich wäre?	• Skill-basiert • Items als Situationen und Verhalten • Kompetenzen sollten situationsabhängig abgefragt werden • Fokus auf Kompetenzen, die sich an Beispielsituationen und Verhaltensweisen festmachen lassen
Was fehlt Ihnen methodisch an den bisher von Ihnen einsetzten Tests?	• Möglichkeit abzulesen, auf welcher Entwicklungsstufe ein Teilnehmer bezüglich einer spezifischen Kompetenz ist • Die Entwicklungsstufen sollten sich nicht auf interkulturelle Kompetenz im Allgemeinen beziehen, sondern auf konkrete Entwicklungsbereiche
Was erwarten Sie von einem neuen Potenzialanalysetool? (Attribute)	• Dass man am Tool die Kompetenzentwicklung ablesen kann • Entscheidungsfragen anstatt Selbsteinschätzungen
Welche spezifischen Anforderungen stellen Sie an ein Potenzialanalysetool, damit es in Ihrem interkulturellen Training hilfreich ist?	• Es soll zeigen können, welche Mitarbeiter geeignet sind, in das Ausland zu gehen • Es müsste zwei Messungen/Abfragen geben—direkt nach dem Training und nach vier, acht oder 12 Wochen—Lernerfolg setzt auch Umsetzung voraus—nur Wissenserweiterung alleine sagt noch nichts über Anwendbarkeit aus • Konkrete Angaben zum Datenschutz in der Einleitung • Größeres Bewusstsein bei den Teilnehmern schaffen, wie wichtig es ist, den Test spontan auszufüllen • Dass sich die Teilnehmer sicher fühlen • Die Sprache des Tests soll die Sprache des Managements sein, also nicht zu psychologisch, nicht zu universitär • Dass das bipolare Kulturkonzept aufgebrochen wird • Reliabilität und Validität • Kulturübergreifender Test • Geringe Tendenzen zur sozialen Erwünschtheit • Es sollte abgefragt werden, wo Menschen bereit sind, sich zu ändern, wo gar keine Bereitschaft besteht und wo Teilnehmer sich flexibel fühlen • Push-Kompetenzen: Focus on goals, inner purpose, clarity of communication, exposing intentions (Fähigkeit, gut rüber zu bringen, warum etwas gemacht werden soll) • Balance zwischen Push- und Pull-Kompetenzen
Closing part	
Open questions, thanks	-

Appendix B

Pre-print Version of Paper A

Appendix B provides a pre-print version of Paper A: Schnabel, D., Kelava, A., Seifert, L., & Kuhlbrodt, B. (2015). Konstruktion und Validierung eines multimethodalen berufsbezogenen Tests zur Messung interkultureller Kompetenz [Development and validation of a job-related multimethod Test to Measure Intercultural Competence]. *Diagnostica*, *61*, 3–21. doi:10.1026/0012-1924/a000110

Pre-print refers to the originally in 2012 submitted version of the paper before any peer-review. Please note, that there are consequently substantial differences between the pre-print and the published versions.

Titel in deutscher Sprache

Konstruktion und Validierung eines multimethodalen berufsbezogenen Tests zur Messung interkultureller Kompetenz

Zusammenfassung in deutscher Sprache

Interkulturelle Kompetenz wird als globale Verhaltensorientierung mit mehrdimensionaler Struktur verstanden. Auf dieser Basis wurde der Test zur Messung Interkultureller Kompetenz (TMIK) entwickelt, der anhand von Selbsteinschätzungs- und Situationsbeurteilungsitems die multimethodale sowie differenzierte Messung interkultureller Kompetenz erstmalig ermöglicht. An $N = 641$ Personen wurden die Messeigenschaften der Skala sowie die Ausprägung interkultureller Kompetenz in Abhängigkeit von vier Außenkriterien überprüft. Mithilfe des ESEM-Verfahrens (Exploratory Structural Equation Modeling; Asparouhov & Muthén, 2009) konnte eine 17-faktorielle Struktur bestätigt werden. Die Ergebnisse zeigen, dass das Instrument gute psychometrische Eigenschaften besitzt. Die Gesamtskala wies hypothesenkonforme Beziehungen mit den Außenkriterien auf, was die Kriteriumsvalidität bestätigt. Für 16 Faktoren konnten wie erwartet konvergente Zusammenhänge mit Items eines konstruierten Situational Judgement Tests festgestellt werden, was die Konstruktvalidität stützt. Wie vorhergesagt erzielten Personen, die schon einmal an einem interkulturellen Training teilgenommen haben, eine längere Zeit im Ausland verbracht haben oder interkulturell involviert waren, höhere Mittelwerte im Großteil der Faktoren. Mögliche Einsatzbereiche sind die Personaldiagnostik und -entwicklung sowie die interkulturelle Kompetenzforschung.

Schlüsselwörter in deutscher Sprache

Interkulturelle Kompetenz, Personaldiagnostik, Situationsbeurteilungstest, multimethodal

1. Text

Durch die Internationalisierung von Arbeitsprozessen sowie durch die einschlägigen gesellschaftspolitischen Veränderungen, gewinnt interkulturelle Kompetenz über Lebensbereiche hinweg an Bedeutung (Erpenbeck, 2012; Sinicrope, Norris, & Watanabe, 2007; Thomas, 2006). Im Zuge dessen, wurden multiple Definitionen als auch einige empirische Überprüfungen interkultureller Kompetenz vorgenommen (z. B. Bennett, 1993; Bolten, 2007b; Fantini & Tirmizi, 2006; Ruben, 1989; Thomas, 2006). Trotz eines hohen Heterogenitätsgrades (Rathje, 2006), besteht die wesentliche Überschneidung unterschiedlicher Auffassungen darin, dass interkulturelle Kompetenz ein mehrdimensionales Konstrukt ist und eine entscheidende Rolle im Umgang mit Personen aus anderen Kulturen hat (z. B. Fantini & Tirmizi, 2006; Hammer, Bennett & Wiseman, 2003; Kealey & Ruben, 1983). Der vorliegende Artikel gibt zunächst einen Überblick über die theoretischen Ansätze sowie bestehenden Messverfahren interkultureller Kompetenz und stellt eine eigene Definition interkultureller Kompetenz vor. Darauf aufbauend wird die Entwicklung und Validierung eines multimethodalen Tests zur Messung interkultureller Kompetenz (TMIK) beschrieben.

Interkulturelle Kompetenz

Kompetenzen befähigen ein Individuum dazu, in unerwarteten und unbekannten Situationen, erfolgreich neuartige Probleme zu lösen (Erpenbeck & von Rosenstiel, 2007; Weinert, 2001). Sie werden als vom Individuum erlernbare persönlichkeitsrelevante Faktoren verstanden (Erpenbeck, 2012) und können in Fach- und Methodenkompetenzen, sozial-kommunikative Kompetenzen und personale Kompetenzen unterteilt werden. Dem übergeordnet ist die Handlungskompetenz (Erpenbeck & von Rosenstiel, 2007; Kulturministerkonferenz, 2011).

Bolten (2007a) differenziert zur Definition interkultureller Kompetenz zwischen Listen-, Struktur- und Prozessmodellen. Bei Listenmodellen werden verschiedene Merkmale interkultureller Kompetenz gesammelt (Brislin, 1981; Ruben, 1976). In Strukturmodellen sind die Merkmale interkultureller Kompetenz affektiven, kognitiven und konativen Kategorien zugeordnet (Dauner, 2011; Gertsen, 1990; Ting-Toomey, 1993). Mehrere Autoren nehmen an, dass *eine* interkulturelle Kompetenz *verschiedene* Fähigkeiten umfasst, die benötigt werden, um angemessen und effektiv mit Personen aus anderen Kulturen zu

interagieren (z. B. Bergemann & Bergemann, 2005; Fantini & Tirmizi, 2006; Hammer et al., 2003; Müller & Gelbrich, 2004). Bolten (2007a) argumentiert, dass interkulturelle Kompetenz als prozessuales und nicht als abzugrenzendes Konstrukt verstanden und aufgrund der vielfältigen Bezüge zu anderen Kernkompetenzen als kontextspezifische Handlungskompetenz definiert werden muss (Bolten, 2007b). Eine weitere wichtige Gruppe stellen Entwicklungsmodelle interkultureller Sensitivität dar (z. B. Bennett, 1993; Bennett & Bennett, 2003; Lustig & Koester, 2003). Während interkulturelle Kompetenz sich darauf bezieht, dass eine Person sich angemessen verhält, umfasst interkulturelle Sensitivität die Fähigkeit, kulturelle Unterschiede wahrzunehmen (Hammer et al., 2003).

Die Anzahl und Art der Teilfacetten innerhalb verschiedener interkultureller Kompetenzmodelle variiert beträchtlich. Zu denen am häufigsten untersuchten Dimensionen gehören: Ambiguitätstoleranz (Fantini & Tirmizi, 2006; Koester & Olebe, 1988; Ruben 1976), Cultural Awareness (Chen & Starosta, 1997; Thomas, Kammhuber, & Layes, 1997; Triandis, 1977), Empathie (Fantini & Tirmizi, 2006; Koester & Olebe, 1988; Ruben, 1976), Flexibilität (Fantini & Tirmizi, 2006; Kelley & Meyers, 1995), Kommunikationsfähigkeit (Chen & Starosta, 1997; Fantini & Tirmizi, 2006), Offenheit für neue Erfahrungen (Chen & Starosta, 1997; Kelley & Meyers, 1995), Fähigkeit zum Perspektivenwechsel (Stüdlein, 1997) und Respekt gegenüber Verschiedenheit (Chen & Starosta, 1997; Fantini & Tirmizi, 2006). Die Heterogenität der Modelle ist auf das in der Literatur verbreitete Kriterienproblem interkultureller Kompetenz zurückzuführen (Stahl, 1998). Interkulturelle Kompetenz ist weder in ihrer Stabilität, ihrer Universalität, ihrer Generalisierbarkeit oder als einheitlicher Begriff bestätigt (Bolten, 2007a; Deardorff, 2006; Rathje, 2006).

Da keine der bestehenden Definitionen als alleinige Operationalisierungsgrundlage des zu entwickelnden Tests inhaltlich ausreicht, haben die Autoren dieser Arbeit eine eigene Arbeitsdefinition interkultureller Kompetenz erstellt. Dabei soll in Anlehnung an das Verständnis der Leistungsmotivation nach Schuler und Prochaska (2000) interkulturelle Kompetenz als globale Verhaltensorientierung verstanden werden, an der verschiedene Kompetenzen Anteil haben. Eine Person berücksichtigt dabei stetig Interkulturalität und Internationalität bei der Wahl ihrer Verhaltensweisen.

Interkulturelle Kompetenz wird als globale Verhaltensorientierung verstanden, an der eine Vielzahl von Teilkompetenzen aus den Kompetenzgruppen „Soziale Kompetenz“, „Personale Kompetenz“ und „Methoden Kompetenz“ beteiligt sind. Diese Teilkompetenzen beziehen sich auf die von einem Individuum erlernbaren Fähigkeiten, die im interkulturellen Kontext einzeln oder im Zusammenspiel zur Meisterung von neuartigen

Situationen sowie zum Lösen von Problemen befähigen und das Verhalten direkt beeinflussen. Ein interkultureller Kontext liegt dann vor, wenn ein Individuum mit Personen aus anderen Kulturen im In- und Ausland zusammentrifft und (persönlich oder virtuell) mit diesen zusammenarbeitet.

Messung interkultureller Kompetenz

Hinsichtlich der Messmethodik interkultureller Kompetenz können punktuelle von systemisch-prozessualen sowie indirekte von direkten Testverfahren unterschieden werden (Bolten, 2007a; Sinicrope et al., 2007). Zu den direkten Methoden gehören interkulturelle Assessment Center (z. B. Bolten, 2001; Kühlmann & Stahl, 1998; Müller-Neumann, 2005; Stumpf, Thomas, Zeutschel, & Ruhs, 2003) und Interviews (z. B. Byram, 1997; Fantini & Tirmizi, 2006; Strafford, 2003). Einschränkungen in der Ökonomie, Objektivität und Generalisierbarkeit direkter Messinstrumente führen nach wie vor zu einer bevorzugten Verwendung indirekter Verfahren (Bolten, 2007a; Sebald, 2008; Sinicrope et al., 2007). Im Folgenden werden die sieben am umfassendsten empirisch geprüften Instrumente beschrieben, die interkulturelle Kompetenz anhand von indirekten Selbst- oder Fremdeinschätzungen messen. Eine Ausnahme bildet das Intercultural Competence Assessment (Fantini & Tirmizi, 2006), bei dem verschiedene Methoden verwendet werden. Ein Vergleich der Merkmale der Instrumente befindet sich in Tabelle 1.
»Tabelle 1 hier einfügen«

Bei der *Behavioral Assessment Scale for Intercultural Communication (*BASIC*;* Koester & Olebe, 1988) müssen Beobachter Testpersonen bezüglich der Items *Display of Respect, Interaction Posture, Orientation to Knowledge, Empathy, Individualistic Roles, Relational Role Orientation, Task-related Role Orientation, Interaction Management* und *Tolerance of Ambiguity* einschätzen. Zur Überprüfung der Konstruktvalidität wurde der Gesamtscore von BASIC mit einem Item zur globalen kommunikativen Effektivität korreliert (r = .62).

Das Gesamtergebnis des *Intercultural Development Inventory* (IDI; Hammer, 2008; Hammer et al., 2003) zur Messung interkultureller Sensitivität zeigt die Position eines Individuums auf einem Entwicklungskontinuum mit den folgenden fünf Stufen: *Denial/Defence, Reversal, Minimization, Acceptance/Adaptation* und *Encapsulated Marginality*. Von den fünf Dimensionen korrelierten drei signifikant mit der *Worldmindedness Scale* (r = -.29 für *Denial/Defense*, r = .03 für *Acceptance/Adaptation*, r = .01 für *Cultural Marginality Scale*) und mit der *Intercultural Anxiety Scale* (r = .16 für

Denial/Defense, r = -.13 für *Acceptance/Adaptation*, r = .14 für *Cultural Marginality Scale*), was die Konstruktvalidität eingeschränkt bestätigte (Hammer, 2008).

Das *Cross-Cultural Adaptability Inventory* (CCAI; Kelley & Meyers, 1995) misst die Fähigkeit eines Individuums sich verschiedenen Kulturen anzupassen. CCAI umfasst vier Dimensionen: *Emotional Resiliance, Flexibility and Openness, Perceptual Acuity* und *Personal Autonomy.* Davis und Finney (2003) versuchten die Faktor-Struktur zu reproduzieren, was jedoch nicht gelang. Es liegen keine Ergebnisse zur Validität vor.

Der *Prospector* (Spreitzer, McCall & Mahoney, 1997) dient der Messung des Potenzials von zukünftigen internationalen Führungskräften bezüglich der Dimensionen *Sensitive to Cultural Differences, Business Knowledge, Courage, Brings Out the Best in People, Integrity, Insightful, Committed, Takes Risks, Seeks Feedback, Uses Feedback, Is Culturally Adventurous, Seeks Learning Opportunities, Open to Criticism* und *Flexibility.* Die Kriteriumsvalidität konnte für N = 838 Manager aus 21 Ländern bestätigt werden (Spreitzer et al., 1997).

Das *Intercultural Sensitivity Inventory* (ICSI; Bhawuk & Brislin, 1992) misst die Fähigkeit das eigene Verhalten in Abhängigkeit von der aktuellen kulturellen Umgebung zu verändern. Die Validität des Instruments wurde anhand von gemittelten Expertenratings erfasst, durch welche die Testpersonen in zwei Gruppen geteilt wurden. Dieselbe Gruppenteilung konnte anhand des Tests festgestellt werden.

The Intercultural Sensitivity Index (ISI; Olson & Kroeger, 2001) bringt die Entwicklungsstufen des IDI (Hammer, 2008; Hammer et al., 2003) mit der globalen Kompetenz in einem Instrument zusammen. Olson und Kroeger (2001) berichten ausschließlich deskriptive Informationen zum Test. Williams (2005) hat in einer Untersuchung das ISI verwendet und kam zu einer internen Konsistenz von α = .67.

Das *Intercultural Competence Assessment* (INCA; Fantini & Tirmizi, 2006) erfasst die gängigerweise zu interkultureller Kompetenz zählenden Teilfacetten Ambiguitätstoleranz, Verhaltensflexibilität, Kommunikationsbewusstsein, Wissenserwerb, Offenheit und Empathie. Untersuchungen zur Validität sind geplant (Sinicrope et al., 2007).

Im Großteil der Verfahren werden stabile Eigenschaften, Einstellungen oder interkulturelle Sensitivität erfasst (z. B. Bhawuk & Brislin, 1992; Hammer et al., 2003; Hannigan, 1990;

Olson & Kroeger, 2001; Ruben, 1989), so dass diese im beruflichen Kontext für den Personalauswahlkontext, nicht aber für Personalentwicklungszwecke herangezogenen werden können (Erpenbeck, 2012). Auch muss mit Verzerrungseffekten durch sozial erwünschtes Antwortverhalten gerechnet werden (Sebald, 2008). Verhaltensbezogene Kompetenzen werden überwiegend in direkten Verfahren beleuchtet, die in ihrer Zuverlässigkeit und Ökonomie eingeschränkt sind (Bolten, 2007a; Sebald, 2008; Sinicrope et al., 2007). Kombinierte Verfahren bilden interkulturelle Kompetenz am differenziertesten ab (Fantini & Tirmizi, 2006; Pruegger & Rogers, 1994; Straffon, 2003). Daher wird ebenfalls im *Test zur Messung interkultureller Kompetenz* (TMIK) eine multimethodale Herangehensweise gewählt, die zwei Methoden der indirekten Messung – Selbsteinschätzungsitems und Situationsbeurteilungsitems – verbindet.

Entwicklung des Messinstruments

Die Entwicklung des TMIK erfolgte in mehreren Schritten. Durch explorative Experteninterviews (Bogner, Littig, & Menz, 2005; Lamnek, 1995) und eine umfassende Literaturrecherche konnten 101 Teilkompetenzen gesammelt werden, die im Zusammenhang mit interkultureller Zusammenarbeit stehen. In einer Online-Vorstudie mit $N = 150$ international tätigen Personalmanagern konnten anhand einer explorativen Faktorenanalyse 25 Teilkompetenzen extrahiert werden. Im Sinne einer externalen Konstruktionsstrategie wurden in die erste Testversion nur Teilkompetenzen integriert, für die eine akzeptable interne Konsistenz ($\alpha < .75$; Bortz, 2006; Walsh & Betz, 2000) und ein Zusammenhang mit relevanten Erfolgskriterien ermittelt werden konnte.

Ein Testformat, das sowohl ökonomisch ist als auch die Vorteile eines direkten Instruments hat, ist der Situationsbeurteilungstest (Situational Judgment Test, SJT; McDaniel, Hartman, Whetzel, & Grubb, 2006). Bisher liegt kein Situationsbeurteilungstest zur Messung interkultureller Kompetenz vor. Ein Situationsbeurteilungstest besteht aus erfolgskritischen Situationen mit entsprechenden Antwortalternativen. Testpersonen müssen sich dann für eine Antwortalternative entscheiden (McDaniel et al., 2006). Situationsbeurteilungstests werden als besonders praxisnah, ökonomisch, valide und robust gegen Verzerrungseffekte bewertet (z. B. Hooper, Cullen, & Sackett, 2006; Lievens, Peeters, & Schollaert, 2008; McDaniel et al., 2006; McDaniel & Nguyen, 2001; Weekly & Ployhart, 2006).
Bei der Konstruktion des SJTs wurde sich an dem empfohlenen Vorgehen nach McDaniel und Whetzel (2005) orientiert. Als Basis für die Items dienten 20 Critical Incidents

(Flanagan, 1954) von Expatmanagern, Expatriates sowie interkulturellen Trainern, die in ein entsprechendes Format gebracht wurden. Es wurden textbasierte Situationsbeschreibungen und verhaltens- („Wie würden Sie sich am wahrscheinlichsten in dieser Situation verhalten?“) statt wissensbezogene („Welche Antwort ist die beste?“) Antwortalternativen gewählt, da letztere höher mit kognitiven Variablen korrelieren (McDaniel, Hartman, & Grubb, 2003; McDaniel et al., 2006) und stärker kulturabhängig sind (Nguyen, McDaniel, & Whetzel, 2005).

Aus den zuvor erläuterten Schritten wurde die erste Testversion generiert. Zur Messung der 17 Kompetenzfacetten wurden 79 eigens formulierte Selbsteinschätzungsitems, von denen 6 Items invertiert waren, und 17 Situationsbeurteilungsitems integriert. Für die Selbsteinschätzungsstatements wurde eine sechsstufige Likertskala mit den Abstufungen „*Trifft voll und ganz zu*“, „*Trifft zu*“, „*Trifft eher zu*“, „*Trifft eher nicht zu*“, „*Trifft nicht zu*“ und „*Trifft überhaupt nicht zu*“ verwendet, während die Probanden bei den Situationsbeurteilungsfragen, eine von vier Antwortkategorien auswählen sollten, die am ehesten auf sie zutrifft.

Hypothesen zum Testmodell

Es wird angenommen, dass die 17 durch Selbsteinschätzungsitems gemessenen Teilfacetten interkultureller Kompetenz anhand einer Faktorenanalyse durch eindeutig interpretierbare Faktoren abgebildet werden können.
Außerdem wird von einer Second-Order-Faktorenstruktur ausgegangen, die eine Zuordnung der 17 Kompetenzfacetten zu den folgenden sechs Dimensionen ermöglicht:

1. *Kommunikation im internationalen Kontext: Flexibilität in der Kommunikation* (4 Items + 1 SJT)*, Klarheit in der Kommunikation* (4 Items + 1 SJT)*, Empathie in der Kommunikation* (7 Items + 1 SJT)*, Perspektivenwechsel in der Kommunikation* (5 Items + 1 SJT)
2. *Lernen im internationalen Kontext: Bereitschaft zur Anwendung einer Fremdsprache* (4 Items + 1 SJT)*, Lernbereitschaft* (5 Items + 1 SJT)*, Gezieltes Sammeln von Informationen* (3 Items + 1 SJT)
3. *Soziale Interaktion im internationalen Kontext: Integration in Gruppen* (6 Items + 1 SJT)*, Aufbau und Pflege von Kontakten* (5 Items + 1 SJT)*, Aufbau eines beruflichen*

Netzwerks (3 Items + 1 SJT)

4. *Selbstmanagement im internationalen Kontext: Strategisches Problemlösen* (5 Items + 1 SJT) und *Zielorientierung* (4 Items + 1 SJT)
5. *Selbsterkenntnis im internationalen Kontext* : *Bewusstsein der eigenen Kultur* (6 Items + 1 SJT)*, Reflektion der eigenen Kultur* (4 Items + 1 SJT)
6. *Synergien schaffen im internationalen Kontext: Metakommunikation* (4 Items + 1 SJT)*, Mediation unterschiedlicher Interessen* (5 Items + 1 SJT)*, Verschiedenheit gezielt nutzen* (5 Items + 1 SJT)

Des Weiteren soll die interne Konsistenz der Kompetenzfacetten sowie des Gesamttests im befriedigenden bis sehr guten Bereich liegen.

Validierungshypothesen

Zunächst soll festgestellt werden, ob die Situationsbeurteilungsitems auch tatsächlich konvergent mit den Kompetenzfacetten korrelieren, zu denen diese anhand von Expertenratings zugeordnet worden sind. Gleichzeitig kann so auch die Konstruktvalidität festgestellt werden. Dies soll für alle Kompetenzfacetten des TMIK gelten, die mit zwei Methoden gemessen werden.

Des Weiteren soll die Kriteriumsvalidität des TMIK anhand der folgenden vier Außenkriterien überprüft werden:

Dauer eines vergangenen Auslandsaufenthalts. Durch Dissonanz-Erfahrungen im Ausland erlernt ein Individuum mit kultureller Unterschiedlichkeit umzugehen (Bhawuk & Brislin, 1992, Busch, 2007; Thomas, Kinast, & Schroll-Machl, 2006). Folglich wird erwartet, dass Personen, die eine längere Zeit im Ausland verbracht haben, einen höheren Gesamtwert interkultureller Kompetenz im Selbsteinschätzungsteil des TMIK haben als jene, die lediglich eine kurze Zeitdauer im Ausland verbracht haben.

Vergangene Teilnahme an einem interkulturellen Training. Eine Teilnahme an einem interkulturellen Training trägt dazu bei, durch gezielte Einübung von interkulturell kompetentem Verhalten, die Anpassung und Leistung eines Arbeitnehmers im Ausland zu erhöhen (Morris & Robie, 2001). Es wird daher erwartet, dass Personen, die bereits an einem interkulturellen Training teilgenommen haben, einen höheren Mittelwert in den 17 Kompetenzfacetten sowie einen höheren Gesamtwert im Selbsteinschätzungsteil des TMIK

aufweisen als Personen, die noch nie an einem interkulturellen Training teilgenommen haben.

Interkulturelle Involviertheit. Beruflich oder private Beschäftigung mit anderen Kulturen führt zu einer erhöhten Sensibilisierung für interkulturelle Kompetenzen (Loboda, 2003; Schacher, 2011). Dies soll zur Folge haben, dass eine interkulturelle Involviertheit mit einem höheren Mittelwert in den 17 Kompetenzfacetten sowie einem höheren Gesamtergebnis des Selbsteinschätzungsteils im TMIK einhergeht.

Interkulturelle Kompetenz Global. Zur Feststellung der Kriteriumsvalidität kann ein globales Maß interkultureller Kompetenz verwendet werden (Koester & Olebe, 1988). Die Annahme besteht, dass Personen, die sich als insgesamt interkulturell kompetent beschreiben, auch einen höheren Mittelwert in den 17 Kompetenzfacetten sowie einen höheres Gesamtergebnis im Selbsteinschätzungsteil des TMIK aufweisen.

Methode

Stichprobe

Für die Onlinefragebogenstudie zur Testkonstruktion und -validierung wurde eine studentische Stichprobe rekrutiert. Dazu wurden die E-Mail-Verteilerlisten der Fachschaften von deutschen Hochschulen genutzt. In der Auswertung wurden zwei Fälle ausgeschlossen, bei denen nur für einen Teil des Fragebogens Daten vorlagen. Insgesamt betrug der Stichprobenumfang dann $N = 641$ Personen, wovon 448 (69.90 %) weiblich und 193 (30.10 %) männlich waren. Im Durchschnitt waren die Probanden 28.25 Jahre ($SD = 9.29$) alt, wobei die Alterspanne von 18 bis 65 Jahren reichte. Die weiblichen Personen waren im Durchschnitt 27.45 ($SD = 8.39$) und die männlichen Personen 30.12 ($SD = 10.91$) Jahre alt. Der Altersunterschied zwischen den Geschlechtern war signifikant ($p < .05$; $t(639) = -3.37$, $d = .27$). Insgesamt gaben 574 Probanden an, der Deutschen Nationalität anzugehören. Die restlichen 67 Personen verteilten sich auf 32 weitere Nationen. 430 Befragte gingen ausschließlich dem Studium nach, während die restlichen 211 Personen aktuell überwiegend andere Tätigkeiten ausübten (Ausbildung = 1, Anstellung = 137; Selbstständigkeit = 57). Die Abfrage des höchsten Bildungsgrades ergab, dass 282 Personen einen Schulabschluss, 77 Personen eine abgeschlossene Berufsausbildung, 250 Personen ein abgeschlossenes Hochschulstudium, 21 Personen eine Promotion und 11 Personen einen anderen Abschluss

vorweisen konnten. Unter allen Studienteilnehmern wurden drei Amazon-Gutscheine verlost (zwei Gutscheine je 25 Euro und ein Gutschein im Wert von 50 Euro).

Erfassung der Außenkriterien

Neben dem zuvor beschriebenen neu entwickelten Instrument zur Messung interkultureller Kompetenz wurden die vier Außenkriterien mit jeweils einem Item im Fragebogen erfasst. Dabei wurden die zwei folgenden Antwortformate verwendet: sechsstufige Likertskala („*Ich halte mich für interkulturell kompetent.“ Trifft voll und ganz zu* bis *Trifft überhaupt nicht zu*) und Forced-Choice (z. B. *„Haben Sie schon einmal an einem interkulturellen Training teilgenommen?“ Ja*/*Nein*). Bei der Formulierung der Items für die Außenkriterien wurde sich an vergangenen Studien orientiert (Bhawuk & Brislin, 1992; Koester & Olebe, 1988; Schacher, 2011).

Ergebnisse

Faktorielle Struktur

Das ESEM-Verfahren (*Exploratory Structural Equation Modeling*; Asparouhov & Muthén, 2009) wurde verwendet, um sowohl die Faktorenstruktur als auch die Güte des Basismodells zu ermitteln. Die vorhergesagte 17-faktorielle Lösung konnte im Wesentlichen reproduziert werden. Aufgrund von Faktorladungen < .20 wurden vier Items ausgeschlossen (Bortz, 2006). Neun Items wurden anderen Faktoren zugeordnet als ursprünglich angenommen. Dies hatte zur Folge, dass die Kompetenzfacette *Verschiedenheit Gezielt Nutzen* wegfiel und *Integration in Gruppen* in zwei Faktoren aufgespaltet wurde. Die thematische Verschiebung im Faktor *Metakommunikation* führte außerdem dazu, dass eine adäquatere Bezeichnung, nämlich *Produktive Zusammenarbeit Ermöglichen,* gewählt wurde.
Das finale Strukturgleichungsmodell bestand aus 75 Selbsteinschätzungsitems und wurde auf seine Modellgüte untersucht. Der Chi-Quadrat-Wert des gefitteten Modells betrug 2579.85 ($df = 1636$); $p < .001$. Das Verhältnis zwischen dem Chi-Quadrat-Wert und den Freiheitsgraden ergab $\chi^2/df = 1.6$, was als sehr gut zu bewerten ist (Bollen & Long, 1993). Der Comparative Fit Index (CFI) lag mit .96 über dem Cut-Off-Wert von .95 und der Tucker Lewis Index (TLI) mit .93 über dem empfohlenen Wert von .90, was ebenfalls auf eine gute Modellpassung hinweist (Bentler & Bonnett, 1980; Hu & Bentler, 1998). Der Root Mean Square Error of Approximation (RMSEA) betrug .030 (bei RMSEA 90 % KI = .028), und

das Standardisierte Root Mean Square Residual (SRMR) .017, wodurch eine sehr gute Modellpassung angenommen werden kann (Cudeck & Browne, 1993).

Zur Absicherung des psychometrischen Modells wurden mehrere alternative Modelle gefittet. Tabelle 2 zeigt den Vergleich der 17-faktoriellen Lösung mit einem Modell, das 16 und einem Modell, das 15 Faktoren beinhalten sollte. Neben den zuvor erwähnten Fit-Statistiken wurden als Entscheidungskriterien ebenfalls der Akaike (AIC), der Bayesian (BIC) sowie der Sample-Size Adjusted BIC betrachtet. Besonders der deutlich niedrigere AIC, BIC und Sample-Size Adjusted BIC im 17-faktoriellen Modell sprechen eindeutig für eine Beibehaltung desselbigen.

»Tabelle 2 hier einfügen«

Außerdem sollte eine Second-Order Struktur der Faktoren überprüft werden, um die vorhergesagte Zuordnung der Teilkompetenzen zu den sechs Kompetenzbereichen bestätigen zu können. Dazu wurde eine konfirmatorische Faktorenanalyse durchgeführt[1], die neben den 17 Kompetenzfacetten auf Selbsteinschätzungsbasis mit den zugehörigen 16 Situationsbeurteilungsitems sechs weitere latente Variablen beinhaltete. Für dieses Modell ergab sich folgende Güte: $\chi^2(3972) = 7492.519$, $p < .001$; $\chi^2/df = 1.80$; RMSEA = .037, RMSEA 90 % KI = .036; SRMR = .066; CFI = .85. Alle Fit-Indices bis auf der CFI wiesen auf eine akzeptable bis sehr gute Modellpassung hin. Marsh, Hau, und Wen (2004) argumentieren, dass bei mehr als 5 Faktoren und über 50 Items das überwiegend postulierte Cut-Off-Kriterium von Fit-Indices > .90 zu restriktiv ist.

Statistische Kennwerte der Items und Skalen

Tabelle 3 zeigt die statistischen Kennwerte der 17 Kompetenzskalen sowie der Gesamtskala interkultureller Kompetenz. Die interne Konsistenz der Kompetenzfacetten liegt mit Cronbachs Alpha zwischen .69 und .90 (Guttmans Split-Half-Koeffizient .62–.88) im akzeptablen bis sehr guten Bereich. Die Reliabilität der Gesamtskala ist mit $\alpha = .96$ als hervorragend zu bewerten.

»Tabelle 3 hier einfügen«

[1] Es ist mit der aktuellen Fassung des ESEM-Verfahrens in Mplus nicht möglich, Modelle höherer Ordnung zu spezifizieren und zu schätzen. Daher wurde auf die Alternative der Standard-Konfirmatorischen Faktorenanalyse zurückgegriffen.

Die Trennschärfe lag bei allen Selbsteinschätzungsitems über der Untergrenze von $r_{i} = .30$ (Moosbrugger & Kelava, 2012). Die Werte bewegten sich in einem Bereich zwischen r_i = .37 und r_i = .63. Weiterhin wurden für alle Items eine Itemschwierigkeit zwischen .45 und .79 ermittelt, was im zulässigen Rahmen von $P_i = .20$ und $P_i = .80$ liegt.

Des Weiteren wurden die Interkorrelationen der 17 Faktoren überprüft. Während die überwiegende Anzahl der Faktoren bedeutsame Korrelationen zwischen $r = .15$ und $r = .69$ aufwiesen, lagen vier Zusammenhänge über .70: *Empathie in der Kommunikation* und *Perspektivenwechsel in der Kommunikation* ($r = .70$, $p < .001$), *Aufbau von Beziehungen* und *Integration in Gruppen* ($r = .79$, $p < .001$), *Aufbau von Beziehungen* und *Aufbau eines beruflichen Netzwerks* ($r = .70$, $p < .001$) sowie *Aufbau von Beziehungen* und *Aufbau und Pflege von Kontakten* ($r = .83$, $p < .001$). Jeweils kein signifikanter Zusammenhang konnte zwischen *Integration in Gruppen* und den folgenden vier Faktoren festgestellt werden: *Gezieltes Sammeln von Informationen* ($r = .00$, $p = .87$)*, Strategisches Problemlösen* ($r = .01$, $p = .60$)*, Reflektion der eigenen Kultur* ($r = .01$, $p = .73$) sowie *Perspektivenwechsel in der Kommunikation* ($r = .05$, $p = .07$).

Validierung

Konstruktvalidität durch den multimethodalen Ansatz. Tabelle 4 zeigt die Korrelationen zwischen den 17 latenten Faktoren, bestehend aus den Selbsteinschätzungsitems im Test und den 16 Situationsbeurteilungsitems. 16 Faktoren weisen einen positiven Zusammenhang mit den inhaltlich entsprechenden Situationsbeurteilungsitems auf, wobei die Stärke des Zusammenhangs eher gering ist. Keine signifikante Korrelation konnte zwischen *Flexibilität in der Kommunikation* (*FK*) und dem passenden Situationsbeurteilungsitem (*SJT_FK*) festgestellt werden.

»Tabelle 4 hier einfügen«

Dauer des vergangenen Auslandsaufenthalts. Um zu untersuchen, ob Personen, die eine längere Zeit im Ausland verbracht haben, höhere Werte im Selbsteinschätzungsfragebogen erzielen, wurde eine Varianzanalyse durchgeführt. Aufgrund der hohen Anzahl an Gruppenvergleichen, wurde für den gesamten Validierungsteil eine Bonferroni-Holm Korrektur des Signifikanzniveaus vorgenommen (Holm, 1979). Als signifikant gelten demnach Ergebnisse auf einem Niveau von $p < .001$.

Es konnte ein bedeutsamer Effekt für den Einfluss der Länge eines Auslandsaufenthaltes auf den Gesamtwert interkultureller Kompetenz im Selbsteinschätzungsfragebogen festgestellt werden ($F(3, 337) = 34.06$, $p < .001$, $\eta^2 = .86$). Hochbergs GT2 Posthoc-Test ergab, dass Personen, die bereits einen Auslandsaufenthalt von über drei Monaten absolviert hatten interkulturell kompetenter waren ($M = 4.40$, $SD = 0.45$) als diejenigen, die lediglich bis zu zwei Wochen ($M = 3.90$, $SD = 0.43$) oder zwischen zwei und vier Wochen im Ausland waren ($M = 4.11$, $SD = 0.46$). Kein signifikanter Unterschied in der interkulturellen Kompetenz ergab sich allerdings zwischen einem Auslandsaufenthalt von ein bis drei Monaten ($M = 4.29$, $SD = 0.49$) und einem Auslandsaufenthalt von einer Dauer über drei Monaten.[2]

Vergangene Teilnahme an einem interkulturellen Training. Testteilnehmer, die bereits an einem interkulturellen Training teilgenommen haben, waren insgesamt auch interkulturell kompetenter im Selbsteinschätzungsfragebogen ($M = 4.37$, $SD = 0.45$) als Personen, die noch nie ein interkulturelles Training erhalten haben ($M = 4.14$, $SD = 0.48$); ($p < .001$; $t(639) = 7.87$; $d = .69$). Des Weiteren wurden Mehrgruppenvergleiche für die 17 Kompetenzfacetten durchgeführt, um zu prüfen, ob sich auch diese in ihrer Ausprägung in Abhängigkeit von der Teilnahme an einem interkulturellen Training unterscheiden. Für 13 Faktoren war dies der Fall (s. Tabelle App-A des Anhangs). Dabei lagen die Mittelwerte der latenten Faktoren in der Gruppe „*Training*“ jeweils über denen in der Gruppe „*kein Training*“. Kein signifikanter Unterschied zwischen den Gruppen konnte für die Kompetenzfacetten *Zielorientierung* ($\chi^2(1) = 6.19$, $p = .01$), *Perspektivenwechsel in der Kommunikation* ($\chi^2(1) = 1.54$, $p = .21$), *Empathie in der Kommunikation* ($\chi^2(1) = .37$, $p = .24$) und *Strategisches Problemlösen* ($\chi^2(1) = 2.84$, $p = .09$) festgestellt werden.

Interkulturelle Involviertheit. Der Hypothese entsprechend erzielten Personen, die angaben sich beruflich oder privat mit interkulturellen Themen zu beschäftigen einen höheren Gesamtwert der Selbsteinschätzung interkultureller Kompetenz ($M = 4.31$, $SD = 0.47$) als Testteilnehmer, die angaben nicht interkulturell involviert zu sein ($M = 3.90$, $SD = 0.44$); ($p < .001$; $t(639) = 8.69$; $d = .70$). Die Mehrgruppen-Vergleiche für die 17 Kompetenzfacetten führten zu 13 bedeutsamen Unterschieden zwischen den Gruppen *interkulturelle Involviertheit* und *keine interkulturelle Involviertheit.* Dabei waren die Mittelwerte der

[2] Des Weiteren wurden Mehrgruppen-Vergleiche für die 17 Faktoren und das Außenkriterium Dauer des vergangenen Auslandsaufenthaltes durchgeführt, deren Ergebnisse extern eingesehen werden können.

latenten Faktoren höher bei Personen, die sich beruflich oder privat mit interkulturellen Themen beschäftigten (s. Tabelle App-B des Anhangs). Keine Unterschiede wurden für das *Strategische Problemlösen* ($\chi^2(1) = 1.40$, $p = .24$), die *Zielorientierung* ($\chi^2(1) = 3.22$, $p = .07$), die *Produktive Zusammenarbeit ermöglichen* ($\chi^2(1) = 4.94$, $p = .03$) und die *Flexibilität in der Kommunikation* ($\chi^2(1) = 8.62$, $p = .003$) festgestellt.

Globales Maß interkultureller Kompetenz. Alle 17 Kompetenzfacetten korrelierten wie erwartet positiv mit dem globalen Maß interkultureller Kompetenz (siehe Tabelle 5). Die Zusammenhänge lagen überwiegend im mittleren Bereich. Auch der Gesamtwert des Tests korrelierte positiv mit dem globalen Maß interkultureller Kompetenz ($r = .58$, $p < .001$).

»Tabelle 5 hier einfügen«

Diskussion

Ziel der vorgestellten Untersuchung war es, einen Test zu entwickeln, der auf einem Verständnis interkultureller Kompetenz als globale Verhaltensorientierung (Schuler & Prochaska, 2000) basiert, so dass dem hochdimensionalen Charakter interkultureller Kompetenz Rechnung getragen werden kann. Zur Entwicklung des Modells interkultureller Kompetenz wurde eine phänomenologisch-expertenbasierte sowie empirische Strategie verwendet. Diese führte zur Postulierung und Überprüfung eines aus 17 Kompetenzen bestehenden Modells sowie einer Second-Order-Struktur der Faktoren. Um den Bedarf nach einem ökonomischen, reliablen und validen Instrument in der internationalen Personalauswahl und -entwicklung zu decken (Bolten, 2007a), wurde ein multimethodaler Test geschaffen, der sowohl Selbsteinschätzungs- als auch Situationsbeurteilungsitems enthält.

Zusammenfassung der Ergebnisse. Das vorhergesagte 17-faktorielle Modell interkultureller Kompetenz konnte reproduziert werden; die Skalen wiesen dabei eine zufriedenstellende bis sehr gute Reliabilität auf. Auch ein Second-Order-Modell aus Selbsteinschätzungs- und Situationsbeurteilungsitems, in welchem die 17 Kompetenzfacetten den sechs vorhergesagten Kompetenzbereichen zugeordnet wurden, ergab einen befriedigenden Fit zu den Daten.

Die primären Abweichungen zum vorhergesagten Modell wurden aufgrund der inhaltlichen Stimmigkeit übernommen. Demnach wurde auch die Aufspaltung des Faktors *Integration in*

Gruppen in zwei Faktoren zugelassen, so dass nachträglich ein Faktor *Aufbau vertrauensvoller Beziehungen* entstand. Streng genommen ist die Fähigkeit, sich ohne weiteres in einer Gruppe aus unbekannten Personen zu positionieren (Integrationsfähigkeit; Niermeyer, 2006) von der Fähigkeit, schnell Beziehungen zu den Personen in dieser Gruppe aufzubauen (Rapport; Stumm & Pritz, 2009) zu trennen.

Die Korrelationen der Situationsbeurteilungsitems mit den latenten Variablen führten für 16 von 17 Faktoren zu den vorhergesagten positiven Zusammenhängen, so dass die konvergente Validität des Instruments größtenteils bestätigt werden konnte. Einzig der Faktor *Kommunikative Flexibilität* korrelierte nicht signifikant mit dem entsprechenden Situationsbeurteilungsitem. Dies entspricht der generellen Schwierigkeit wechselseitige Kommunikation mit einem SJT abzubilden (vgl. Watzlawick, Beavin, & Jackson, 2011).

Zur Überprüfung der Kriteriumsvalidität wurde der Gesamttestwert in einen Zusammenhang mit vier Außenkriterien gebracht. Wie vorhergesagt, erzielten Personen einen höheren Gesamtwert im Test, die a) eine längere Zeit als 2 Wochen am Stück im Ausland verbracht haben, b) bereits an einem interkulturellen Training teilgenommen haben, c) sich privat oder beruflich mit interkulturellen Themen beschäftigten und d) ihre globale interkulturelle Kompetenz als hoch einschätzten.

Die Mehrgruppenvergleiche für die 17 Faktoren und die Außenkriterien *Teilnahme an einem interkulturellen Training* und *interkulturelle Involviertheit* führten für die überwiegende Anzahl an Faktoren zu hypothesenkonformen Ergebnissen. Einzig für die Faktoren *Strategisches Problemlösen* und *Zielorientierung* konnten keine der vorgesagten Mittelwertsunterschiede festgestellt werden. Auffällig ist, dass beide Kompetenzen zwar von den Experten als für die internationale Zusammenarbeit hoch relevant bewertet wurden, jedoch in interkulturellen Trainings und Coachings bislang zu wenig Beachtung fanden. Dies sollte zukünftig stärker geschehen, zu mal beide Kompetenzen, wie auch alle anderen 15 Faktoren, in einem positiven Zusammenhang mit der globalen Einschätzung interkultureller Kompetenz standen.

Limitationen. Der TMIK eignet sich vor allem für den beruflichen Kontext. Zwar kann angenommen werden, dass auch Studenten internationale Berufserfahrung gesammelt haben und die interkulturelle Zusammenarbeit im Studium eine wichtige Rolle spielt, jedoch bleibt offen, ob die Ergebnisse auf international tätige Mitarbeiter generalisierbar sind. Weiterhin

kann davon ausgegangen werden, dass die Zusammenhänge zwischen den Faktoren und den Situationsbeurteilungsitems noch stärker bei international tätigen Mitarbeitern abbildbar wären, die besser als Studenten mit kritischen Situationen in der internationalen Zusammenarbeit vertraut sind. Die Konstruktvalidierung des *Tests zur Messung interkultureller Kompetenz* wird erschwert durch die geringe Vergleichbarkeit bestehender Konzepte und Instrumente (Rathje, 2006). Daher wurde zur Validierung des TMIK kein anderes Instrument, sondern der Zusammenhang mit Außenkriterien sowie der multimethodale Vergleich herangezogen. Weitere Untersuchungen sind nötig, um die psychometrische Güte der Situationsbeurteilungsitems zu überprüfen. Die Zuordnung der Kompetenzen zu den Situationen durch Experten ist zwar ein schlüssiges Vorgehen, das jedoch nicht frei von Subjektivität ist (Findeisen, Kim & Dietz, 2012). Umso bedeutsamer ist der Zusammenhang mit der Selbsteinschätzung der Kompetenzen, so dass dieser eingehender untersucht werden sollte.

Ausblick. Die Annahme, dass es keine allgemeingültige interkulturelle Kompetenz gibt, sondern dass interkulturelle Kompetenz sowohl kultur- (Bolten, 2007a) als auch kontextspezifisch ist (Arasaratnam & Doerfel, 2005) sollte auch für den TMIK beleuchtet werden. Einerseits wäre es deshalb wichtig, den Test bei international tätigen Mitarbeitern einzusetzen. Dabei könnte überprüft werden, ob sich das 17-faktorielle Kompetenzmodell auch in einer anderen Stichprobe reproduzieren lässt. Andererseits sollte der Einfluss von Kultur auf das Kompetenzmodell sowie der kulturübergreifende Einsatz des Tests untersucht werden. Dies könnte auch die Feststellung eines möglichen Zusammenhangs zwischen den interkulturellen Kompetenzen und interkulturellen Präferenzen einer Person (z. B. Hofstede & Bond, 1984) beinhalten. Die interkulturellen Kompetenzen im Test wurden bereits mit weiteren Außenkriterien in einen Zusammenhang gebracht, was zu vielversprechenden Ergebnissen führte. Daher wäre es sinnvoll, weitere Studien anzuschließen, die auch die Vorhersagekraft des Ergebnisses im TMIK für Erfolg und Wohlbefinden im Ausland beleuchten (vgl. Müller & Gelbrich, 2004).

Anwendung des Tests. Der TMIK erfüllt den Anspruch eines reliablen und validen Instruments, das einen wichtigen Beitrag zur interkulturellen Personaldiagnostik und -entwicklung leistet. Der TMIK sollte vorwiegend im berufsbezogenen Kontext angewendet werden und eignet sich zur umfassenden Analyse im Rahmen von Coachings oder Trainings. Außerdem sollte mit einem Instrument zur Messung interkultureller Kompetenz, der die

psychometrischen Standards erfüllt, Vorschub geleistet werden, interkultureller Kompetenz zukünftig mehr Bedeutung in der empirischen Wissenschaft zu schenken.

2. Literaturverzeichnis

Arasaratnam, L. A. & Doerfel, M. L. (2005). Intercultural communication competence: Identifying key components from multicultural perspectives. *International Journal of Intercultural Relations, 29*, 137–163.

Asparouhov, T. & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 397–438.

Bhawuk, D. & Brislin, R. (1992). The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International Journal of Intercultural Relations, 16,* 413–436.

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Faige (Ed.), *Education for the Intercultural Experience* (2nd ed., pp. 21–77). Yarmouth, ME: Intercultural Press.

Bennett, J. M. & Bennett, M. J. (2003). Developing intercultural sensitivity. In D. Landis, J. M. Bennett & M. J. Bennett (Eds.). *Handbook of Intercultural Training* (3rd ed., pp. 147–165). Thousand Oaks, CA: Sage.

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

Bogner, A., Littig, B., & Menz, W. (2005). *Das Experteninterview: Theorie, Methode, Anwendung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Bollen, K. A. & Long, J. S. (Eds.). (1993). *Testing Structural Equation Models.* Newbury Park, CA: Sage Publications.

Bolten, J. (2001). Interkulturelle Assessment-Center. In W. Sarges (Hg.), *Weiterentwicklungen der Assessment Center-Methode* (2. überarb. Aufl., S. 213–218). Göttingen: Hogrefe.

Bolten, J. (2007a). Was heißt Interkulturelle Kompetenz? Perspektiven für die internationale Personalentwicklung. In V. Künzer & J. Berninghausen (Hrsg.), *Wirtschaft als interkulturelle Herausforderung* (S. 21–42). Frankfurt am Main: IKO.

Bolten, J. (2007b). *Interkulturelle Kompetenz.* Erfurt: Thüringer Landeszentrale für politische Bildung.

Bortz, J. (2006). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.). Heidelberg: Springer Medizin.

Brislin, R. W. (1981). *Cross-Cultural Encounters. Face-to-Face Encounters*. New York, NY: Pergamon.

Busch, D. (2007). *Interkulturelle Mediation: Eine theoretische Grundlegung triadischer Konfliktbearbeitung in interkulturell bedingten Kontexten* (2. Aufl.). Frankfurt am Main: Lang.

Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence.* Philadelphia, PA: Multilingual Matters.

Chen, G.-M. & Starosta, W. G. (1998). *Foundations of Intercultural Communication.* Needham Heights, MA: Allyn & Bacon.

Cudeck, R. & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research, 18,* 147–167.

Dauner, A. B. (2011). Die organisierende Funktion von Kompetenz – Entwurf eines zwei-Ebenen-Modells interkultureller Kompetenz. *Inter Culture Journal*, *10*(14), 25–46.

Davis, S. L. & Finney, S. J. (2003, April). *Examining the Psychometric Properties of the Cross-Cultural Adaptability Inventory*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Deardorff, D. K. (2006). *Policy Paper zur Interkulturellen Kompetenz*. Gütersloh: Bertelsmann Stiftung.

Erpenbeck, J. (2012). Siegeszug trotz Kritik: Warum Kompetenzdenken notwendig ist. *Wirtschaftspsychologie Aktuell, 19*, 19–22.

Erpenbeck, J. & von Rosenstiel, L. (Hrsg.). (2007). *Handbuch Kompetenzmessung. Erkennen, Verstehen und Bewerten von Kompetenzen in der Betrieblichen, Pädagogischen und Psychologischen Praxis* (2. Aufl.). Stuttgart: Schäffer-Poeschel.

Fantini, A. & Tirmizi, A. (2006). Exploring and assessing intercultural competence. *World Learning Publications,* Paper 1. Verfügbar unter http://digitalcollections.sit.edu/worlslearning_publications/1

Findeisen, A., Kim, C. & Dietz, F. (2012). Führungspotenziale messen. *Wirtschaftspsychologie Aktuell, 19,* 9–12.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327–358.

Gertsen, M. C. (1990). Intercultural competence and expatriates. *The International Journal of Human Resource Management, 1*, 341–362.

Hammer, M. R. (2008). The Intercultural Development Inventory (IDI): An approach for assessing and building intercultural competence. In M. A. Moodian (Ed.), *Contemporary Leadership and Intercultural Competence: Understanding and*

Utilizing Cultural Diversity to Build Successful Organizations (pp. 245–259). Thousand Oaks, CA: Sage.

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations, 27,* 421–443.

Hannigan, T. P. (1990). Traits, attitudes, and skills that are related to intercultural effectiveness and their implications for cross-cultural training: A review of the literature. *International Journal of Intercultural Relations, 14,* 89–111.

Hofstede, G. & Bond, M. H. (1984). Hofstede's Culture Dimensions. An independent validation using Rokeach's Value Survey. *Journal of Cross-Cultural Psychology, 15*, 417–433.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70.

Hooper, A. C., Cullen, M. J., & Sackett, P. R. (2006). Operational threats to the use of SJTs: Faking, coaching, and retesting issues. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational Judgement Tests: Theory, Measurement and Application* (pp. 205–232). London: Lawrence Erlbaum Associates.

Hu, L. & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453.

Kealey, D. J. & Ruben, B. D. (1983). Cross-cultural personnel selection: Criteria, issues, and methods. In D. Landis & R. W. Brislin (Eds.), *Handbook of Intercultural Training (Vol.1): Issues in Theory and Design* (pp. 155–175). New York, NY: Pergamon Press.

Kelley, C. & Meyers, J. (1995). *The Cross-Cultural Adaptability Inventory (manual)*. Minneapolis, MN: National Computer Systems.

Koester, J. & Olebe, M. (1988). The behavioral assessment scale for intercultural communication effectiveness. *International Journal of Intercultural Relations, 12*, 233–246.

Kulturministerkonferenz (2011). *Handreichung für die Erarbeitung von Rahmenlehrplänen der Kultusministerkonferenz für den berufsbezogenen Unterricht in der Berufsschule und ihre Abstimmung mit Ausbildungsordnungen des Bundes für anerkannte Ausbildungsberufe.* Sekretariat der Kultusministerkonferenz: Berlin.

Kühlmann, T. M. & Stahl, G. K. (1998). Diagnose interkultureller Kompetenz: Entwicklung

und Evaluierung eines Assessment-Centers. In C. I. Barmeyer & J. Bolten (Hrsg.), *Interkulturelle Personalorganisation* (S. 213–224). Sternenfels: Verlag für Wissenschaft und Praxis.

Lamnek, S. (1995). *Qualitative Sozialforschung* (3. Aufl.). Weinheim: Beltz.

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review, 37*, 426–441.

Loboda, J. (2003). Entwicklung eines Tests zur Messung Interkultureller Handlungskompetenz (TIHK). Unveröffentlichte Diplomarbeit an der Universität Regensburg. Verfügbar unter www.psychologie.uni-regensburg.de/Thomas/lehre/PWP_IK/DALoboda_Gesamt.pdf

Lustig, M.W. & Koester, J. (2003). *Intercultural competence: Interpersonal communication across cultures* (4th ed.). Boston, MA: Allyn & Bacon.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 320–341.

McDaniel, M. A., Hartman, N. S., & Grubb, W. L. (2003, April). *Situational Judgment Tests, Knowledge, Behavioral Tendency, and Validity: A Meta-Analysis.* Paper presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2006). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*, 63–91.

McDaniel, M. A. & Nguyen, N. T. (2001). Situational judgment test: A review of practice and constructs assessed. *International Journal of Selection and Assessment, 9,* 103–113.

McDaniel, M. A. & Whetzel, D. L. (2005, June). *Situational Judgment Tests.* Paper presented at the 29th Annual IPMAAC Conference on Personnel Assessment, Orlando, FL.

Morris, M. A. & Robie, C. (2001). A meta-analysis of the effects of cross-cultural training on expatriate performance and adjustment. *International Journal of Training and Development, 5*, 112–125.

Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion* (2. neubearbeitete Aufl.). Berlin: Springer.

Müller, S. & Gelbrich, K. (2004). *Interkulturelles Marketing*. München: Vahlen.

Müller-Neumann, C. (2005). Assessment-Center interkulturell. In V. Fischer, M. Springer, & I. Zacharaki (Hrsg.), *Interkulturelle Kompetenz. Fortbildung – Transfer – Organisationsentwicklung* (S. 159–179). Schwalbach: Wochenschau-Verlag.

Niermeyer, R. (2006). *Soft Skills: Das Kienbaum Trainingsprogramm* (2. Aufl.). München: Haufe-Lexware.

Nguyen, N., McDaniel, M. A., & Whetzel, D. L. (2005, April). *Subgroup Differences in Situational Judgment Test Performance: A Meta-Analysis*. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.

Olson, C. L. & Kroeger, K. R. (2001). Global Competency and Intercultural Sensitivity. *Journal of Studies in International Education, 5*, 116–137.

Pruegger, V. J. & Rogers, T. B. (1994). Cross-cultural sensitivity training: Methods and assessments. *International Journal of Intercultural Relations*, *18*, 369–387.

Rathje, S. (2006). Interkulturelle Kompetenz – Zustand und Zukunft eines umstrittenen Konzepts. *Zeitschrift für Interkulturellen Fremdsprachenunterricht, 11*(3), 1–21.

Ruben, B. D. (1976). Assessing communication competency for intercultural adaptation. *Group & Organization Studies, 1*, 334–354.

Ruben, B. D. (1989). The study of cross-cultural competence: Traditions and contemporary issues. *International Journal of Intercultural Relations*, *13*, 229–240.

Schacher, C. (2011). *Der Perspektivenwechsel in interkulturellen Konfliktsituationen.* Unveröffentlichtes Manuskript, Technische Universität Darmstadt.

Schuler, H. & Prochaska, M. (2000). Entwicklung und Konstruktvalidierung eines berufsbezogenen Leistungsmotivationstests. *Diagnostica, 46*, 61–72.

Sebald, J. A. P. (2008). *Effektivität und Angemessenheit – Messung Interkultureller Kompetenz im Assessment Center.* Unveröffentlichte Dissertation, Friedrich-Schiller-Universität Jena.

Sinicrope, C., Norris, J., & Watanabe, Y. (2007). Understanding and assessing intercultural competence: A summery of theory, research, and practice. *Second Language Studies, 26*, 1–58.

Spreitzer, G. M., McCall, Jr., M. W., & Mahoney, J. D. (1997). Early identification of international executive potential. *Journal of Applied Psychology, 82*, 6–29.

Stahl, G. K. (1998). *Internationaler Einsatz von Führungskräften*. München: Oldenburg.

Straffon, D. A. (2003). Assessing the intercultural sensitivity of high school students attending an international school. *International Journal of Intercultural Relations, 27*, 487–501.

Stumpf, S., Thomas, A., Zeutschel, U. & Ruhs, D. (2003). Assessment Center als Instrument zur Förderung der Handlungskompetenz von Fach- und Führungskräften in der internationalen Jugendarbeit. In Internationaler Jugendaustausch- und Besucherdienst der Bundesrepublik Deutschland (IJAB) e.V. (Hrsg.), *Forum Jugendarbeit International 2003* (S. 70–91). Münster: Votum.

Stumm, G. & Pritz, A. (Hrsg.). (2009). *Wörterbuch der Psychotherapie*. Wien: Springer.

Stüdlein, Y. (1997). *Management von Kulturunterschieden. Phasenkonzept für internationale strategische Allianzen.* Wiesbaden: Deutscher Universitätsverlag.

Ting-Toomey, S. (1993). Communicative resourcefulness: An identity negotiation perspective. In R. L. Wiseman & J. Koester (Eds.), *Intercultural Communication Theory* (pp. 72–111). Newbury Park, CA: Sage.

Thomas, A. (2006). Interkulturelle Handlungskompetenz – Schlüsselkompetenz für die moderne Arbeitswelt. *Arbeit*, *15*, 114–125.

Thomas, A., Kammhuber, S., & Layes, G. (1997). Interkulturelle Kompetenz. Ein Handbuch für internationale Einsätze der Bundeswehr. *Untersuchungen des Psychologischen Dienstes der Bundeswehr, 32*, 7–190.

Thomas, A., Kinast, E.-U., & Schroll-Machl, S. (2006). Entwicklung interkultureller Handlungskompetenz von international tätigen Fach- und Führungskräften durch interkulturelle Trainings. In K. Götz (Hrsg.) *Interkulturelles Lernen / Interkulturelles Training* (S. 91–114). München: Rainer Hampp Verlag.

Triandis, H. C. (1977). *Interpersonal Behavior.* Monterey, CA: Brooks/Cole.

Walsh, W. & Betz, N. (2000). *Tests and Assessment.* Upper Saddle River, NJ: Prentice Hall.

Watzlawick, P., Beavin, J. H., & Jackson, D. D. (2011). *Menschliche Kommunikation: Formen, Störungen, Paradoxien* (12. Aufl.). Bern: Huber.

Weekly, J. A. & Ployhart, R. E. (2006). *Situational Judgment Tests: Theory, Measurement and Application.* London: Lawrence Erlbaum Associates.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim: Beltz.

Williams, T. R. (2005). Exploring the impact of study abroad on students' intercultural communication skills: Adaptability and sensitivity. *Journal of Studies in International Education, 9*, 356–371.

3. Anhang

Tabelle App-A: Mehrgruppen-Vergleiche für die Faktoren und das Außenkriterium Teilnahme an einem interkulturellen Training

	N = 641; *Training* = 189, *Kein Training* = 462			
	Chi-Quadrat-Differenztest	Untergruppen	*M*	*SE*
BAF	$\chi^2(1) = 44.99$, $p < .001$; $d = .51$	Training	0.55	0.08
		Kein Training	0.00	0.00
GSI	$\chi^2(1) = 14.77$, $p < .001$; $d = .33$	Training	0.31	0.07
		Kein Training	0.00	0.00
AB	$\chi^2(1) = 15.62$, $p < .001$; $d = .33$	Training	0.28	0.07
		Kein Training	0.00	0.00
IG	$\chi^2(1) = 22.02$, $p < .001$; $d = .41$	Training	0.50	0.09
		Kein Training	0.00	0.00
FK	$\chi^2(1) = 28.20$, $p < .001$; $d = .40$	Training	0.27	0.05
		Kein Training	0.00	0.00
KK	$\chi^2(1) = 13.29$, $p < .001$; $d = .28$	Training	0.23	0.06
		Kein Training	0.00	0.00
MI	$\chi^2(1) = 11.21$, $p < .001$; $d = .25$	Training	0.24	0.07
		Kein Training	0.00	0.00
RK	$\chi^2(1) = 36.48$, $p < .001$; $d = .51$	Training	0.70	0.10
		Kein Training	0.00	0.00
ABN	$\chi^2(1) = 49.31$, $p < .001$; $d = .49$	Training	0.60	0.09
		Kein Training	0.00	0.00
LB	$\chi^2(1) = 25.61$, $p < .001$; $d = .44$	Training	0.36	0.06
		Kein Training	0.00	0.00
APK	$\chi^2(1) = 28.56$, $p < .001$; $d = .40$	Training	0.44	0.08
		Kein Training	0.00	0.00
BK	$\chi^2(1) = 41.27$, $p < .001$; $d = .55$	Training	0.60	0.08
		Kein Training	0.00	0.00
PZ	$\chi^2(1) = 7.28$, $p < .001$; $d = .18$	Training	0.15	0.06
		Kein Training	0.00	0.00

Anmerkungen: BAF = Bereitschaft zur Anwendung einer Fremdsprache, GSI = Gezieltes Sammeln von Informationen, AB = Aufbau von Beziehungen, IG = Integration in Gruppen, FK = Flexibilität in der Kommunikation, KK = Klarheit in der Kommunikation, MI = Mediation unterschiedlicher Interessen, RK = Reflektion der eigenen Kultur, ABN = Aufbau eines beruflichen Netzwerks, LB = Lernbereitschaft, APK = Aufbau und Pflege von Kontakten, BK = Bewusstsein der eigenen Kultur und PZ = Produktive Zusammenarbeit ermöglichen.

Tabelle App-B: Mehrgruppen-Vergleiche für die Faktoren und das Außenkriterium Interkulturelle Involviertheit

N = 641; *Interkulturelle Involviertheit* = 520, *Keine interkulturelle Involviertheit* = 121

	Chi-Quadrat-Differenztest	Untergruppen	*M*	*SE*
BAF	$\chi^2(1) = 43.87$, $p < .001$; $d = .36$	Interkulturelle Involviertheit	0.73	0.09
		Keine interkulturelle Involviertheit	0.00	0.00
GSI	$\chi^2(1) = 31.57$, $p < .001$; $d = .21$	Interkulturelle Involviertheit	0.44	0.09
		Keine interkulturelle Involviertheit	0.00	0.00
EK	$\chi^2(1) = 27.99$, $p < .001$; $d = .21$	Interkulturelle Involviertheit	0.38	0.08
		Keine interkulturelle Involviertheit	0.00	0.00
AB	$\chi^2(1) = 15.19$, $p < .001$; $d = .18$	Interkulturelle Involviertheit	0.33	0.08
		Keine interkulturelle Involviertheit	0.00	0.00
IG	$\chi^2(1) = 39.21$, $p < .001$; $d = .24$	Interkulturelle Involviertheit	0.72	0.13
		Keine interkulturelle Involviertheit	0.00	0.00
PK	$\chi^2(1) = 14.58$, $p < .001$; $d = .20$	Interkulturelle Involviertheit	0.32	0.07
		Keine interkulturelle Involviertheit	0.00	0.00
KK	$\chi^2(1) = 22.62$, $p < .001$; $d = .17$	Interkulturelle Involviertheit	0.31	0.08
		Keine interkulturelle Involviertheit	0.00	0.00
MI	$\chi^2(1) = 50.33$, $p < .001$; $d = .25$	Interkulturelle Involviertheit	0.45	0.08
		Keine interkulturelle Involviertheit	0.00	0.00
RK	$\chi^2(1) = 67.73$, $p < .001$; $d = .39$	Interkulturelle Involviertheit	0.97	0.11
		Keine interkulturelle Involviertheit	0.00	0.00
ABN	$\chi^2(1) = 27.29$, $p < .001$; $d = .20$	Interkulturelle Involviertheit	0.54	0.12
		Keine interkulturelle Involviertheit	0.00	0.00
LB	$\chi^2(1) = 31.27$, $p < .001$; $d = .22$	Interkulturelle Involviertheit	0.40	0.08
		Keine interkulturelle Involviertheit	0.00	0.00
APK	$\chi^2(1) = 49.49$, $p < .001$; $d = .24$	Interkulturelle Involviertheit	0.55	0.10
		Keine interkulturelle Involviertheit	0.00	0.00
BK	$\chi^2(1) = 67.44$, $p < .001$; $d = .27$	Interkulturelle Involviertheit	0.56	0.09
		Keine interkulturelle Involviertheit	0.00	0.00

Anmerkung: BAF = Bereitschaft zur Anwendung einer Fremdsprache, GSI = Gezieltes Sammeln von Informationen, EK = Empathie in der Kommunikation, AB = Aufbau von Beziehungen, IG = Integration in Gruppen, PK = Perspektivenwechsel in der Kommunikation, KK = Klarheit in der Kommunikation, MI = Mediation unterschiedlicher Interessen, RK = Reflektion der eigenen Kultur, ABN = Aufbau eines beruflichen Netzwerks, LB = Lernbereitschaft, APK = Aufbau und Pflege von Kontakten und BK = Bewusstsein der eigenen Kultur.

4. Tabellen inkl. Tabellentitel

Tabelle 1. Vergleich einiger Merkmale direkter Verfahren zur Messung interkultureller Kompetenz

Merkmale	BASIC	IDI	CCAI	ICSI	Prospector	ISI	INCA
N	236	4763	653	139	838	52	50
Probanden	Studierende	Normal-bevölkerung /Studierende	Normal-bevölkerung /Studierende	Studierende	Angestellte mit und ohne Führungs-verantwortung	Universitäts-angestellte	International tätige Ehrenamtliche
Anzahl der Dimensionen (Anzahl der Items)	1 (9)	5 (50)	4 (50)	4 (46)	14 (116)	9 (–)	4 (211)
Anzahl der Methoden (Art der Methode)	1 (Fremd-einschätzung)	1 (Selbst-einschätzung)	1 (Selbst-einschätzung)	1 (Selbst-einschätzung)	1 (Selbst-einschätzung)	1 (Selbst-einschätzung)	3 (Biografischer Fragebogen, Rollenspiel, Szenario)
Mittlere interne Konsistenz α	.82	.80–.85	.68–.90	.83	.70–.92	–	.82–.89

Anmerkungen: – keine Information; BASIC = Behavioral Assessment Scale for Intercultural Communication (Koester & Olebe, 1988); IDI = Intercultural Development Inventory (Hammer, 2008; Hammer et al., 2003); CCAI = Cross-Cultural Adaptability Inventory (Kelley & Meyers, 1995); ICSI = Intercultural Sensitivity Inventory (Bhawuk & Brislin, 1992), Prospector (Spreitzer et. al., 1997); ISI = The Intercultural Sensitivity Index (Olson & Kroeger, 2001); INCA = Intercultural Competence Assessment (Fantini & Tirmizi, 2006)

Tabelle 2. Vergleich der Fit-Statistiken für Modelle mit unterschiedlicher Faktorenanzahl

Anzahl der Faktoren	***AIC***	***BIC***	**Sample-Size Adjusted *BIC***	χ^2/df	***RMSEA***	***SRMR***	***CFI***
17	113539.20	119292.05	115199.55	1.58	0.030	0.017	0.961
16	119366.73	125458.76	121124.97	1.61	0.031	0.018	0.955
15	119531.21	125056.44	121125.87	1.72	0.034	0.020	0.944

Tabelle 3. Kennwerte der 17 Kompetenzfacetten und der Gesamtskala

Skalen	***M***	***SD***	**α**	**Guttmans Split-Half-Koeffizient**
BAF	3.77	1.04	.81	.79
EK	4.57	0.69	.89	.88
GSI	4.80	0.84	.86	.75
AB	3.98	0.96	.84	.71
IG	3.59	1.04	.84	.80
PK	4.33	0.75	.69	.62
FK	4.56	0.67	.74	.75
KK	4.22	0.75	.82	.82
LB	4.60	0.64	.81	.65
MI	4.36	0.69	.82	.73
ABN	3.60	0.86	.77	.72
SP	4.57	0.70	.78	.79
APK	4.12	0.79	.80	.75
ZO	4.43	0.73	.86	.80
PZ	4.42	0.57	.74	.63
RK	3.91	1.06	.90	.80
BK	4.10	0.81	.84	.82
TMIK	4.25	0.49	.96	.90

Anmerkungen: N = 641; M = Mittelwert; SD = Standardabweichung; α = Interne Konsistenz; BAF = Bereitschaft zur Anwendung einer Fremdsprache, EK = Empathie in der Kommunikation, GSI = Gezieltes Sammeln von Informationen, AB = Aufbau von Beziehungen, IG = Integration in Gruppen, PK = Perspektivenwechsel in der Kommunikation, FK = Flexibilität in der Kommunikation, KK = Klarheit in der Kommunikation, LB = Lernbereitschaft, MI = Mediation unterschiedlicher Interessen, ABN = Aufbau eines beruflichen Netzwerks, SP = Strategisches Problemlösen, APK = Aufbau und Pflege von Kontakten, ZO = Zielorientierung, PZ = Produktive Zusammenarbeit ermöglichen, RK = Reflektion der eigenen Kultur, BK = Bewusstsein der eigenen Kultur; TMIK = Test zur Messung interkultureller Kompetenz.

Tabelle 4. Korrelationen der Faktoren mit den inhaltlich passenden Situationsbeurteilungsitems

	SJT _BAF	SJT _EK	SJT _GSI	SJT _IG	SJT _PK	SJT _FK	SJT _KK	SJT _LB	SJT _MI	SJT _ABN	SJT _SP	SJT _APK	SJT _ZO	SJT _PZ	SJT _RK	SJT _BK
BAF	.61***	n.s.	.11*	.17***	n.s.	.15**	.10*	.27***	n.s.	.10*	n.s.	.25**	n.s.	n.s.	.17***	.20***
EK	.15***	.11*	.09*	.13**	.20***	n.s.	n.s.	.14**	n.s.	n.s.	n.s.	n.s.	n.s.	.11*	.10*	.11**
GSI	.13**	n.s.	.42***	n.s.	n.s.	n.s.	n.s.	.28***	n.s.	.12**	n.s.	.11*	.11*	n.s.	.16***	n.s.
AB	.45***	n.s.	n.s.	.30***	n.s.	.10*	.13**	.17***	n.s.	n.s.	n.s.	.31***	n.s.	n.s.	.14**	.16***
IG	.37***	n.s.	n.s.	.27***	n.s.	.09*	.14**	.14**	n.s.	n.s.	n.s.	.29***	n.s.	n.s.	.11*	.13**
PK	.12*	n.s.	.11*	n.s.	.40***	n.s.	n.s.	.15**	n.s.	n.s.	n.s.	n.s.	n.s.	.18***	.15**	.10*
FK	.17***	n.s.	.15**	.15**	.19***	n.s.	n.s.	.21***	n.s.	n.s.	n.s.	.10*	n.s.	.14**	n.s.	n.s.
KK	.27***	n.s.	.12*	.17***	n.s.	.15**	.17**	.13**	n.s.	.09*.	n.s.	.14**	n.s.	.13**	.13**	.19***
LB	.26***	n.s.	.17***	.17***	.14**	.13**	n.s.	.33***	n.s.	.24***	n.s.	.23***	n.s.	.17***	.18***	.23***
MI	.24***	n.s.	n.s.	.17***	.21***	n.s.	n.s.	.17***	.08*	n.s.	n.s.	.13**	n.s.	.19***	.18***	.17***
ABN	.32***	n.s.	n.s.	.25***	n.s.	n.s.	n.s.	.19***	n.s.	.16***	n.s.	.20***	n.s.	.11**	.16***	.20***
SP	n.s.	n.s.	n.s.	.12**	.18***	n.s.	n.s.	.15**	n.s.	.11*	.18***	.09*	.10*	.11*	.13**	n.s.
APK	.35***	.09*	n.s.	.25***	n.s.	n.s.	n.s.	.23***	n.s.	.17***	n.s.	.43***	n.s.	.14**	.19***	.20***
ZO	.17***	n.s.	.13**	.11*	.10*	n.s.	n.s.	.17***	n.s.	.11**	n.s.	n.s.	.28***	.09*	.10*	.14***
PZ	.27***	n.s.	.11*	.20***	.12*	.16**	.10*	.22***	n.s.	.14**	n.s.	.12**	n.s.	.21***	n.s.	.11*
RK	.14**	.10*	.12**	.14**	.16***	n.s.	-.09*	.16***	n.s.	n.s.	n.s.	.13**	n.s.	n.s.	.21***	.17***
BK	.25***	n.s.	n.s.	.15**	.11*	n.s.	n.s.	.09*	n.s.	n.s.	n.s.	.14**	n.s.	n.s.	.13**	.28***

Anmerkungen: n.s. = nicht signifikant; * $p < .05$, ** $p < .01$; *** $p < .001$; SJT = Situational Judgment Test; BAF = Bereitschaft zur Anwendung einer Fremdsprache, EK = Empathie in der Kommunikation, GSI = Gezieltes Sammeln von Informationen, AB = Aufbau von Beziehungen, IG = Integration in Gruppen, PK = Perspektivenwechsel in der Kommunikation, FK = Flexibilität in der Kommunikation, KK = Klarheit in der Kommunikation, LB = Lernbereitschaft, MI = Mediation unterschiedlicher Interessen, ABN = Aufbau eines beruflichen Netzwerks, SP = Strategisches Problemlösen, APK = Aufbau und Pflege von Kontakten, ZO = Zielorientierung, PZ = Produktive Zusammenarbeit ermöglichen, RK = Reflektion der eigenen Kultur und BK = Bewusstsein der eigenen Kultur

Tabelle 5. Korrelation der Faktoren mit einem globalen Maß interkultureller Kompetenz

Faktoren	**Globales Maß interkultureller Kompetenz**
Bereitschaft zur Anwendung einer Fremdsprache	.49
Empathie in der Kommunikation	.27
Gezieltes Sammeln von Informationen	.38
Aufbau vertrauensvoller Beziehungen	.44
Integration in Gruppen	.31
Perspektivenwechsel in der Kommunikation	.25
Flexibilität in der Kommunikation	.28
Klarheit in der Kommunikation	.38
Lernbereitschaft	.40
Mediation unterschiedlicher Interessen	.42
Aufbau eines beruflichen Netzwerks	.38
Strategisches Problemlösen	.20
Kontakte knüpfen und aufrecht erhalten	.43
Zielorientierung	.22
Produktive Zusammenarbeit ermöglichen	.30
Reflektion der eigenen Kultur	.49
Bewusstsein der eigenen Kultur	.57

Anmerkung: Alle Korrelationen sind auf einem Niveau von $p < .001$ signifikant.

Appendix C

Pre-print Version of Paper B

Appendix C provides a pre-print version of Paper B: Schnabel, D., Kelava, A., & Van de Vijver, F. J. R. (in press). The effects of using collaborative assessment with students going abroad: Intercultural competence development, self-understanding, self-confidence, and stages of change. *Journal of College Student Development.*

Pre-print refers to the originally in 2013 submitted version of the paper before any peer-review. Please note, that there are consequently substantial differences between the pre-print and the published versions.

The effects of using collaborative assessment with students going abroad: Intercultural competence development, self-understanding, self-confidence, and stages of change.

Abstract

In this study collaborative assessment (Finn, 1996, 2007; Fischer, 1994) was examined in counseling N = 820 German students who were going abroad and who were exposed to the Test to Measure Intercultural Competence (TMIC; Schnabel, Kelava, Seifert, & Huber, in press). A randomized pretest-posttest control group design was used. The control group did not get any test feedback. The remaining groups received only written feedback or a collaborative assessment intervention (SHORT). Repeated measures linear mixed modeling showed that participating in SHORT positively influences students' self-appraisal of their intercultural competence (TMIC-SA), their values on three stages of change, as well as their self-understanding, self-confidence, and perceived benefit from test participation.

Keywords: Collaborative assessment, test feedback, intercultural competence, Test to Measure Intercultural Competence, student counseling, student development.

Nowadays, global mobility of students is a common practice in most parts of the world. However, establishing development programs for students in higher education is still at an early stage (Straub, Nothnagel, & Weidemann, 2010). The most widespread development activities in the field are culture-specific or cross-cultural awareness group trainings that are offered to students before going abroad (Landis & Bhagat, 1996). These trainings are rather generic with little focus on the needs, strengths, or weaknesses of the individual (Mendenhall, Stahl, Ehnert, Oddou, Osland, & Kühlmann, 2004). Personal coaching (or counseling) might be more fruitful, though less chosen in higher education due to its time and cost intensity (Vulpe, 2004). This shows that there is a great need for personalized, yet economic interventions in intercultural competence development of students. Although using psychological tests in intercultural competence development is a well-known practice with Expatriates (for an overview see Deardorff, 2010), to our knowledge no empirically grounded guidelines exist to date, which enable counselors to meaningfully interpret and communicate personal results from intercultural competence assessments to foster students' development.

Therefore, this article focuses on the personal benefit after filling in the Test to Measure Intercultural Competence (TMIC; Schnabel, Kelava, Seifert, & Hubert, in press) and receiving no versus written versus written plus oral collaborative test feedback as a brief intervention. In the following, we will 1) introduce the main topics intercultural competence and collaborative assessment, 2) outline the newly adapted collaborative test feedback intervention SHORT, and 3) present results of the development effects of collaborative test feedback in a nonclinical setting.

Intercultural Competence

Intercultural competence consists of several facets, which enable a person to successfully face unknown challenges while living, working, or studying in different cultures (e.g., Bolten, 2007; Earley & Ang, 2003; Fantini & Tirmizi, 2006; Schnabel et al., in press). These competence facets can be developed and are thus not as stable and determined as cross-

cultural relevant personality traits are (cf. Fantini & Tirmizi, 2006; Kelley & Meyers, 1995; Koester & Olebe, 1988; Ruben 1976). Schnabel et al. (in press) showed that intercultural training experiences, private or professional involvement with an intercultural topic and longer stays abroad positively influence the level of intercultural competence in students. This indicates that individuals who operate in a diverse environment can and should consciously take actions to develop their intercultural competence. The most common development method is taking part in an intercultural training (e.g., Brislin & Bhawuk, 1999; Landis & Bhagat, 1996; Triandis, 1995). Culture-specific or cross-cultural trainings aim at preparing individuals for successfully handling intercultural situations (Earley, 1987). Hereby, they concentrate mainly on building knowledge (Mendenhall et al., 2004). Intercultural coaching then focuses on how this knowledge can be applied in the context of the coachee. The needs and development issues of the coachee are thus central (Rosinski & Abbott, 2006). In practice assessment instruments like the Intercultural Preference Tool (ICUnet.AG, 2008) are mostly used as part of the training or coaching, but not as a developmental intervention per se.

A wide range of instruments in the field of intercultural psychology has been developed so far, which differ greatly in their psychometric qualities and operationalization of the construct (Gabrenya, Griffith, Moukarzel, Pomerance, & Reid, 2012). Established instruments either measure personality traits (e.g., Multicultural Personality Questionnaire; Van Oudenhoven & Van der Zee, 2002), intercultural sensitivity (e.g., Intercultural Development Inventory; Hammer, Bennett, & Wiseman, 2003), or cultural intelligence (e.g., Cultural Intelligence Scale; Van Dyne, Ang, & Koh, 2008). As popular the concept intercultural competence theoretically is in literature as large is the gap for instruments, which actually measure intercultural competence. Schnabel et al. (in press) contribute to closing that gap with their recently developed German Test to Measure Intercultural Competence (TMIC). It assesses 17 competence facets, which belong to six second-order factors (communication, learning, social interaction, self-management, self-knowledge, and building synergies). Three

studies showed that TMIC is a well-validated innovative instrument in the field (Schnabel & Kelava, 2013; Schnabel et al., in press; Schnabel, Seifert, & Kelava, 2013), which highlights the potential of TMIC to serve as a basis for a collaborative assessment feedback intervention.

Collaborative Assessment

Collaborative assessment (Fischer, 1994), also called therapeutic assessment (Finn, 1996, 2007), is a highly individualized approach of using psychological tests in counseling. Hereby, test results are interpreted in the light of the personal experiences and situation of the testee who collaboratively works together with the assessor to increase the unique benefit of the assessment process (Finn & Tonsager, 1997; Fischer, 2000). The testees thus become the co-assessors who share their opinions in an open and trustful dialogue (Craddick, 1975). Therefore, collaborative assessment goes beyond pure information gathering and classification modes of assessment (Finn & Tonsager, 1997). Waiswol (1995) stated that therapeutic assessment, with its transformative nature, functions like a brief intervention (Miller & Rollnick, 2002). Three basic human motives regulate the individual's transformation (Finn & Kamphuis, 2006). First, the need for self-verification (Swann & Read, 1981), which is addressed in collaborative assessment through confirmation of the testee's own view of their personal self. Second, the need for self-enhancement (Sedikides & Gregg, 2008), which is fulfilled when the testee feels valued. Third, the need for self-efficacy-self-discovery (Bandura, 1994), which is satisfied when the testee learns novel aspects about themselves and about - so far - unsolved problems (Finn & Kamphuis, 2006).

Research on the impact of collaborative assessment is still emerging (Riddle, Byers, & Grimesey, 2002). Poston and Hanson (2010) realized the strong need for giving an empirical overview on the treatment utility of assessment procedures in counseling and conducted a meta-analysis including 17 studies dealing with psychological assessment as a therapeutic intervention. Sixty-six percent of treatment group means were significantly higher than the means in the control and comparison group; $d = 0.423$, 95% CI [0.321, 0.525]. Furthermore,

moderate treatment group effects have been found for therapy outcomes (d = 0.367, 95% CI [0.256, 0.478]). Early reports of clinicians indicated several positive outcomes after sharing test results with their clients. These included an increased self-esteem, a greater self-awareness and self-understanding, and a higher motivation to seek mental health therapy (Finn & Butcher, 1991). In their first study on therapeutic effects of providing feedback on the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Hathaway & McKinley, 1942), Finn and Tonsager (1992) showed that individuals who received collaborative assessment feedback reported less symptomatic distress as well as an increase of self-esteem and hope for solving their problems in comparison to the control group. These results were also replicated in a second study with Australian students (Newman & Greenway, 1997). When compared with a group that received no feedback at all, Allen, Montgomery, Tubman, Frazier, and Escovar (2003) found a positive effect on self-esteem, self-understanding, self-liking, self-competence, and feedback satisfaction after providing collaborative assessment feedback about the Million Index of Personality Styles (Million, Weiss, Million, & Davis, 1994). According to Finn and Kamphius (2006), the main goal of therapeutic assessment is that individuals reach a better understanding about themselves and their situation.

Concerning the mode of providing assessment feedback, Folds and Gazda (1966) found that test participants who received individual feedback were more satisfied with test interpretation procedure than those getting a written report. Their three feedback groups also significantly changed in the self-concept, but there was no interaction between the degree of change and the type of test interpretation. Holmes (1964) showed that feedback administered via a counselor increased the perceived benefit of the test results in students when compared to written feedback via mail. Interestingly, results suggest that testing without any feedback would be as meaningless as receiving no treatment at all (Poston & Hanson, 2010). In a more recent study Lance and Krishnamurthy (2003) examined the effect of a combined written and oral feedback approach on client satisfaction. According to their results, oral feedback alone

does not lead to a greater satisfaction than written feedback, which makes the combined oral and written feedback most preferable.

Yet unexamined but proposed by Finn and Kamphius (2006) is the relation between collaborative assessment and Prochaska's transtheoretical model of stages of change (TTM; cf. Prochaska, DiClemente, & Norcross, 1992; DiClemente & Prochaska, 1998). TTM is used to explain and predict the intention of an individual to change their behavior. Particularly, Finn and Kamphius (2006) assume that collaborative assessment positively influences the localization of an individual on one of the early three stages of change.

Poston and Hanson (2010) identified the question of how and why collaborative assessment is beneficial for an individual as a key research direction. As previous studies showed, this question can only be answered if 1) different feedback types (e.g., written vs. oral) are compared with each other and 2) effects of these feedback types on relevant outcome variables are examined. In this study we aim to follow this procedure, which will be outlined in our study purpose and in our hypotheses.

Study Purpose

This study specifically adds to the existing research in having the following two main purposes: First, to strengthen the *external validity* of collaborative assessment in highlighting its effects for students' self-related and intercultural competence development in a nonclinical setting. Second, to enrich the utility (cf. Hayes, Nelson, & Jarrett, 1987; Meyer et al., 2001) of (intercultural) personality and competence assessments in student counseling, coaching, and training practice. Hereby, we believe that the way feedback is framed, that is the mode of providing feedback, will strongly influence the individual's decision making-processes, which are crucial for any further development (Pope, 1992). In accordance with previous findings reviewed above the following hypotheses should be tested:

Hypothesis 1: In comparison to no or written feedback alone, combined written and oral collaborative assessment feedback positively influences the self-appraisal of ones intercultural competence.

Hypothesis 2a-c: Students receiving combined written and oral collaborative assessment feedback will show a better self-understanding, a higher self-confidence, and a greater perceived benefit of taking part in the TMIC after the treatment than individuals without any or with written feedback.

Hypothesis 3a-c: Combined written and oral collaborative assessment feedback will have a positive effect on an individual's stage of change. On average, testees in the intervention group will have higher means on the contemplation and action stage and lower means on the precontemplation stage (Prochaska & DiClemente, 1982) following the feedback than participants in the comparison or control group.

Hypothesis 4a-b: Students in the collaborative assessment feedback group report higher satisfaction with the feedback given than testees in the written feedback group. The more satisfied participants are with the collaborative feedback session the higher is their gain in intercultural competence after the session.

Method

Sampling Procedures

All study participants were recruited from a German technical university. Hereby, we cooperated with the department for International and External Affairs, which organizes in- and outbound activities of the university as well as intercultural development activities. The sample consisted of students who were already accepted for a year-abroad-program in Europe, Asia, or America, which would start approximately four months after we began with the study (April 2013). Out of all outgoing students, we randomly selected individuals for our three experimental groups. At T0 $n = 351$ students took part who would receive written feedback only after the post-test, $n = 396$ filled in the survey in the written feedback condition, and $n =$

73 students participated in the treatment group. When comparing the number of participants between T0 and T1 (see Table 1), the following dropout rates are to be derived: 56% (n_1), 89% (n_2), and 43% (n_3). Outgoing students are a very attractive sample for different kinds of studies. Therefore, these students receive a high number of invitations to various research projects. Together with the fact that the participation was voluntary, this could be a reason for the high drop-off rates in this study.

The invitation to the follow-up survey was sent two weeks after the first participation. Students were given two weeks to fill in the follow-up survey. Two reminders were sent in between. At T0 all participants could sign up for a lottery with the possibility to win one of three Amazon vouchers (two 25 Euro and one 50 Euro voucher) or an Apple Ipod Nano.

Sample Size, Structure, and Power

All together, $N = 820$ students, 480 females and 327 males (13 are missing), with an average age of 23.37 ($SD = 3.89$) years took part in the study at T0. In total, $N = 233$ students participated also in the follow-up study. The 133 men (57%) and 100 women (43%) were in average 23.47 ($SD = 2.89$) years old. Table 1 shows age and sex for all three subgroups as well as results concerning education and intercultural experiences, which were similar across the three groups.

To examine the achieved power, we exemplarily conducted a post-hoc analysis. Effect size f, α-level, and total sample size refer to the repeated measures, within-between interaction ANOVA with the variables TMIC-SA total score (T0 vs. T1) and feedback group (no vs. written vs. oral feedback; see results part). The resulting power ($1-\beta$) = 1.00 was highly satisfactory.

Measures and Covariates

Test to Measure Intercultural Competence (TMIC). The Test to Measure Intercultural Competence (TMIC; Schnabel et al., in press) contains two scales. TMIC-SJT consists of 17 situational judgment items, which assess the individual's behavioral preferences

in work-related critical situations in the international context. TMIC-SJT was not included in this study for two reasons: Firstly, students would have benefited less from the assessment process, if they had been confronted with situations, which are relatively unknown for them. Secondly, completing the whole TMIC is very time-consuming, which – in combination with the feedback – could have probably led to increased drop-out rates.

TMIC-SA measures with 75 self-appraisal items (six-point Likert scale from „does not apply at all“ to „fully applies“) an individual's self-concept concerning the before mentioned six second-order factors of intercultural competence. In the following we will provide an example item for each of the six second-order factors: “The way I address something depends on the person I am talking to.”[1] (Communication), “When planning a trip abroad I use various sources of information.” (Learning), “When I join a group for the first time I quickly build relationships with other group members.” (Social interaction), “When I plan something I usually then go on to achieve my aim.” (Self-management), “I am good at mediating between people with conflicting interests.” (Creating synergies), and “I make an effort to understand to what extent my behavior is shaped by culture.” (Self-knowledge).

Exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) led to a very good model fit for the 17 competence facets: $\chi^2(1636) = 2579.85$, $p < .001$; $\chi^2/df = 1.58$; RMSEA = .031 (RMSEA 90% CI = .029–033); SRMR = .017; CFI = .955; TLI = .927. The six-second order factor structure was supported by a confirmatory factor analysis (CFA): $\chi^2(3987) = 8280.09$, $p < .001$; $\chi^2/df = 2.08$; RMSEA = .040 (RMSEA 90% CI = .040–.046); SRMR = .076; CFI = .820. Cronbach's alpha of the TMIC-SA indicated a very high internal consistency of the scale ($\alpha = .96$; Schnabel et al., in press). Additionally, factorial invariance was established for a Portuguese (Brazil) version of the TMIC (Schnabel, Kelava, & Seifert, 2013).

[1] All example items in the following were translated from German to English for the purpose of this article. However, these English items are not validated.

Stages of change. In the early versions of measuring the stages of change in the transtheoretical model (cf. Prochaska et al., 1992), DiClemente and Prochaska (1998) proposed an approach, in which each individual was exclusively categorized to one of six stages of change. At the same time, the authors themselves showed through several studies that a) only four of the proposed six stages of change were found in a factor analysis and b) these factors were correlated (Abellanas & McLellan, 1993; DiClemente & Hughes, 1990; McConnaughy, Prochaska, & Velicer, 1983). A distinct categorization of an individual to one of the four stages of change is thus not recommended (Fecht, Heidenreich, Hoyer, Lauterbach, & Schneider, 1998). There are only few German questionnaires, which measure the transtheoretical model. Moreover, all of them are related to drug or alcohol abuse. Therefore, items had to be adapted in context. The original items came from Fecht et al. (1998), Hoyer, Heidenreich, Fecht, Lauterbach, and Schneider (2003), and Hannöver, Rumpf, Meyer, Hapke, and John (2001). As stage four, maintenance, was not possible to attain through the collaborative feedback intervention, only the following three stages were assessed: Precontemplation (e.g., “I guess I have weaknesses in the area of intercultural competence, but there is nothing that I really have to change about it.”), contemplation (e.g., “I should inform myself about how I could possibly foster my intercultural competence.”), and action (e.g., “I work hard on changing myself.”). The four items per stage were measured on a six-point Likert scale.

Treatment benefit variables. The procedure of previous therapeutic assessment evaluation studies (e.g., Allen, et al., 2003; Finn & Tonsager, 1992; Newman & Greenway, 1997**)** was used as an orientation for the assessment of the treatment benefit variables in this study. Because of the specific focus on intercultural competence, we adapted items from the well-established Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1979), from the Self-Liking/Self-Competence Scale (SLCS-R; Tafarodi & Swann, 1995, 2001), and from a subscale of the Assessment Questionnaire-2 (Finn & Tonsager, 1992, 1997), the New Self-awareness/Understanding. Self-understanding was thus measured with six items (e.g., “I know

what is important when interacting with people from other cultures."), self-confidence with four items (e.g., "I feel competent enough to deal with problems, which are arising from working together with people from other cultures."), and perceived benefit from taking part in the test with five items (e.g., "In taking part in the present study, I learned something about myself."). Like in earlier studies, items concerning the degree of satisfaction with the feedback were developed separately and exclusively for our purpose. General satisfaction was measured in the written and oral collaborative assessment group with seven questions (e.g., "Now, that I got feedback concerning my intercultural competences, I know, which steps I have to take in order to improve."). The scale related to the specific satisfaction with the collaborative test feedback contained five items (e.g., "My true self was well reflected in the feedback talk."). For all items a six-point Likert scale was used. Psychometric properties of each scale at T0 are presented in Table 2. The German items are available from the first author.

Research Design and Experimental Intervention

The study is based on a multivariate randomized pretest-posttest control group design with a treatment, a comparison, and a control group. The control group did not get any test feedback or other treatment. The other groups received only written feedback or written plus oral collaborative assessment feedback (SHORT). Variables of interest are the total score of intercultural competence measured with the TMIC Self-appraisal scale (TMIC-SA; Schnabel et al., in press) as well as students' stages of change, self-understanding, self-confidence, perceived benefit from test participation, and satisfaction with the assessment and feedback process (e.g., Allen et al., 2003; Finn & Tonsager, 1992; Hannöver et al., 2001). Table 3 gives an overview about groups, variables, and the pretest-/posttest-design.

The written feedback report was administered immediately after a person had finished the survey. A brief description of all 17 TMIC facets was presented along with the personal score and an according categorization in comparison with another group of 641 German university students (Schnabel et al., in press). Thus, each participant could read from the report,

weather their values in the 17 competence facets and in the overall TMIC-SA scale was below, on, or above average. All students could save and keep the report.

The oral collaborative assessment session SHORT was held on the phone. The same assessor was deployed for all feedback sessions. She followed a highly structured feedback guideline and was instructed to use a language free of jargon (Finn, 1996; Mosak & Gushurst, 1972). The maximum duration of a session was one hour. A graphical representation of the results was sent to all participants in preparation for the phone call and was used during the session to discuss results. The collaborative feedback guideline was partly adapted from the manual for using the MMPI-2 as a therapeutic intervention (Finn, 1996; Fischer & Finn, 2008) and contained important aspects of motivational interviewing techniques (Miller & Rollnick, 2002). Moreover, general recommendations for providing psychological test feedback were respected (e.g., American Psychological Association, 1986, 2010). SHORT is an acronym and stands for the key steps (cf. Poston & Hanson, 2010; Waiswol, 1995) in the collaborative feedback session, which will be described in detail in the following.

Start and how we will proceed. The first step is to make sure that the general framework is established for conducting the collaborative feedback session. Confidentiality has to be assured to the assessee who should sit in a quite place without possible disturbances. Then, the assessor gives an overview about the session content and process as well as encourages any kind of questions or remarks during the session.

Orientation. Firstly, the assessor has to clarify the roles in the session. This also emphasizes the collaborative nature of the feedback. Hereby, the assessor is the expert of the TMIC, whereas the assessees are the experts in themselves. As such they are motivated to help the assessor in correctly interpreting the TMIC results. In thanking the assessee for the willingness to share thoughts and in communicating the personal gain of knowledge through participation in the session, the assessor creates an equal relationship (or partnership) with the assessee. Secondly, the assessor asks the assessee to share previous experiences with

psychological tests. Negative incidents should be treated with a high degree of empathy. Before filling in the TMIC, participants in the *collaborative feedback* group had the chance to formulate questions about what they want to learn from the TMIC results. The assessor reads these questions out loud in the feedback session and asks for further ones. Additionally, previous experiences abroad and future international plans are to be discussed. Finally, information concerning the purpose and content of the TMIC along with the results graphic are explained to the assessee.

Reflection. The assessees are asked to take some time to review their results. They are then prompted to explain, if and how they find themselves fairly well represented in the profile. They should also name results, which are surprising or seemingly unrelated to their self-image. The assessor takes notes. Greatest therapeutic effects are expected when test feedback is ordered according to the client's existing self-concept (Finn, 1996; Schroeder, Hahn, Finn, & Swann, 1993). Therefore, results are discussed in the following order: 1) high values in competence facets with high acceptance of assessee, 2) low values in competence facets with high acceptance of assessee, 3) high and low values in competence facets with low acceptance of assessee. Generally, the assessor has to ask for real-life examples concerning each facet discussed and take reference to the individual assessment questions formulated by as well as the background of the assessee. If an assessee does not accept a result, there are several possibilities how to deal with that (for extensive recommendations see Finn, 1996). Examples would be to ask how the result has to look like in the opinion of the assessee, to communicate that results might be wrong, or to formulate a take-home message.

Targets. In the last step remaining questions are answered. As adding goal setting to feedback enhances performance after the feedback (Balcazar, Hopkins, & Suarez, 1986) the assessee then has to create two goals they want to attain in the future. These goals should be related to the discussed results in the TMIC. The collaborative feedback session closes with communicating the possibility to give feedback about the session and to be contacted in the

future. Also, the assessor should thank the assessee for their openness and reassure confidentiality.

Results

Internal Consistency of the Scales

The test-retest reliability (after two weeks) in the *no feedback* group was high ($r = .89$, $p = .000$). As shown in Table 2 Cronbach's Alpha led for all used scales to satisfactory internal consistencies.

Total Score of the TMIC-SA

Taking the following multiple mean comparisons into consideration, the authors decided to employ a Bonferroni-Holm correction (Holm, 1979) on the whole results part. Accordingly, results are handled as significant as soon as they reach a $p \leq .01$ level. All means and standard deviations can be found in Table 2.

An analysis of variance revealed no significant difference in the TMIC-SA total score between the three experimental groups at T0 ($F(2, 815) = 1.03$, $MSE = 0.23$, $p = .357$). The repeated measures linear mixed model resulted in a significant change of the TMIC-SA total score over time ($F(1, 228) = 3.92$, $MSE = 0.81$, $p = .021$, $\eta^2 = .03$). As it is shown in the following strong interaction effect, this change was dependent from the type of feedback a person received ($F(1, 228) = 17.65$, $MSE = 0.49$, $p = .000$, $\eta^2 = .14$). Whereas no significant difference in the TMIC-SA total score across time was found for individuals in the *written feedback* group ($\Delta M = 0.003$, 95% CI [-0.012, 0.006]), TMIC-SA total score significantly increased for the *collaborative assessment feedback* group; $\Delta M = 0.224$, 95% CI [0.215, 0.234], which supports hypothesis 1. Additionally, the TMIC-SA total score slightly decreased in the *no feedback* group across time: $\Delta M = -0.045$, 95% CI [-0.05, -0.041].

Moreover, there was a significant effect of *feedback group* on intercultural competence at T1 ($F(2, 227) = 8.05$, $MSE = 1.89$, $p = .000$, $\eta^2 = .07$). Group differences were assessed by the help of a-priori defined contrasts, which showed that the intervention group differed from

both the *no feedback* as well as the *written feedback* group at T1 ($t(227) = -3.79$, $p = .000$, $d = -.50$).

Therapeutic Benefit Variables

At starting point T0 there were no differences between the three experimental groups concerning perceived benefit from TMIC participation ($F(2, 803) = 2.29$, $MSE = 2.29$, $p = .057$), self-confidence ($F(2, 809) = 0.84$, $MSE = 0.46$, $p = .434$), and self-understanding ($F(2, 809) = 0.72$, $MSE = 0.50$, $p = .488$).

Three repeated measures linear mixed models examined the therapeutic effects of a TMIC assessment in relation to the received feedback. All three therapeutic benefit variables significantly changed over time. For perceived benefit from TMIC participation ($F(1, 226) = 21.82$, $MSE = 7.52$, $p = .000$, $\eta^2 = .09$) and self-confidence ($F(1, 226) = 10.77$, $MSE = 1.47$, $p = .001$, $\eta^2 = .05$) the effect was moderate and for self-understanding with $F(1, 226) = 100.29$, $MSE = 19.10$, $p = .000$, $\eta^2 = .31$ very large. Additionally, we found three respectable interaction effects, which illustrate that all three therapeutic effect variables differed across the experimental groups (see Table 4). Participants who received no or collaborative assessment feedback perceived the benefit after taking part in the TMIC a second time as higher than at the first time. Moreover, the perceived benefit of taking part in the assessment process decreased after individuals got written feedback. Self-confidence beliefs were lower at T1 for the *no feedback* group and higher for the *written feedback* and *collaborative assessment feedback* group. The largest interaction effect was found for self-understanding with individuals having the highest increase in self-understanding after collaborative assessment feedback. Self-understanding also grew as a consequence of written feedback. Lower values were found at T1, if no feedback at all was provided.

At T1 the three feedback groups significantly differed in their perceived benefit from participating in the TMIC ($F(2, 225) = 23.36$, $MSE = 22.24$, $p = .000$, $\eta^2 = .17$), in their self-confidence ($F(2, 225) = 5.28$, $MSE = 2.64$, $p = .006$, $\eta^2 = .05$), and in their self-understanding

($F(2, 225) = 30.01$, $MSE = 19.28$, $p = .000$, $\eta^2 = .21$). As predicted the therapeutic effect was highest for individuals who took part in a collaborative feedback session as could be shown by the following contrasts: a) $t(225) = -6.80$, $p = .000$, $d = -.91$ (perceived benefit from TMIC), b) $t(225) = -3.25$, $p = .001$, $d = -.43$ (self-confidence), and c) $t(225) = -7.29$, $p = .000$, $d = -.97$ (self-understanding).

Stages of Change

Firstly, differences between the experimental groups concerning all three stages of change were examined at T0. This revealed similar starting values in the precontemplation stage across groups ($F(2, 809) = 0.27$, $MSE = 0.19$, $p = .765$), but different ones in the contemplation ($F(2, 809) = 16.48$, $MSE = 12.83$, $p = .000$, $\eta^2 = .04$) and in the action stage ($F(2, 809) = 14.51$, $MSE = 11.62$, $p = .000$, $\eta^2 = .04$). For group comparisons Hochberg's GT2 post-hoc test was chosen, which led to three subgroups for action that all significantly ($p \leq .01$) differed from each other and two for contemplation. In the latter stage significant differences were only attained between the *collaborative feedback* and the *written feedback* as well as the *no feedback* group ($p = .000$).

Secondly, repeated measures linear mixed models were computed for all three stages of change. In the precontemplation stage only the interaction effect between time and feedback type was found to be significant (see Table 4). Whereas participants in the *no feedback* and in the *written feedback* group scored higher on the precontemplation stage at T1, values for the *collaborative assessment feedback* decreased at T1, which is in line with hypothesis 3a. Furthermore, values of contemplation significantly increased over time ($F(1, 226) = 6.40$, $MSE = 1.54$, $p = .012$, $\eta^2 = .03$; $\Delta M = 0.147$; 95% CI [0.139, 0.157]). However, the interaction effect (time x feedback group) did not hold for the Bonferroni-Holm correction (see Table 4). Concerning the action stage, a significant main ($F(1, 226) = 31.09$, $MSE = 7.73$, $p = .000$, $\eta^2 = .12$) and interaction effect resulted. As shown in Table 4, individuals who received written or

collaborative assessment feedback reached higher values at T1 whereas survey participants without any feedback had lower scores in action at T1, which partly supports hypothesis 3c.

As mentioned before, only for the precontemplation stage values were similar at T0 across groups. Therefore, mean comparisons at T1 were only computed for this stage. The experimental groups differed significantly in their precontemplation at T1 ($F(2, 224) = 10.77$, $MSE = 7.05$, $p = .000$, $\eta^2 = .09$). The contrast $t(224) = 4.47$, $p = .000$, $d = .60$ shows that the lowest mean was found for the *collaborative feedback* group when compared with the *written feedback* and the *no feedback* group.

Evaluation of the Collaborative Assessment Feedback

To examine if the evaluation of the feedback increases after participation in the collaborative assessment feedback session, a paired *t*-test was conducted. At T0 after participants received a written feedback report, the average evaluation was $M = 3.78$, $SD = 0.57$. After the collaborative feedback session, the evaluation significantly ($t(32) = -9.97$, $p = .000$, $d = -3.52$) increased on $M = 4.97$, $SD = 0.70$. Furthermore, the average rating of the collaborative feedback session was high $M = 5.24$, $SD = 0.82$. The overall evaluation of the feedback significantly correlated with the specific rating of the feedback talk ($r = .82$, $p = .000$). Additionally, the more satisfied individuals were with the general and specific feedback, the higher was their TMIC-SA total score at T1 ($r = .43$, $p = .013$; $r = .38$, $p = .029$). Moreover, the evaluation of the collaborative feedback session was highly related to the perceived benefit of taking part in the TMIC ($r = .86$, $p = .000$), moderately to the self-confidence at T1 ($r = .35$, $p = .047$), and not significantly to the self-understanding in T1 ($r = .33$, $p = .060$).

Discussion

The present article contributes to the research on intercultural competence development of students in higher education in showing that collaborative assessment feedback might be applied as a personalized, yet economic intervention. Additionally, we could show that the purpose of instruments in the field of Intercultural, Industrial, and Organizational Psychology

goes beyond selecting and classifying individuals (Finn & Tonsager, 1997). The utility of such instruments increases enormously, as they can function as a brief intervention per se (Hayes et al., 1987; Miller & Rollnick, 2002; Waiswol, 1995). Such an intervention has been developed for this study. SHORT is a one-hour telephone feedback session, which follows the requirements of collaborative assessment (Fischer, 1994) and contains elements of motivational interviewing (Miller & Rollnick, 2002). It was applied with university students to prepare them for studying and living abroad. In a randomized pretest-posttest control group design the collaborative feedback intervention SHORT was compared with written and no feedback at all. Variables of interest were the total scores of the self-appraisal scale of the Test to Measure Intercultural Competence (TMIC-SA; Schnabel et al., 2013), the first three stages of change in the transtheoretical model (Prochaska & DiClemente, 1982), therapeutic benefit as well as general and specific satisfaction with the feedback (e.g., Allen, et al., 2003; Finn & Tonsager, 1992; Newman & Greenway, 1997).

Key Findings

Except for the contemplation stage, all hypotheses were supported. Only for individuals who received combined written and oral collaborative assessment feedback effects were consistent, that is they scored higher on TMIC-SA, all therapeutic benefit variables, and on the action stage at T1. Students' values on the precontemplation stage moreover decreased after they participated in SHORT, which shows that the intervention changes an individual's perception of having no need to change their problematic behavior. Whereas there is a tendency that values slightly decreased across time in the *no feedback* group, effects were somewhat inconsistent with students who got a written feedback report. Written feedback had no effect on TMIC-SA total score and a negative on the perceived benefit of taking part in the test. A potential reason could be that participants cannot envision how to use the feedback personally, if an assessor does not guide them. Self-confidence, self-understanding, and action were positively influenced by written feedback. Interestingly, individuals who received only a

written report about their intercultural competences were even more unwilling to improve their weaknesses at T1 than they were at T0. This shows that written assessment feedback is not necessarily better than receiving no feedback at all, which is in line with previous findings (Lance & Krishnamurthy, 2003). Apparently, written feedback alone even has the potential to harm the development of an individual in some cases. A possible explanation could be derived from attribution theory (Weiner, 1985). We assume that individuals receiving written feedback attribute a negative outcome to external factors like the test characteristics or to stable aptitude factors, which they think are unchangeable anyways. It seems that only individuals who received further explanations through collaborative feedback can fully understand why they achieved a certain result and that they are able to develop, if only they want to.

As hypothesized, satisfaction with the assessment increased after receiving collaborative assessment feedback. Moreover, satisfaction with the feedback session was positively related to the gain of intercultural competence.

From these results we might infer some central goals, which can be attained through SHORT: a) broaden the self-understanding students, b) decrease students' passive acceptance of their weaknesses, and c) increase the self-appraisal of students' intercultural competence.

Limitations and Future Directions

As promising results are, as cautiously they have to be interpreted. This study focuses on the benefits of collaborative assessment for students going abroad. These results cannot simply be generalized to other important populations like adult Expatriates. Also, the gain in intercultural competence is rather subjective. TMIC uses self-appraisal to measure intercultural competences. When we want to know, if there is an objective improvement after participating in SHORT, using 360-degree feedback (Ward, 1997) would be a fruitful approach. Hereby, an individual is evaluated by several counterparts (e.g., their peers, their leaders/professors etc.). Moreover, the way feedback is given and accepted is highly influenced by culture. Especially challenging and interesting would be to conduct SHORT in Asian countries like China where

communication style is rather indirect (Hofstede & Bond, 1984). Also, we still do not know enough about the unique components of collaborative assessment, which might foster or hinder positive effects for individuals. This leaves several questions unanswered: Which effect has the graphic representation of results for example using a radar chart? Is there a difference between collaborative assessment feedback administered in person, on the phone, or through new media? Do effects vary if the assessor is masculine or feminine? As long as these and many other questions are open, no general conclusion in favor of collaborative assessment can be drawn. Nevertheless, results show that, especially when there are time and/or budget restrictions, collaborative feedback might be a way to go beyond pure knowledge-building in intercultural competence development of students and hereby sensitize them for their future encounters abroad.

References

Abellanas, L., & McLellan, A. T. (1993). „Stages of Change“ by drug problem in concurrent opioid, cocaine, and cigarette users. *Journal of Psychoactive Drugs, 25*, 307−313. doi:10.1080/02791072.1993.10472288

Allen, A., Montgomery, M., Tubman, J., Frazier, L., & Escovar, L. (2003). The effects of assessment feedback on rapport-building and self-enhancement processes. *Journal of Mental Health Counseling, 25*, 165−181. Retrieved from: http://essential.metapress.com/content/lw7h84q861dw6ytj/

American Psychological Association (1986). *Guidelines for computer-based tests and interpretations.* Washington, DC: American Psychological Association.

American Psychological Association (2010). *Ethical principals of psychologists and code of conduct.* Washington, DC: American Psychological Association.

Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 397–438. doi:10.1080/10705510903008204

Balcazar, F., Hopkins, B. L., & Suarez, Y. (1986). A critical, objective review of performance feedback. *Applied Psychology*, *66*, 314−323. doi:10.1300/J075v07n03_05

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behaviour* (Vol. 4, pp.71−81). New York, NY: Academic Press.

Bolten, J. (2007). Was heißt Interkulturelle Kompetenz? Perspektiven für die internationale Personalentwicklung [What does intercultural competence mean? Perspectives for international personnel development]. In V. Künzer & J. Berninghausen (Eds.), *Wirtschaft als interkulturelle Herausforderung* [Economy as an intercultural challenge] (pp. 21–42). Frankfurt am Main, Germany: IKO.

Brislin, R. W., & Bhawuk, D. P. S. (1999). Cross-cultural training: Research and innovations. In J. Adamopoulos & Y. Kashima (Eds.), *Social psychology and cultural context* (pp. 205−216). Thousand Oaks, CA: Sage.

Craddick, R. A. (1975). Sharing oneself in the assessment procedure. *Professional Psychology, 6,* 279−282. Retrieved from http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1976-11275-001

Deardorff, D. K. (2010). *The Sage handbook of intercultural competence.* Thousands Oaks, CA: Sage.

DiClemente, C. C., & Hughes, S. O. (1990). Stages of change profile in outpatient alcoholism treatment. *Journal of Substance Abuse, 2*, 217−235. doi:10.1016/S0899-3289(05)80057-4

DiClemente, C. C., & Prochaska, J. O. (1998). Toward a comprehensive, transtheoretical model of change: Stages of change and addictive behaviors. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors* (2nd ed., pp. 3−24). New York, NY: Plenum Press.

Earley, P. C. (1987). Intercultural training for managers: A comparison of documentary and interpersonal methods. *Academy of Management Journal*, *30*, 685−698. Retrieved from http://www.jstor.org/stable/256155

Earley, P. C., & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures.* Palo Alto, CA: Stanford University Press.

Fantini, A., & Tirmizi, A. (2006). Exploring and assessing intercultural competence. *World Learning Publications*, *1,* 1−75. Retrieved from http://digitalcollections.sit.edu/worlslearning_publications/1

Fecht, J., Heidenreich, T., Hoyer, J., Lauterbach, W., & Schneider, R. (1998). Veränderungsstadien bei stationärer Alkoholentwöhungsbehandlung – Probleme der Diagnostik [Stages of change with stationary alcohol withdrawal – Problems in

diagnostics]. *Verhaltenstherapie und Psychosoziale Praxis, 30,* 403–419. Retrieved from http://www.dgvt-verlag.de/

Finn, S. E. (1996). *Manual for using the MMPI-2 as a therapeutic intervention*. Minneapolis, MN: University of Minnesota Press.

Finn, S. E. (2007). *In our clients' shoes: Theory and techniques of therapeutic assessment*. Mahwah, NJ: Earlbaum

Finn, S. E., & Butcher, J. N. (1991). Clinical objective personality assessment. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *The clinical psychology handbook* (2nd ed., pp. 362–373). New York, NY: Pergamon Press.

Finn, S. E., & Kamphuis, J. H. (2006). Therapeutic Assessment with the MMPI-2. In J. N. Butcher (Ed.), *MMPI-2: A practitioners guide* (pp. 165–191). Washington, D.C.: APA Books.

Finn, S. E., & Tonsager, M. E. (1992). Therapeutic effects of providing MMPI-2 test feedback to college students awaiting therapy. *Psychological Assessment, 4*, 278–287. Retrieved from http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1993-04052-001

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment, 9,* 374–385. doi:10.1037/1040-3590 .9.4.374

Fischer, C. T. (1994). *Individualizing psychological assessment.* Mahwah, NJ: Erlbaum.

Fischer, C. T. (2000). Collaborative, individualized assessment, *Journal of Personality Assessment, 74*, 2–14. doi:10.1207/S15327752JPA740102

Fischer, C. T., & Finn, S. E. (2008). Developing the life meaning of psychological test data: Collaborative and therapeutic approaches. In R. P. Archer & S. R. Smith (Eds.), *Personality Assessment* (pp. 379–404). New York, NY: Routledge.

Folds, J. H., & Gazda, G. M. (1966). A comparison of the effectiveness and efficiency of three methods of test interpretation. *Journal of Counseling Psychology, 13*, 318–324. doi:10.1037/h0023708

Gabrenya, W. K., Griffith, R. L., Moukarzel, R. G., Pomerance, M. H., & Reid, P. A. (2012). Theoretical and practical advances in the assessment of cross-cultural competence. In D. D. Schmorrow & D. M. Nicholson (Eds.), *Advances in design for cross-cultural activities* (pp. 317–332). Boca Raton, FL: Taylor Francis.

Hannöver, W., Rumpf, H.-J., Meyer, C., Hapke, U., & John, U. (2001). Der Fragebogen zur Änderungsbereitschaft bei Alkoholkonsum (RCQ-D) [Readiness to Change Questionnaire for Consuming Alcohol (RCQ-D)]. In A. Glöckner-Rist, F. Rist, & H. Küfner (Eds.), *Elektronisches Handbuch zu Erhebungsinstrumenten im Suchtbereich (EHES) [Electronical handbook for assessment instruments for the addictions (EHES)]* (Version 1.00). Mannheim, Germany: Zentrum für Umfragen, Methoden und Analysen (ZUMA).

Hathaway, S. R., & McKinley, J. C. (1942). *Minnesota Multiphasic Personality Inventory.* Minneapolis, MN: University of Minnesota Press.

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations, 27,* 421–443. Retrieved from http://www.sciencedirect.com/science/article/pii/S0147176703000324

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, *42*, 963–974. doi:10.1037/0003-066X.42.11.963

Hofstede, G., & Bond, M. H. (1984). Hofstede's Culture Dimensions. An independent validation using Rokeach's Value Survey. *Journal of Cross-Cultural Psychology, 15*, 417–433. doi:10.1177/0022002184015004003

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70. Retrieved from http://www.jstor.org/stable/4615733

Holmes, J. E. (1964). The presentation of test information to college freshman. *Journal of Counseling Psychology, 11*, 54–58. doi:10.1037/h0046594

Hoyer, J., Heidenreich, T., Fecht, J., Lauterbach, W., & Schneider, R. (2003). Stadien der Veränderung in der stationären Alkoholentwöhnungstherapie [Stages of change with stationary alcohol withdrawal]. *Verhaltenstherapie, 13*, 31–38. Retrieved from http://psylux.psych.tu-dresden.de/i2/klinische/studium/literatur/195m.pdf

ICUnet.AG (2008). *Handbuch zum Intercultural Preference Tool* [*Intercultural Preference Tool manual*]. Göttingen, Germany: Hogrefe.

Kelley, C., & Meyers, J. (1995). *The Cross-Cultural Adaptability Inventory (manual).* Minneapolis, MN: National Computer Systems.

Koester, J., & Olebe, M. (1988). The Behavioral Assessment Scale for Intercultural Communication Effectiveness. *International Journal of Intercultural Relations, 12*, 233–246. doi:10.1016/0147-1767(88)90017-X

Lance, B. R., & Krishnamurthy, R. (2003, March). *A comparison of three modes of MMPI-2 test feedback*. Paper presented at the annual meeting of the Society for Personality Assessment, San Francisco, CA.

Landis, D. E., & Bhagat, R. S. (1996). Handbook of intercultural training. Thousand Oaks, CA: Sage.

McConnaughy, E. A., Prochaska, J. O., & Velicer, W. F. (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research, and Practice, 20*, 368–375. doi:10.1037/h0090198

Mendenhall, M. E., Stahl, G. K., Ehnert, I., Oddou, G., Osland, J. S., & Kühlmann, T. M. (2004). Evaluation studies of cross-cultural training programs. In D. Landis, J. M. Bennett, & M. J. Bennett (Eds.), *Handbook of intercultural training* (pp. 129–143).

Thousand Oaks, CA: Sage.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological Testing and Psychological Assessment. A Review of Evidence and Issues. *American Psychologist, 56*, 128–165. doi:10.1037//003-066X.56.2.128

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing. Preparing people for change*. New York, NY: Guilford Press.

Million, T., Weiss, L., Million, C., & Davis, R. D. (1994). *MIPS: Million Index of Personality Styles manual.* San Antonio, TX: Psychological Corp.

Mosak, H. H., & Gushurst, R. S. (1972). Some therapeutic uses of psychological testing. *American Journal of Psychotherapy, 26,* 539–546. Retrieved from http://psycnet.apa.org/psycinfo/1973-07580-001

Newman, M. L., & Greenway, P. (1997). Therapeutic effects of providing MMPI-2 test feedback to clients at a university counseling service. *Psychological Assessment, 9,* 122–131. doi:10.1037/1040-3590.9.2.122

Pope, K. S. (1992). Responsibilities in providing psychological test feedback to clients, *Psychological Assessment, 4*, 268–271. doi:10.1037/1040-3590.4.3.268

Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment, 22*, 203–212. doi:10.1037/a0018679

Prochaska, J. O., & C. C. DiClemente (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research and Practice*, *19*(3), 276–288. doi:10.1037/h0088437

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change. Applications to addictive behaviors. *American Psychologist, 47*, 1102–1114. doi:10.1037/0003-066X.47.9.1102

Riddle, B. C., Byers, C. C., & Grimesey, J. L. (2002): Literature review of research and practice in collaborative assessment, *The Humanistic Psychologist, 30*, 33–48. doi:10.1080/08873267.2002.9977021

Rosenberg, M. (1979). *Conceiving the self.* New York, NY: Basic Books.

Rosinski, P., & Abbott, G. (2006). Intercultural coaching. In Passmore, J. (Ed.), *Excellence in coaching. The industry guide* (pp. 153–169). London, UK: Kogan Page.

Ruben, B. D. (1976). Assessing communication competency for intercultural adaptation. *Group & Organization Studies, 1*, 334–354. doi:10.1177/105960117600100308

Schnabel, D. B. L., & Kelava, A. (2013, July). *Development and validation of the German Test to Measure Intercultural Competence (TMIC): A combined-method approach.* Paper presented at the 12th European Conference on Psychological Assessment, San Sebastian, Spain.

Schnabel, D., Kelava, A., Seifert, L., & Huber, B. (in press). *Konstruktion und Validierung eines multimethodalen berufsbezogenen Tests zur Messung interkultureller Kompetenz* [Construction and validation of an occupational multimethod test to measure intercultural competence]. *Diagnostica*.

Schnabel, D. B. L., Seifert, L., & Kelava, A. (2013). *How generalizable is intercultural competence across cultures? Examining psychometric properties, measurement invariance, and construct validity of the Test to Measure Intercultural Competence (TMIC) in Germany and Brazil.* Manuscript submitted for publication.

Schroeder, D. G., Hahn, E. D., Finn, S. E., & Swann, W. B., Jr. (1993, June). *Personality feedback has more impact when mildly discrepant form self views.* Paper presented at the fifth annual convention of the American Psychological Society, Chicago, IL.

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science, 3*, 102–116. doi:10.1111/j.1745-6916.2008.00068.x

Straub, J., Nothnagel, S., & Weidemann, A. (2010). *Wie lernt man interkulturelle Kompetenz? Theorien, Methoden und Praxis in der Hochschulbildung. Ein Handbuch*. [*How can intercultural competence be learned? Theories, methods, and practice in higher education. A handbook*]. Bielefeld, Germany: transcript.

Swann, W. B., Jr., & Read, S. J. (1981). Self-verification processes: How we sustain our self-conceptions. *Journal of Experimental Social Psychology, 17*, 351–372. doi: 10.1016/0022-1031(81)90043-3

Tafarodi, R. W., & Swann, W. B., Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment, 65,* 322–342. doi:10.1207/s15327752jpa6502_8

Tafarodi, R. W., & Swann, W. B., Jr. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences, 31*, 653–673. doi: 10.1016/S0191-8869(00)00169-0

Triandis, H. C. (1995). Culture specific assimilators. In S. Fowler & M. Mumford (Eds.), *Intercultural sourcebook* (Vol. 1, pp. 179–186). Yarmouth, ME: Intercultural Press.

Van Dyne, L., Ang, S., & Koh, C. (2008). Development and Validation of the CQS. In S. Ang & L. Van Dyne (Eds.), *Handbook of Cultural Intelligence: Theory, measurement and application* (pp. 16–38). Armonk, NY: M.E. Sharpe.

Van Oudenhoven, J. P., & Van der Zee, K. I. (2002). Predicting multicultural effectiveness of international students: The Multicultural Personality Questionnaire. *International Journal of Intercultural Relations, 26*, 679–694. doi:10.1016/S0147-1767(02)00041-X

Vulpe, T. (2004, November/December). Training for intercultural competence: Avoiding failures in international assignments. *CERC news, 132*, 1–2. Retrieved from http://www.international.gc.ca/cfsi-icse/cil-cai/pdf/tic-aci-eng.pdf

Waiswol, N. M. (1995). Projective techniques as psychotherapy. *American Journal of Psychotherapy*, *49*, 244–259. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7677204

Ward, P. (1997). *360-degree feedback*. London, UK: CIPD Publishing.

Weiner, B. (1985). An attribution theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573. doi:10.1037/0033-295X.92.4.54

Tables

Table 1

Socio-demographic Characteristics and Intercultural Experiences divided by Subgroups

		NFB (n_1 = 156)	WFB (n_2 = 44)	OFB (n_3 = 33)
Age		*M* = 23.55, *SD* = 3.02	*M* = 22.93, *SD* = 2.59	*M* = 23.82, *SD* = 2.60
Sex	Male	93 (60%)	25 (57%)	15 (46%)
	Female	63 (40%)	19 (43%)	18 (55%)
Target degree	Bachelor	55 (35%)	19 (43%)	15 (46%)
	Master	51 (33%)	15 (34%)	11 (33%)
	Diploma	3 (2%)	2 (5%)	2 (6%)
	Ph.D.	18 (11%)	4 (9%)	3 (9%)
	Other	29 (19%)	4 (9%)	2 (6%)
Study area	Humanities	70 (45%)	17 (39%)	8 (24%)
	Sciences	44 (28%)	13 (29%)	6 (18%)
	Engineering	28 (18%)	10 (23%)	16 (49%)
	Other	14 (9%)	4 (9%)	2 (9%)
Intercultural training experiences	Yes	47 (30%)	10 (23%)	9 (27%)
	No	109 (70%)	34 (77%)	24 (73%)
Intercultural Involvement	Yes	94 (60%)	24 (55%)	19 (58%)
	No	62 (40%)	20 (46%)	14 (42%)
Previous experiences abroad	Internship	52 (15%)	12 (14%)	12 (17%)
	Project	49 (14%)	13 (15%)	9 (13%)
	Studying	75 (21%)	17 (20%)	11 (15%)
	Vacation	132 (38%)	27 (43%)	30 (42%)
	Work	42 (12%)	7 (8%)	9 (13%)

Note. NFB = No feedback group; WFB = Written feedback group; OFB = Oral feedback group. The difference in age between the three groups was not significant: $F(2, 228) = 1.05$, $p = .35$.

Table 2

Means, Standard Deviations, and Cronbach's Alpha for all Dependent Variables

	NFB (n_1 = 156)		WFB (n_2 = 44)		OFB (n_3 = 33)		Total (N = 820)
	M	*SD*	*M*	*SD*	*M*	*SD*	α
TMIC-SA T0	4.43	0.43	4.22	0.52	4.42	0.46	.95
TMIC-SA T1	4.38	0.45	4.19	0.63	4.65	0.42	
Precontemplaction T0	2.68	0.79	2.76	1.03	2.47	0.75	.66
Precontemplation T1	2.83	0.78	2.83	0.99	2.12	0.68	
Contemplation T0	3.38	0.87	3.35	0.81	3.92	0.63	.75
Contemplation T1	3.40	0.90	3.38	0.95	4.27	0.79	
Action T0	3.38	0.94	2.93	0.93	3.89	0.74	.70
Action T1	3.40	0.85	3.26	0.92	4.59	0.65	
Perceived benefit from TMIC T0	3.50	0.81	3.70	0.98	3.80	0.74	.89
Perceived benefit from TMIC T1	3.69	0.99	3.48	0.98	4.88	0.90	
Self-confidence T0	4.39	0.72	4.08	0.60	4.39	0.68	.79
Self-confidence T1	4.36	0.73	4.24	0.68	4.75	0.61	
Self-understanding T0	3.65	0.88	3.30	0.69	3.44	0.77	.89
Self-understanding T1	3.59	0.82	3.66	0.84	4.77	0.67	

Note. NFB = No feedback group; WFB = Written feedback group; OFB = Oral feedback group.

Table 3

Multivariate Randomized Pretest-Posttest Control Group Design

	Group	Pretest	Treatment	Posttest
R	G_1 (n_1 = 156)	$\overline{Y}_{1\text{-}7,\ pre}$	No feedback (x_0)	$\overline{Y}_{1\text{-}7,\ post}$
	G_2 (n_2 = 44)	$\overline{Y}_{1\text{-}8,\ pre}$	Written feedback (x_1)	$\overline{Y}_{1\text{-}8,\ post}$
	G_3 (n_3 = 33)	$\overline{Y}_{1\text{-}8,\ pre}$	Collaborative test feedback (x_2)	$\overline{Y}_{1\text{-}9,\ post}$

Note. R = Randomized; G = Group; $\overline{Y}$ = Means of Dependent Variables; $\overline{Y}_1$ = TMIC-SA total score, $\overline{Y}_2$ = Precontemplation stage, $\overline{Y}_3$ = Contemplation stage, $\overline{Y}_4$ = Action stage, $\overline{Y}_5$ = Self-understanding, $\overline{Y}_6$ = Self-confidence, $\overline{Y}_7$ = Perceived benefit from TMIC participation, $\overline{Y}_8$ = General satisfaction with feedback, $\overline{Y}_9$ = Specific satisfaction with collaborative assessment feedback.

Table 4

Mean Differences and Confidence Intervals for the Interaction Effect Time x Feedback

		NFB (n_1 = 156)	WFB (n_2 = 44)	OFB (n_3 = 33)
Perceived benefit from TMIC	Δ*M*	0.187	-0.297	1.079
$F(2, 226) = 25.31$, $MSE = 8.73$,	95% CI	[0.164, 0.210]	[-0.343, -0.252]	[1.029, 1.128]
$p = .000$, $\eta^2 = .18$				
Self-confidence	Δ*M*	-0.037	0.109	0.356
$F(2, 226) = 8.05$, $MSE = 1.10$,	95% CI	[-0.039, -0.035]	[0.105, 0.114]	[0.351, 0.361]
$p = .000$, $\eta^2 = .07$				
Self-understanding	Δ*M*	-0.066	0.286	1.324
$F(2, 226) = 69.55$, $MSE = 13.24$,	95% CI	[-0.061, -0.072]	[0.298, 0.275]	[1.336, 1.312]
$p = .000$, $\eta^2 = .38$				
Precontemplation stage	Δ*M*	0.154	0.057	-0.383
$F(2, 226) = 6.82$, $MSE = 1.91$,	95% CI	[0.158, 0.150]	[0.066, 0.050]	[-0.374, -0.392]
$p = .001$, $\eta^2 = .06$				
Contemplation stage	Δ*M*	0.014	0.045	0.382
$F(2, 226) = 3.77$, $MSE = 0.91$,	95% CI	[0.005, 0.023]	[0.027, 0.063]	[0.363, 0.403]
$p = .021$, $\eta^2 = .03$				
Action stage	Δ*M*	-0.079	0.333	0.735
$F(2, 226) = 20.07$, $MSE = 4.99$,	95% CI	[-0.068, -0.089]	[0.355, 0.312]	[0.758, 0.711]
$p = .000$, $\eta^2 = .15$				

Note. NFB = No feedback group; WFB = Written feedback group; OFB = Oral feedback group.

Appendix D

Pre-print Version of Paper C

Appendix D provides a pre-print version of Paper C: Schnabel, D., Kelava, A., Van de Vijver, F. J. R., & Seifert, L. (2014). *Examining psychometric properties, measurement invariance, and construct validity of a short version of the Test to Measure Intercultural Competence (TMIC-S) in Germany and Brazil.* Manuscript submitted for publication.

Pre-print refers to the originally in 2014 submitted version of the paper before any peer-review. Please note, that there are consequently substantial differences between the pre-print and the published versions.

Examining Psychometric Properties, Measurement Invariance, and Construct Validity of a Short Version of the Test to Measure Intercultural Competence (TMIC-S) in Germany and Brazil

Abstract

The goal of this study was to develop and validate a short version (TMIC-S) of the Test to Measure Intercultural Competence (TMIC; Schnabel, Kelava, Seifert, & Kuhlbrodt, 2014). TMIC-S measures six malleable abilities, which support handling novel or difficult cross-cultural situations. The short TMIC-S version, comprising 25 self-report and six situational judgment items, was administered to 1032 Germans and 769 Brazilians. Confirmatory factor analysis (CFA) showed a good fit of the six-factor multimethod model in both samples. Measurement invariance was examined by multigroup CFA, which showed metric and scalar invariance of the TMIC-S. Construct validity was shown as the German and Brazilian version correlated with the Cultural Intelligence Scale (Van Dyne, Ang, & Koh, 2008). Additionally, prior intercultural experience was positively associated with latent TMIC-S means in both samples, highlighting criterion validity. Thus, TMIC-S is a valid instrument, which is economic at the same time. It can be applied in training, coaching, and counseling settings as well as in personnel selection.

Keywords: intercultural competence; multimethod test; situational judgment test; measurement invariance; cross-cultural generalizability.

1 Introduction

Intercultural competence (ICC) has been of scientific interest for decades (cf. Bennett, 1993; Kealey & Ruben, 1983; Sinicrope, Norris, & Watanabe, 2007). However, despite the increasing economic and political relevance of intercultural collaboration (Leung, Ang, & Tan, 2014; Sheahan, 2005), defining and measuring ICC appeared to be elusive (Ang et al., 2007). The present study set out to develop and test a new, theory-based instrument that focuses on intercultural competences in the behavioral domain. The instrument is tested in a cross-cultural framework.

1.1 The Concept of Intercultural Competence

Researchers in various disciplines have developed ICC definitions (Byram, 1997; Fantini & Tirmizi, 2006; Gudykunst, 1994; Kim, 1992; Lambert, 1994; Thomas, 2003), which differ in focus, extent, and conceptualization (Spritzberg & Changnon, 2009). Bolton (2007a) established a threefold taxonomy for ICC definitions: 1) listing models in which different characteristics of intercultural competence are simply collected (e.g., Brislin, 1981; Ruben, 1976); b) structure models in which the characteristics of intercultural competence are assigned to affective, cognitive, and behavioral categories (e. g., Dauner, 2011; Gertsen, 1990; Ting-Toomey, 1993); c) procedural models in which intercultural competence is defined as context-specific competence to act due to its manifold connections with other core competences (e.g., Bolton, 2007b). Spitzberg and Changnon (2009) distinguished five types of competence models: compositional, co-orientational, developmental, adaptational, and causal process. The latter two model types assume that intercultural competence consists of several *related* components. In causal process models the nature of those relations is defined via correlations tested in empirical research.

Several authors assumed that ICC is a heterogeneous construct involving multiple dimensions, which are necessary to interact with people from other cultures adequately and

effectively (e.g., Bergemann & Bergemann, 2005; Fantini & Tirmizi, 2006; Hammer, Bennett, & Wiseman, 2003; Müller & Gelbrich, 2004, Spitzberg & Changnon, 2009). Intercultural competence is in theory often defined as an ability or a skill (Spitzberg & Changnon, 2009). However, as exemplarily shown in Table 1, prominent measurement approaches operationalize intercultural competence merely as stable personality traits (Chen & Starosta, 1997; Kelley & Meyers, 1995; Ruben, 1976), cultural intelligence (Earley & Ang, 2003; Van Dyne, Ang, & Koh, 2008), sensitivity to cultural differences (Hammer, 2008; Hammer et al., 2003), or a combination of those (Fantini & Tirmizi, 2006; Koester & Olebe, 1988; Yamazaki & Kayes, 2004). This obviously creates a gap between the conceptualization and the measurement of intercultural competence. In a recent overview of intercultural competence, Leung et al. (2014) reviewed models that are based on a) traits, b) attitudes, or c) capabilities. Whereas several models were described that cover traits and attitudes, only one model (Cultural Intelligence Scale by Van Dyne et al., 2008) was identified, which focuses on malleable abilities. Mixed-models that integrate two or more concepts (e.g., traits and capabilities) in their measurement approach do not distinguish potential predictive or hierarchical relations between traits, attitudes, and capabilities (cf. Leung et al., 2014). Schnabel, Kelava, Seifert, and Kuhlbrodt (2014) developed an onion model of intercultural competence (cf. Schuler & Prochaska, 2001) that places malleable behavior-related abilities at the core of intercultural competence. Personality traits, knowledge, and attitudes are positioned at the outside layer of the onion model as they are understood as related antecedents, which influence the nature of intercultural competences, but are not intercultural competences per se. Locating malleable capabilities at the heart of intercultural competence has one major advantage: In contrast to stable characteristics of a person a malleable construct is open to training (Schnabel et al., 2014). The theoretical basis of Schnabel et al.'s (2014) model integrates different features of existing theories, such as the understanding of

(intercultural) competence as an ability, which can be learned, which helps individuals to master intercultural situations (e.g., Earley & Ang, 2003; Fantini & Tirmizi, 2006; Weinert, 2001), which involves multiple facets, and which triggers an individual's global behavioral orientation (Schuler & Prochaska, 2001). These aspects also serve as the operationalization basis for the Test to Measure Intercultural Competence (TMIC; Schnabel et al., 2014). Specifically, intercultural competence is seen as a global orientation of behavior. Multiple facets are involved, which belong to one of the three following competence groups: a) "social competence" (e.g., communication competences), b) "personal competence" (e.g., learning competences), and c) "methodological competence" (e.g., self-management competences). These competences can be acquired, are directly influencing human behavior, and enable an individual – together or separately – to handle novel intercultural situations or problems while interacting with people from other cultures. Table 2 gives an overview of the intercultural competence model of Schnabel et al. (2014).

1.2 Measuring Intercultural Competence

Along with the variety of theoretical approaches there are numerous instruments to measure ICC. Examples are given in Table 1.

Recently, Gabrenya, Moukarzel, Pomerance, Griffith, and Deaton (2011) analyzed 34 instruments that aim to measure intercultural competence. Results concerning face validity, construct validity (convergent and divergent), and criterion validity (congruent and predictive) were taken into consideration. Validation results were then categorized in poor, moderate, and good. Only for seven tests validity findings have been published with five having satisfactory criterion validity (e.g., Cultural Intelligence Scale, Van Dyne et al., 2008; Multicultural Personality Questionnaire, Van Oudenhoven & Van der Zee, 2002).

A large number of those scales incorporate both personality traits and competences (Ang et al., 2007). Additionally, these instruments use exclusively Likert scale based self-

report items. Thus, there is a strong need for performance-based measures, such as situational judgment items, which gauge all major aspects of intercultural competence (Leung et al., 2014).

Representing a sophisticated attempt to measure ICC using a multimethod approach that refers to malleable competences rather than stable personality traits, Schnabel et al. (2014) recently published the reliable and valid Test to Measure Intercultural Competence (TMIC), a German instrument. The TMIC is so far the only multimethod ability-based behavior-related measure of ICC, which makes it especially promising for training, coaching, and counseling situations (Gabrenya, Griffith, Moukarzel, Pomerance, & Reid, 2012). The TMIC assesses 17 competence facets, which are assigned to the following six second-order factors: Communication, learning, social interaction, self-knowledge, self-management, and creating synergies. Following the recommendation of using more than one method to assess intercultural competence (Deardorff, 2004; Gabrenya et al., 2011; Leung et al., 2014), the TMIC combines 75 self-report questions (TMIC-SA) with 17 situational judgment items (TMIC-SJT; see Appendix for a selection of items). Self-report measures aim at measuring the self-concept of a person, whereas situational judgment items uncover behavioral preferences (Bledow & Frese, 2009). A situational judgment test (SJT; McDaniel, Hartman, Whetzel, & Grubb, 2006) consists of critical incident scenarios with a fixed number of behavior alternatives as answering options. Testees must then choose one of those options (McDaniel et al., 2006). Situational judgment tests are known to have the advantages of direct measurements on the one hand and as being particularly practical, economic, valid, and robust against biases on the other hand (e.g., Hooper, Cullen, & Sackett, 2006; Lievens, Peeters, & Schollaert, 2008; McDaniel et al., 2006; McDaniel & Nguyen, 2001; Weekly & Ployhart, 2006).

Schnabel et al. (2014) reported convergent validity of the TMIC as both parts correlated

with cultural intelligence (Cultural Intelligence Scale; CQS; Van Dyne et al., 2008); $r = .67$, $p < .001$ for TMIC-SA and $r = .26$, $p < .001$ for TMIC-SJT as well as with global competence (Intercultural Sensitivity Index; ISI; Olson & Kroeger, 2001); $r = .63$, $p < .001$ for TMIC-SA and $r = .26$, $p < .001$ for TMIC-SJT. Criterion validity of TMIC-SA and TMIC-SJT was shown in using four external criteria of intercultural prior knowledge (Bhawuk & Brislin, 1992; Koester & Olebe, 1988; Morris & Robie, 2001). Furthermore, a positive, strong correlation ($r = .49$, $p < .001$) was found between the overall values of the TMIC-SA and the TMIC-SJT, which shows that both parts measure one intercultural competence construct without making each other redundant as self-concept is covered in one and behavioral preferences in the other method (Bledow & Frese, 2009). These results provide validity support for the TMIC in Germany. Thus, it is worth addressing its cross-cultural applicability.

To do so a short version of the TMIC was created (TMIC-S). The conceptual starting point for TMIC-S was the six second-order factor model described before. To select first-order factors for the TMIC-S, we followed three main strategies: (1) we chose facets, which are most influential concerning intercultural competent behavior, (2) we excluded facets, which might have been related contentwise to certain culture dimensions, and (3) we decided on those six competencies, which functioned together as an intercultural competence model and thus fit the data well. Concerning the first strategy we considered existing research results, which found empathy (e.g., Fantini & Tirmizi, 2006; Koester & Olebe, 1988; Ruben, 1976), cultural awareness (e.g., Chen & Starosta, 2000; Thomas, Kammhuber, & Layes, 1997; Triandis, 1977) as well as broad cultural knowledge acquisition (e.g., Deardorff, 2004; Van Dyne et al., 2008) to be crucial characteristics of an individual in the intercultural context. Moreover, we clearly aimed at creating a competence measure, which might hold value across cultures. Therefore, the second strategy consisted of excluding competences, which have the potential to be driven rather by culture values than by ability. For example,

we considered clarity in communication to be influenced by communication style (Hall & Hall, 1993; Thomas, 1991), building trusting relationships and building professional networks by task- versus relationship-orientation (Bass, 1990; Thomas, 1991), and integration in groups by individualism versus collectivism (Hofstede & Bond, 1984). The final TMIC-S version included the following intercultural competence facets: Sensitivity in communication, cross-cultural information seeking, socializing, cultural identity reflection, goal setting, and mediation of interests. Overall, 25 self-report items as well as six situational judgment items were selected. A short description of the six facets, example self-report items in German, Portuguese, and English as well as all six situational judgment items in English can be found in the Appendix.

1.3 Cross-Cultural Generalizability of Intercultural Competence

In contrast to constructs like personality (e.g., Cheung, 2009; Church, 2001; McCrae et al., 2005) or leadership styles (e.g., Brodbeck et al., 2000; House et al., 1999), only little empirical research has been conducted regarding the actual cross-cultural generalizability of intercultural competence. There are different points of view in the literature about universal versus culture-specific aspects of ICC (Arasaratnam & Doerfel, 2005). Deardorff (2004) conducted a Delphi study with 21 intercultural experts to learn more about definitions, components, and assessment of ICC. Her results indicate that definitions and components of ICC are evaluated as rather independent from culture, whereas generalizing assessment instruments across cultures is seen as challenging by intercultural experts. Interestingly, although several experts seemed to contemplate a context-specific assessment approach to intercultural competence in Deardorff's (2004) early Delphi rounds, they could not agree upon this idea in the final round.

Deardorff (2004) pointed out the need to apply ICC measures in a cross-cultural context. Few researchers have taken up this challenge so far. Zhong (1998) examined factors

influencing the perception of intercultural communication competence with Chinese and American college students and found no differences between the two cultures' perceptions of intercultural communication competence items. Matsumoto et al. (2003) investigated the generalizability of the Intercultural Adjustment Potential Scale (ICAPS), which was originally intended for Japanese sojourners and immigrants by applying it to other cultures. The results indicated that the concept of intercultural adjustment is rather culture-neutral and that international use of the ICAPS is possible. The Intercultural Sensitivity Scale (ISS; Chen & Starosta, 2000), originally developed in the US, was applied in Germany by Fritz, Mollenberg, and Chen (2002). These authors were able to replicate the theoretical model and found satisfying results using confirmatory factor analysis (CFA). However, a second replication study of Fritz, Graf, Hentze, Mollenberg, and Chen (2005) produced deviating results that questioned the previously proposed transferability. Tamam (2010) applied the ISS in a non-western context, namely Malaysia, and found that 21 items of the original 24 items were loading on three factors (interaction attentiveness and respect, interaction openness, and interaction confidence) instead of the originally proposed five factors.

The studies reviewed before have one major characteristic in common: they were aiming to attain psychometric equivalence solely by replicating factor structures in a different culture than the original without investigating equivalence of those structures across cultures. Even though some findings are in favor of a universally applicable core set of intercultural competence the basis for deciding whether ICC, as measured by currently employed instruments, can be generalized across cultures is rather shallow. A prerequisite for finding a core set of ICC dimensions would be the availability of instruments in several languages that measure a single ICC construct (Ziegler & Bensch, 2013). This can be achieved by introducing measurement invariance procedures (MI; Meredith, 1993) to cross-cultural research of intercultural competence. A given measurement invariance allows a comparison

between scores on the latent construct in different groups in that it assures that the observed items contribute similarly to the latent construct in terms of their factor loadings and intercepts (Meredith, 1993). Measurement invariance is also one major part of the equivalence concept in cross-cultural research in general (Van de Vijver & Leung, 1997). Construct equivalence (also called structural equivalence) is attained when an instrument measures the same construct in different cultures. Van de Vijver and Leung (2009) suggested that a well validated instrument, which aims to measure intercultural competence, should have good psychometric properties on the one hand and demonstrate measurement equivalence across cultures on the other hand. The next level is measurement unit (or metric) equivalence of two scales. The third level is scalar equivalence and points to identity of both the measurement unit and origin of the measurement scales (Matsumoto & van de Vijver, 2011; van de Vijver & Leung, 1997).

1.4 Study Aim and Hypotheses

The present study takes the multidimensional ability concept as well as the multimethod measurement approach of the Test to Measure Intercultural Competence (Schnabel et al., 2014) as a starting point to develop and validate a short version of the TMIC (TMIC-S). The TMIC-S incorporates a selection of the 17 proposed competence facets, which is based on the six-factor second order model of the TMIC (see Table 2). A short version of the TMIC would be especially useful for application in practice as it is less time consuming than the TMIC long version (Schnabel et al., 2014). The short version avoids cognitive overload of the testee (Eignor, 2013).

Further, most ICC theories and instruments have an Anglo-American developmental background and are therefore exclusively addressed to an English-speaking community (Deardorff, 2009; Martin 1993). The TMIC adds to this research as it was developed in Europe with German being the original language. Also, to our knowledge no attention was

given so far to cross-cultural comparisons of intercultural competence theories and instruments that involve Latin-American cultures. As Brazil belongs to the five countries in the world (the so-called BRICS countries), which currently experience the fastest and strongest economic growth, this is quite astonishing. Brazil is in itself a multicultural society (Gawora, de Souza Die, & Barbosa, 2011) that becomes even more diverse due to international traffic from and to Brazil. This also creates a strong need for intercultural diagnostic instruments in Brazil as internationalization of the society and the labor market rapidly increases (Muritiba, Muritiba, Campanário, & de Albuquerque, 2010). In assessing measurement invariance this study also completes a pioneering task in intercultural competence research.

Taking the research reviewed above as well as the goal of this study into consideration, the following hypotheses are tested in the present study:

Hypothesis 1: The theoretically expected six factors hold for both the German and Brazilian TMIC-S. Moreover, measurement invariance (Meredith, 1993) can be established for the German and Brazilian TMIC-S.

Hypothesis 2: All latent variables of the German and Brazilian TMIC-S show positive correlations with the four dimensions of the Cultural Intelligence Scale (CQS; Van Dyne et al., 2008), indicating a convergent construct validity.

Hypothesis 3: We expect higher means in the latent factors of the TMIC-S for individuals who a) took part in an intercultural training before (Bhawuk & Brislin, 1992; Thomas, Kinast, & Schroll-Machl, 2006), b) are privately or professionally involved in intercultural topics (Loboda, 2003; Schacher, 2011), and c) have stayed abroad at least once for three or more months (criterion validity; Morris & Robie, 2001).

2 Method

2.1 Participants and Procedure

2.1.1 German sample. The German sample was mainly recruited from German universities and student organizations with an international focus (e.g., Erasmus Mundus Programme; European Union, 2014) and consisted of 1037 participants: 597 were women (58%), 429 were men (41%), and 11 data were missing (1%). The average age was 27.96 years (*SD* = 9.47). 631 individuals indicated that they have stayed abroad for more than three months before (61%, 24 missing). Previous participation in an intercultural training program was the case for 687 Germans (66%, 1 missing) and intercultural involvement for 689 (66%, 1 missing).

2.1.2 Brazilian sample. Potential participants were contacted through the Brazilian partner university network of several German universities as well as through international student organizations and communities. Overall, 769 Brazilians took part in the survey. The average age of the 415 women (54%) and 354 men (46%) was 27.38 years (*SD* = 10.61). 436 (57%) participants reported a previous stay abroad with a duration of more than three months. 687 individuals (66%) experienced an intercultural training before and 604 (78%) were privately or professionally involved with other cultures.

2.2 Instruments

2.2.1 Test to Measure Intercultural Competence (TMIC). The Test to Measure Intercultural Competence (TMIC; Schnabel et al., 2014) includes 75 self-report items (six-point Likert scale from *does not apply at all* to *fully applies*) and 17 situational judgment items (McDaniel & Whetzel, 2005). An ESEM-procedure (Exploratory Structural Equation Modeling; Asparouhov & Muthén, 2009) showed a very good model fit, thereby supporting the theoretically driven and empirically grounded 17 competence facets, $\chi^2(1636, N = 641) = 2579.85$, $p < .001$; $\chi^2/df = 1.58$; RMSEA = .031 (90% CI = .029–033); SRMR = .017; CFI = .955; TLI = .927. Confirmatory factor analysis (CFA) showed an acceptable fit of the six second-order factor structure, which combined self-report as well as situational judgment

items, $\chi^2(3987, N = 641) = 8280.09$, $p < .001$; $\chi^2/df = 2.08$; RMSEA = .040 (90% CI = .040–.046); SRMR = .076; CFI = .820, TLI = .810. The internal consistency (Cronbach's α) of the 17 facets ranged between .69 and .90. The overall reliability of the TMIC-SA was found to be excellent (α = .96; Schnabel et al., 2014; Schnabel & Kelava, 2013). When combined with collaborative test feedback, TMIC even served as a brief intervention for students going abroad (Schnabel, 2013; Schnabel, Kelava, & Van de Vijver, 2014).

The Brazilian test version was developed in two steps. First, we checked for construct and method bias of the instrument (Van de Vijver & Leung, 1997). Specifically, we wanted to ensure that the construct being measured does not depend on specific aspects of one culture and that biases in scores are minimized. Second, the process of back-translation (Brislin, 1970) was used in that two bilingual translators with extensive knowledge of the source and target language were involved. A third expert reviewed the final version.

2.2.2 Cultural Intelligence Scale (CQS). The Cultural Intelligence Scale (CQS; Van Dyne et al., 2008) is based on the aforementioned concept of cultural intelligence. With a total of 20 self-report items the four dimensions of cultural intelligence are assessed on a seven-point Likert Scale. A German and Brazilian version of the CQS was created applying the process of back-translation (Brislin, 1970). CQS was selected for validation purposes as it is the only instrument that aims to measure a malleable aspect of intercultural competence.

2.2.3 External Criteria. Previous studies were considered during the formulation of those items (Bhawuk & Brislin, 1992; Koester & Olebe, 1988; Schacher, 2011). Each of the three external criteria was included in the survey with one item using forced choice answer format ("Do you privately or professionally deal with different cultures?", "Have you ever taken part in intercultural training?*",* and "What was the longest time that you ever spend abroad at a stretch?").

2.3 Data Analysis

The online survey software Unipark EFS Survey (QuestBack GmbH, 1999-2012) was used for collecting data. Data analysis was conducted by the statistical software Mplus for Windows (Version 7.1) and IBM SPSS Statistics for Windows (Version 22.0). Confirmatory factor analysis, multigroup SEM-comparisons, and correlations were computed with Mplus (Muthén & Muthén, 1998-2012). IBM SPSS Statistics (IBM, 2013) was used to analyze descriptive statistics.

3 Results

The results part is structured as follows: First, we describe the model fit as well as the psychometric properties of all six TMIC-S dimensions for the German and Brazilian test version separately. Second, we report results of multigroup confirmatory factor analyses (MG-CFA) along with model comparisons that outline measurement invariance of the German and Brazilian TMIC-S. Third, we provide evidence for construct and criterion validity of the TMIC-S. The results part closes with exploratory comparisons of the Brazilian and German samples concerning their intercultural competence.

3.1 Model Fit and Descriptive Results

To examine model fit of the German and Brazilian TMIC-S confirmatory factor analysis (CFA) was computed in each group. All six situational judgment items were integrated as categorical variables. Therefore, the WLSMV estimator (Muthén, du Toit, & Spisic, 1997) was selected.

Satisfactory fit indices were attained in the German sample, $\chi^2(419, N = 1037) = 824.10$, $p < .001$; $\chi^2/df = 1.97$; RMSEA = .036 (90% CI = .033–.040); WRMR = 1.095; CFI = .913; TLI = .904 as well as in the Brazilian sample, $\chi^2(419, N = 769) = 919.35$, $p < .001$; $\chi^2/df = 2.19$; RMSEA = .039 (90% CI = .036–.043); WRMR = 1.168; CFI = .902; TLI = .892. Factor loadings for both groups can be found in Table 3. For the self-report items the cut-off

value for accepted loadings was set at .40. This was reached in both groups. The lowest loading can be found for the factor mediation of interest (item MI4 = .62) in the German group and for the factor goal setting (GS3 = .51) in the Brazilian Group. For the situational judgment items lower loadings were allowed when significant. In both samples SC-SJT loaded weak, yet significant on the factor sensitivity in communication.

Internal consistencies ranged from acceptable to good (see Table 3). Cronbach's alpha was overall a bit lower in the Brazilian sample.

3.2 Measurement Invariance of the TMIC-S in Germany and Brazil

Measurement invariance (Meredith, 1993) of the TMIC-S was investigated using multigroup confirmatory factor analysis (MG-CFA) in Mplus. Three MG-CFAs with varying (nested) parameter restrictions were computed to test for measurement invariance. The WLSMV estimator was used throughout and adjusted chi-square difference tests were applied (Asparouhov & Muthen, 2006). In the configural invariance model all parameters were freely estimated in each group, whereas factor loadings were held equal in both groups for the weak factorial invariance model (metric invariance). To investigate scalar invariance intercepts of the self-report items and thresholds of the situational judgment items were also restricted. The most common procedure of comparing measurement invariance models is to look either at differences in the Comparative Fit Index (CFI) or in the Root Mean Square Error of Approximation (RMSEA; Chen, 2007; Marsh, Nagengast, & Morin, 2012) with the latter one controlling for parsimony (Marsh, 2007). Using the chi-square difference test is also a frequently employed procedure, although problematic in some cases due to its sensitivity to sample size (Bentler & Bonett, 1980; Jöreskog, 1993; Milfont & Fischer, 2010). In the present measurement invariance analysis we followed the above mentioned recommendation of using differences in both the RMSEA and CFI. Hereby, models with a ΔRMSEA $\leq .015$, ΔCFI $\leq .010$, and ΔTLI $\leq .010$ were favored (Chen, 2007). Models were

also compared with a chi-square difference test. Due to the large number of comparisons made, a Bonferroni-Holm-correction (Holm, 1979) was applied. An alpha level of .01 was used. As can be seen in Table 4, some $\Delta\chi^2$ values were significant, yet difficult to interpret given our large sample sizes. Values of ΔRMSEA and ΔTLI pointed clearly to scalar invariance. Finally, values of ΔCFI showed an unexpected pattern in that the CFI value increased from configural invariance to metric invariance and decreased with the transition to scalar invariance. However, given that the CFI value of the configural and scalar invariance model were close to each other, we interpreted the ΔCFI values as essentially supporting scalar invariance. All in all, we found considerable support for scalar invariance.

3.3 Construct Validity of the German and Brazilian TMIC-S

A multigroup CFA without parameter restrictions (configural invariance) revealed a good model fit of the cultural intelligence model in the German and Brazilian samples, $\chi^2(328, N = 1082) = 588.26$, $p < .001$; $\chi^2/df = 1.79$; RMSEA = .050 (90% CI = .044–.057); SRMR = .049; CFI = .939; TLI = .929. The following fit indices were found for the model in which factor loadings were held invariant between the German and Brazilian samples, $\chi^2(344, N = 1082) = 607.16$, $p < .001$; $\chi^2/df = 1.77$; RMSEA = .049 (90% CI = .043–.056); SRMR = .056; CFI = .938; TLI = .931. Delta coefficients of ΔRMSEA = .001, ΔCFI =.001, and ΔTLI = .002 as well as a nonsignificant chi-square difference test for the configural invariance and the metric invariance models, $\chi^2(16) = 18$, $p = .32$, indicated metric invariance of the CQS in this study.

Table 5 shows the intercorrelations of all four CQS factors and all six TMIC-S factors in the German and Brazilian samples. Correlations between the CQS factors and TMIC-S factors were mostly moderate to high, $r = .30 - .64$. Only one high correlation between motivational CQ and cultural identity reflection was found in the German sample, $r = .78$. Lowest correlations were reached for meta-cognitive CQ as well as behavioral CQ with

information seeking of the German TMIC-S, r = .18, and cognitive CQ with goal setting in the Brazilian TMIC-S, r = .25.

To investigate if correlations of the six TMIC-S factors with the four CQS factors are the same in both samples, two further models were computed: In one model all parameters were freely estimated and in the other model the 24 correlation pairs were held equal in both samples. Despite the significant chi-square difference between the models, $\chi^2(24) = 45.49$, p = .005, delta coefficients of ΔRMSEA = .001, ΔCFI =.007, and ΔTLI = .008 strongly suggest that correlations between the TMIC-S and the CQS are equal in both samples.

3.4 Criterion Validity of the German and Brazilian TMIC-S

Table 6 gives an overview about the multigroup SEM comparisons in the German and Brazilian samples. For each external criterion two groups were compared in both samples: (a) involved versus not involved in intercultural matters (hypothesis 3a), (b) experience abroad longer than 3 months versus shorter than 3 months (hypothesis 3b), and (c) previously versus never participated in an intercultural training (hypothesis 3c).

In the German sample individuals with more intercultural involvement had a higher sensitivity in communication, were better information seekers and mediators as well as more advanced in cultural identity reflection and socializing. Brazilians who stated that they privately or professionally deal with other cultures outperformed those with less intercultural involvement in all intercultural competences except for goal setting, thereby supporting hypothesis 3a.

Germans who stayed more than three months abroad before had higher means in information seeking, cultural identity reflection, and socializing. Individuals in the Brazilian sample who reported experiences abroad of longer than 3 months were more sensitive in communication, capable of mediating different interests, vigor in reflecting upon their cultural character, and active in building intercultural relationships. Thus, hypothesis 3b is

partially supported.

Additionally, German training participants were more sociable and advanced in cultural identity reflection. They also had higher values in information seeking and mediation of different interests. In the Brazilian sample intercultural training participation had a positive influence on the latent means of sensitivity in communication, mediation of interests, cultural identity reflection, and socializing. Therefore, hypothesis 3c is supported at large.

3.5 Exploratory Comparison of the TMIC-S Facets in the Two Samples

As scalar invariance was established for the German and Brazilian TMIC-S version, differences in latent means on the intercultural competence dimensions could be tested. For the following latent factors higher means were found for the German group in comparison to the Brazilian: Sensitivity in communication, $M = 0.17$, $SEM = 0.05$; $\chi^2(1) = 15.34$, $p < .001$, $d = .14$, goal setting, $M = 0.40$, $SEM = 0.06$; $\chi^2(1) = 43.90$, $p < .001$, $d = .28$, mediation of interests, $M = 0.36$, $SEM = 0.07$; $\chi^2(1) = 30.50$, $p < .001$, $d = .22$, and socializing, $M = 0.38$, $SEM = 0.05$; $\chi^2(1) = 55.43$, $p < .001$, $d = .32$. The two samples did not differ in information seeking, $M = 0.00$, $SEM = 0.06$; $\chi^2(1) = 1.00$, $p = .32$, $d = .00$, and cultural identity reflection, $M = 0.05$, $SEM = 0.05$; $\chi^2(1) = 1.76$, $p = .19$, $d = .04$. It can be concluded that there were small, yet for some aspects significant cross-cultural differences showing higher scores for German participants.

4 Discussion

The goal of the study was to create a short version of the Test to Measure Intercultural Competence (TMIC-S), which can be applied in Germany and Brazil and can serve as a tool to investigate relevant intercultural competence facets. As in Schnabel et al.'s (2014) model, intercultural competence (ICC) is understood as a behavioral orientation that can be acquired. Self-report as well as situational judgment items were integrated in the TMIC-S to follow the multimethod approach of the original TMIC version.

4.1 Summary of Important Results

For the six-factor model of the German and Brazilian TMIC-S an acceptable fit to the data, good psychometric properties as well as invariance of the factor structure, loadings, and intercepts/thresholds was established, which support hypothesis 1. As expected most TMIC-S factors positively correlated with the four dimensions of the Cultural Intelligence Scale in Germany and Brazil, which underpins hypothesis 2. Results suggested that the 24 correlation pairs were equal in both groups. Concerning the external criteria slightly higher means in most competences of the TMIC-S were found for Brazilians and Germans who attended an intercultural training in the past, were professionally or privately involved with intercultural matters, and who stayed abroad for more than three months before. Together this partially confirms hypothesis 3.

4.2 Implications

Our study has three major implications. First, in establishing measurement invariance this study substantially adds to research about intercultural competence as it leads the way to a core set of intercultural competence in Germany and Brazil. This finding supports the potential of intercultural competence as a universal concept, as measured by the TMIC-S.

Second, factor loadings for some situational judgment items especially in the Brazilian group were below the usual cut-off value of .30 (Kline, 1994). Bledow and Frese (2009) stated that the correlation between the self-report and the situational judgment method could be expected to be no more than moderate. Clearly, both methods should measure the same competence. However, self-report items typically assess the self-concept of a person and situational judgment items refer to behavioral preferences. Thus, both methods are complementary rather than mutually exclusive or in competition. Also, Schnabel et al. (2014) found a moderate correlation between all 75 self-report items and all 17 situational judgment items in the TMIC. The low to moderate relations between both methods are in line with the

lower loadings of the situational judgment items on the homogenous factor as factor loadings can be understood as correlations of an item with a factor (Kline, 1994). As there is a strong need for instruments that integrate two methods to measure ICC (Bolten, 2007b; Deardorff, 2004; Gabrenya et al., 2011; Leung et al., 2014), we decided to keep the situational judgment items despite the lower loadings.

Third, motivational CQ and cultural identity reflection were strongly related in the German sample. Van Dyne et al. (2008) defined motivational CQ as a special kind of intrinsic motivation and self-efficacy, which is directed towards understanding cultural differences. That very cognitive process is required when individuals reflect the cultural identity of themselves or others, especially when interacting with people from other cultures or living abroad (Thomas, 2003). In the CQS motivational CQ is assessed with items such as "I enjoy living in cultures that are unfamiliar to me." and "I enjoy interacting with people from different cultures." (Ang et al., 2007), which could explain the relatively high correlation of motivational CQ and cultural identity reflection.

4.3 Limitations

The present study has two specific limitations. We included a restricted number of constructs in the present study. TMIC (Schnabel et al., 2014) with its two methods and 92 items has many advantages, but it is also a rather long instrument. To avoid participants' fatigue we decided on a selected number of external criteria as well as related scales. Clearly, there are additional constructs such as personality traits or cognitive abilities, which are worth to investigate in future studies of TMIC-S.

Moreover, the present study involved two diverse cultures (Western versus Latin-American). This should be taken as a conceptual starting point for further cross-cultural comparisons, specifically those focused on comparing Western with Asian cultures. Especially, task performance in a situational judgment test seems to differ across cultures,

demographic characteristics, and context (Whetzel, McDaniel, & Nguyen, 2008).

4.4 Future research

As Schnabel et al. (2014) were the first to integrate situational judgments items in a self-report instrument to measure intercultural competence more research is needed to understand the unique functioning of this method. The following questions should be addressed in future studies: 1) Is there a difference in measurement invariance across cultures for self-report measures and situational judgment tests? 2) Which unique components of intercultural competence facets are addressed through self-report items and which through situational judgment items? 3) Which consequences would a manipulation of response instructions (behavior-related versus knowledge related; McDaniel et al., 2006) have on an individual's TMIC-S result?

Additionally, we see a strong need to focus research more on cross-cultural validation of scales in the field to increase understanding about how intercultural competence and related constructs function across borders. For example, to examine construct validity of different language versions of an instrument related scales have to be available in those languages as well. Cultural intelligence was hypothesized to be related to intercultural competence, but the Cultural Intelligence Scale (CQS) was neither available in German nor in Portuguese. Thus, we had to translate the CQS. Fit indices showed that CQS worked quite well in different languages. However, we also found factor intercorrelations of $r > .80$, which queries the proposed latent variables structure. Matsumoto and Hwang (2013) recently argued that there might be a general CQ factor instead of four distinct ones. Clearly, this has to be empirically investigated in the future.

Schnabel et al. (2014) proposed the TMIC to support the development processes of students or employees who will spend a longer period abroad or who intensively interact with individuals from other cultures in their home countries. Like in the original version TMIC-S

incorporates competences that can be developed through interventions such as trainings, coachings, or counseling sessions. This raises the potential to create a holistic approach to intercultural competence development. The process might look as follows: 1) Analyzing the status quo with TMIC-S, 2) defining development goals in a (therapeutic) test feedback session, 3) working on intercultural competences in trainings, coachings, or counseling sessions and 4) analyzing the change in TMIC-S results. Consequently, there would be a need to draft innovative development concepts and to evaluate them by means of longitudinal studies with TMIC-S.

5 Conclusion

The newly developed short version of the Test to Measure Intercultural Competence (TMIC-S) shows a satisfactory model fit and good psychometric properties in a German and Brazilian samples. As factor loadings and intercepts are invariant across groups, TMIC-S can be used to compare Germans and Brazilians concerning their intercultural competence in the future. The TMIC-S deals with the discussion on method effects in the assessment of intercultural competence by integrating two different methods to measure. Thereby, TMIC-S allows evaluations about the self-concept and the behavioral preferences of an individual in an intercultural context. At the same time, the TMIC-S is less time-consuming than the long version and thus reduces fatigue during the assessment process. The TMIC-S can be well applied during assessment or training settings with individuals who wish to interact in an intercultural context.

6 References

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C. et al. (2007). Cultural Intelligence: Its measure and effects on cultural judgment and decision making, cultural adaptation and task performance. *Management and Organization Review*, *3*, 335–371. doi:10.1111/j.1740-8784.2007.00082.x

Arasaratnam, L. A., & Doerfel, M. L. (2005). Intercultural Communication Competence: Identifying key components from multicultural perspectives. *International Journal of Intercultural Relations, 29,* 137–163. doi:10.1016/j.ijintrel.2004.04.001

Asparouhov, T., & Muthén, B. (2006). *Robust Chi Square Difference Testing with mean and variance adjusted test statistics*. Mplus Web Notes: No. 10. May 26, 2006.

Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 397–438. doi:10.1080/10705510903008204

Bass, B. M. (1990). *Bass & Stogdill's handbook of leadership: Theory, research & managerial applications* (3rd ed.). New York, NY: The Free Press.

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (2nd ed., pp. 21–77). Yarmouth, ME: Intercultural Press.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588–606. doi:10.1037/0033-2909.88.3.588

Bergemann, B., & Bergemann, N. (2005). *Interkulturelle Management-Kompetenz: Anforderungen und Ausbildung* [Intercultural management competence: Requirements and education]. Heidelberg, Germany: Physica-Verlag.

Bhawuk, D., & Brislin, R. (1992). The measurement of intercultural sensitivity using the

concepts of individualism and collectivism. *International Journal of Intercultural Relations, 16,* 413–436. doi:10.1016/0147-1767(92)90031-O

Bledow, R., & Frese, M. (2009). A situational judgment test of personal initiative and its relationship to performance. *Personnel Psychology, 62*, 229–258. doi:10.1111/j.1744-6570.2009.01137.x

Bolten, J. (2007a). Was heißt Interkulturelle Kompetenz? Perspektiven für die internationale Personalentwicklung. [What does intercultural competence mean? Perspectives for international personnel development] In V. Künzer & J. Berninghausen (Eds.), *Wirtschaft als interkulturelle Herausforderung* [Economy as an intercultural challenge] (pp. 21–42). Frankfurt am Main, Germany: IKO.

Bolten, J. (2007b). *Interkulturelle Kompetenz* [Intercultural competence]. Erfurt, Germany: Thüringer Landeszentrale für politische Bildung.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology 1*, 185–216. doi:10.1177/135910457000100301

Brislin, R. W. (1981). *Cross-cultural encounters. Face-to-face encounters*. New York, NY: Pergamon.

Brodbeck, F. C., Frese, M., Ackerblom, S., Audia, G., Bakacsi, G., et al. (2000). Cultural variation of leadership prototypes across 22 European countries. *Journal of Occupational and Organizational Psychology, 73,* 1–29. doi:10.1348/096317900166859

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

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. doi:10.1080/10705510701301834

Chen, G. M., & Starosta, W. J. (2000). *The development and validation of the intercultural*

communication sensitivity scale. Paper presented at the annual meeting of National Communication Association. Seattle, WA. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED447525

Cheung, F. M. (2009). The cultural perspective in personality assessment. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (pp. 44–56). New York, NY: Oxford University Press.

Church, A. T. (2001). Personality measurement in cross-cultural perspectives. *Journal of Personality, 69,* 979–1006. doi:10.1111/1467-6494.696172

Dauner, A. B. (2011). Die organisierende Funktion von Kompetenz – Entwurf eines zwei-Ebenen-Modells interkultureller Kompetenz [The organizing function of competence – Draft of a two-level-model of intercultural competence]. *Interculture Journal*, *10*(14), 25–46. Retrieved from http://www.interculture-journal.com/index.php/icj/article/view/115/194

Deardorff, D. K. (2004). *The identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States* (Unpublished doctoral dissertation). North Carolina State University, Raleigh, NC.

Earley, P. C., & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures.* Palo Alto, CA: Stanford University Press.

Eignor, D. R. (2013). *The standards for educational and psychological testing.* Washington, DC: American Psychological Association. doi:10.1037/14047-013

European Union (2014). *Erasmus Mundus Programme.* Retrieved from http://eacea.ec.europa.eu/erasmus_mundus/

Fantini, A., & Tirmizi, A. (2006). Exploring and assessing intercultural competence. *World*

Learning Publications, *1,* 1–75. Retrieved from http://digitalcollections.sit.edu/worlslearning_publications/1

Fritz, W., Graf, A., Hentze, J., Möllenberg, A., & Chen, G. M. (2005). An examination of Chen and Starosta's model of intercultural sensitivity in Germany and United States. *Intercultural Communication Studies, 14*(1), 53–64. Retrieved from http://www.uri.edu/iaics/content/2005v14n1/05%20Wolfgang%20Fritz,%20Andrea%20Graf,%20Joachim%20Hentze,%20Antje%20Mollenberg,%20&%20Guo-Ming%20Chen.pdf

Fritz, W., Möllenberg, A., & Chen, G. M. (2002). Measuring intercultural sensitivity in different cultural contexts. *Intercultural Communication Studies 11*(2), 165–176. Retrieved from http://www.uri.edu/iaics/content/2002v11n2/12%20Wolfgang%20Fritz%20&%20Antje%20Mollenberg.pdf

Gertsen, M. C. (1990). Intercultural competence and expatriates. *The International Journal of Human Resource Management*, *1*, 341–362. doi:10.1080/09585199000000054

Gudykunst, W. B. (1994). *Bridging differences: Effective intergroup communication* (2nd ed.). London, United Kingdom: Sage.

Gabrenya, W. K., Jr., Moukarzel, R. G., Pomerance, M. H., Griffith, H., & Deaton, J. (2011). *A validation study of the Defense Language Office Framework for Cross-cultural Competence*. Technical report, Defense Equal Opportunity Management Institute. Retrieved from http://www.defenseculture.org/Research/EmergeResearchFiles/AssessmentandMeasurement/ValidationDLOFramework.pdf

Gabrenya, W. K., Griffith, R. L., Moukarzel, R. G., Pomerance, M. H., & Reid, P. A. (2012). Theoretical and practical advances in the assessment of cross-cultural competence. In

D. D. Schmorrow & D. M. Nicholson (Eds.), *Advances in Design for Cross-Cultural Activities* (pp. 317–332). Boca Raton, FL: Taylor Francis.

Gawora, D., de Souza Ide, M. H., & Barbosa, R. S. (Eds.). (2011). *Traditionelle Völker und Gemeinschaften in Brasilien. Lateinamerika-Dokumentationsstelle* [Traditional peoples and communities in Brazil. Latin-America documentation centre]. Kassel, Germany: Kassel University Press.

Hall, E. T., & Hall, M. (1983). *Hidden differences. Studies in international communication. How to communicate with the Germans*. Hamburg, Germany: Gruner & Jahr.

Hammer, M. R. (2008). The Intercultural Development Inventory (IDI): An approach for assessing and building intercultural competence. In M. A. Moodian (Ed.), *Contemporary Leadership and Intercultural Competence: Understanding and Utilizing Cultural Diversity to Build Successful Organizations* (pp. 245–259). Thousand Oaks, CA: Sage.

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). The Intercultural Development Inventory: A measure of intercultural sensitivity. *International Journal of Intercultural Relations*, 27, 421–443. doi:10.1016/S0147-1767

Hofstede, G., & Bond, M. H. (1984). Hofstede's Culture Dimensions. An independent validation using Rokeach's Value Survey. *Journal of Cross-Cultural Psychology, 15*, 417–433. doi:10.1177/0022002184015004003

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70. Retrieved from http://www.jstor.org/stable/4615733

Hooper, A. C., Cullen, M. J., & Sackett, P. R. (2006). Operational threats to the use of SJTs: Faking, coaching, and retesting issues. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgement tests: Theory, measurement and application* (pp. 205–232). London, United Kingdom: Lawrence Erlbaum Associates.

House, R. J., Hanges, P. J., Ruiz-Quintanilla, S. A., Dorfman, P. W., Javidan, M., Dickson, M.W., & GLOBE (1999). Cultural influences on leadership: Project GLOBE. In W. Mobley, J. Gessner, & V. Arnold (Eds.), *Advances in global leadership, 1* (pp. 171–233). Stamford, CT: JAI Press.

IBM SPSS Statistics (Version 22.0 for Windows) [Computer software]. Armonk, NY: IBM.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park, CA: Sage.

Kelley, C., & Meyers, J. (1995). *The Cross-Cultural Adaptability Inventory (manual).* Minneapolis, MN: National Computer Systems.

Kealey, D. J., & Ruben, B. D. (1983). Cross-cultural personnel selection: Criteria, issues, and methods. In D. Landis & R. W. Brislin (Eds.), *Handbook of intercultural training (Vol.1): Issues in theory and design* (pp. 155–175). New York, NY: Pergamon Press.

Kim, Y. Y. (1992). Intercultural communication competence: A systems-thinking view. In W. B. Gudykunst & Y. Y. Kim (Eds.), *Readings on communicating with strangers: An approach to intercultural communication* (pp. 371–381). New York, NY: McGraw-Hill.

Kline, P. (1994). *An easy guide to factor analysis.* London, United Kingdom: Routledge.

Koester, J., & Olebe, M. (1988). The Behavioral Assessment Scale for Intercultural Communication Effectiveness. *International Journal of Intercultural Relations, 12*, 233–246. doi:10.1016/0147-1767(88)90017-X

Lambert, R. D. (Ed.) (1994). *Educational exchange and global competence*. New York, NY: Council on International Educational Exchange.

Leung, K., Ang, S., & Tan, M. L. (2014). Intercultural competence. *Annual Review of Organizational Psychology and Organizational Behavior, 1,* 489–519. doi:10.1146/annurev-orgpsych-031413-091229

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*, *37*, 426–441. doi:10.1108/00483480810877598

Loboda, J. (2003). *Entwicklung eines Tests zur Messung Interkultureller Handlungskompetenz (TIHK)* [*Development of a Test to Measure Intercultural Competence*]. Unpublished manuscript, Universitaet Regensburg, Regensburg, Germany.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Upper Saddle River, NJ: Prentice-Hall.

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 774–798). New York, NY: Wiley.

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2012). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*. Advance online publication. doi:10.1037/a0026913

Martin, J. N. (1993). Intercultural communication competence: A review. In R. L. Wiseman & J. Koester (Eds.), *Intercultural communication competence* (pp. 16–29). Newbury Park, CA: Sage.

Matsumoto, D., & Hwang, H. C. (2013). Assessing cross-cultural competence: A review of available tests. *Journal of Cross-Cultural Psychology, 44*, 849–873. doi:10.1177/0022022113492891

Matsumoto, D., LeRoux, J. A., Iwamoto, M., Choi, J., Rogers, D., Tatani, H., & Uchida, H. (2003). The robustness of the Intercultural Adjustment Potential Scale (ICAPS). *International Journal of Intercultural Relations, 27,* 543–562. doi:10.1016/S0147-1767(03)00053-1

Matsumoto, D., & van de Vijver, F. J. R. (Eds.). (2011). *Cross-cultural research methods in psychology.* New York, NY: Cambridge University Press.

McCrae, R. R., Terracciano, A., & 78 members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88,* 547–561. doi:10.1037/0022-3514.88.3.547

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2006). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, *60*, 63–91. doi:10.1111/j.1744-6570.2007.00065.x

McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment test: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, *9*, 103–113. doi:10.1111/1468-2389.00167

McDaniel, M. A., & Whetzel, D. L. (2005, June). *Situational Judgment Tests.* Paper presented at the 29th annual IPMAAC conference on personnel assessment, Orlando, FL.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543. doi:10.1007/BF02294825

Milfont, R., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*, 112–131. Retrieved from http://www.milfont.com/pubs_nz.php

Morris, M. A., & Robie, C. (2001). A meta-analysis of the effects of cross-cultural training on expatriate performance and adjustment. *International Journal of Training and Development, 5*, 112–125. doi:10.1111/1468-2419.00126

Mplus (Version 7.1 for Windows) [Computer software]. Los Angeles, CA: Muthén & Muthén.

Muritiba, P. M., Muritiba, S. N., Campanário, M., & de Albuquerque, L. G. (2010). International HR strategy in Brazilian technology multinationals. *BAR - Brazilian Administration Review*, *7*(4), 325–344. doi:10.1590/S1807-76922010000400002

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*. Conditionally accepted for publication in Psychometrika.

Müller, S., & Gelbrich, K. (2004). *Interkulturelles Marketing* [Intercultural marketing]. München, Germany: Vahlen.

Olson, C. L., & Kroeger, K. R. (2001). Global competency and intercultural sensitivity. *Journal of Studies in International Education, 5*, 116–137. doi:10.1177/102831530152003

Ruben, B. D. (1976). Assessing communication competency for intercultural adaptation. *Group & Organization Studies, 1*, 334–354. doi:10.1177/105960117600100308

Schacher, C. (2011). *Der Perspektivenwechsel in interkulturellen Konfliktsituationen* [Change of perspective in intercultural conflict situations] (Unpublished diploma thesis). Technische Universität Darmstadt, Darmstadt, Germany.

Sheahan, P. (2005). *Generation Y: Thriving and surviving with generation Y at work.* Prahran, Australia: Hardie Grant Books.

Sinicrope, C., Norris, J., & Watanabe, Y. (2007). Understanding and assessing intercultural competence: A summary of theory, research, and practice. *Second Language Studies, 26*, 1–58. Retrieved from http://www.hawaii.edu/sls/uhwpesl/26(1)/Norris.pdf

Schnabel, D., & Kelava, A. (2013, July). *Development and validation of the German Test to Measure Intercultural Competence (TMIC): A combined-method approach.* Paper presented at the 12th European Conference on Psychological Assessment, San

Sebastian, Spain.

Schnabel, D. (2013, September). *The effects of using therapeutic assessment in counselling students who are going abroad: Intercultural competence and self-related development.* Paper presented at the 12th World Congress of Psychological Counselling and Guidance, Istanbul, Turkey.

Schnabel, D., Kelava, A., Seifert, L., & Kuhlbrodt, B. (2014). Konstruktion und Validierung eines multimethodalen berufsbezogenen Tests zur Messung interkultureller Kompetenz [Development and validation of a job-related multimethod test to measure intercultural competence]. *Diagnostica*. Advance online publication. doi:10.1026/0012-1924/a000110

Schnabel, D., Kelava, A., & van de Vijver, F. J. R. (2014). *The effects of using collaborative assessment with students going abroad: Intercultural competence development, self-understanding, self-confidence, and stages of change.* Manuscript submitted for publication.

Schuler, H., & Prochaska, M. (2000). Entwicklung und Konstruktvalidierung eines berufsbezogenen Leistungsmotivationstests [Development and validation of an occupational achievement motivation inventory]. *Diagnostica, 46*, 61–72. doi:10.1026//0012-1924.46.2.61

Spitzberg, B. H., & Changnon, G. (2009). Conceptualizing intercultural competence. In D. K. Deardorff (Ed.), *The SAGE Handbook of Intercultural Competence* (pp. 2-52). Los Angeles, CA: Sage Publications.

Tamam, E. (2010). Examining Chen and Starosta's Model of Intercultural Sensitivity in a multiracial collectivistic country. *Journal of Intercultural Communication Research, 39,* 173–183 doi:10.1080/17475759.2010.534860

Thomas, A. (Ed.). (1991). *Kulturstandards in der internationalen Begegnung* [Cultural

standards in international encounters]. Saarbrücken, Germany: Breitenbach.

Thomas, A. (2003). Interkulturelle Kompetenz - Grundlagen, Probleme und Konzepte [Intercultural Competence – Basics, problems, and concepts]. *Erwägen, Wissen, Ethik, 14*(1)*,* 137–221. Retrieved from http://groups.uni-paderborn.de/ewe/index.php

Thomas, A., Kammhuber, S., & Layes, G. (1997). Interkulturelle Kompetenz. Ein Handbuch für internationale Einsätze der Bundeswehr [Intercultural competence. A manual for international missions of German Federal Armed Forces]. *Untersuchungen des Psychologischen Dienstes der Bundeswehr*, *32*, 7–190.

Thomas, A., Kinast, E.-U., & Schroll-Machl, S. (2006). Entwicklung interkultureller Handlungskompetenz von international tätigen Fach- und Führungskräften durch interkulturelle Trainings [Intercultural competence development of international operating skilled employees and leaders in intercultural trainings]. In K. Götz (Ed.), *Interkulturelles Lernen / Interkulturelles Training* [Intercultural learning / intercultural training] (pp. 91–114). München, Germany: Rainer Hampp Verlag.

Ting-Toomey, S. (1993). Communicative resourcefulness: An identity negotiation perspective. In R. L. Wiseman & J. Koester (Eds.), *Intercultural communication theory* (pp. 72–111). Newbury Park, CA: Sage.

Triandis, H. C. (1977). *Interpersonal behavior.* Monterey, CA: Brooks/Cole.

Unipark EFS Survey [Computer software]. Cologne, Germany: QuestBack GmbH.

Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research.* Newbury Park, CA: Sage.

Van de Vijver, F. J. R., & Leung, K. (2009). Methodological issues in researching intercultural competence. In D. K. Deardorff (Ed.), *The SAGE Handbook of Intercultural Competence* (pp. 404-418). Los Angeles, CA: Sage Publications.

Van Dyne, L., Ang, S., & Koh, C. (2008). Development and Validation of the CQS. In S.

Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence: Theory, measurement and application* (pp. 16–38). Armonk, NY: M. E. Sharpe.

Van Oudenhoven, J. P., & Van der Zee, K. I. (2002). Predicting multicultural effectiveness of international students: The Multicultural Personality Questionnaire. *International Journal of Intercultural Relations, 26*, 679–694. doi:10.1016/S0147-1767(02)00041-X

Weekly, J. A., & Ployhart, R. E. (2006). *Situational Judgment Tests: Theory, measurement and application.* London, United Kingdom: Lawrence Erlbaum Associates.

Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit [Comparative performance tests in schools – a controversial matter of course]. In F. E. Weinert (Ed.), *Leistungsmessungen in Schulen* (pp. 17–31). Weinheim, Gerrmany: Beltz.

Whetzel, D. L., McDaniel, M. A., & Nguyen, N. T. (2008). Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance, 21*, 291–309. doi:10.1080/08959280802137820

Yamazaki, Y., & Kayes, D. C. (2004). An experiential approach to cross-cultural learning: A review and integration of competencies for successful expatriate adaptation. *Academy of Management Learning and Education, 3*, 362–379. Retrieved from http://learningfromexperience.com/media/2010/08/Yamazaki_and_Kayes.pdf

Ziegler, M., & Bensch, D. (2013). Lost in translation: Thoughts regarding the translation of existing psychological measures into other languages. *European Journal of Psychological Assessment, 29*, 81–83. doi:10.1027/1015-5759/a000167

Zhong, M. (1998). Perceived intercultural communication competence in cross-cultural interactions between Chinese and Americans. *Critical Studies, 12*, 161–179. Retrieved from http://books.google.de/books?id=-1c5bSlZ5jUC&printsec=frontcover&hl=de

Appendix

Description of the Factors with Item Examples in German, Portuguese, and English

F	Description	Amount	L	Example items (SA)	SJT English with coding (1) = lowest value to (4) = highest value
SC	To put oneself in the position of another person during communication in order to understand them better; high sensibility for verbal and nonverbal communication aspects.	6 SA + 1 SJT	G	Ich weiß, wie sich andere Personen fühlen, ohne dass sie es mir sagen.	You are working together with a foreign delegation on a project. First of all, a meeting takes pla in order to discuss the further progression of the project and to set important objectives. You hav the impression that the project manager does not directly address which areas they will focus on during the project or talk about which points are most important to them. How are you most like to behave? I wait to get more information after the discussion. (1) I keep quiet until the end of the discussio and then ask the project manager to summarize the most important points again. (2) I listen attentively in order to recognize what the most important points are. (3) I closely observe how th project manager formulates their points and acts during the discussion. (4)
			P	Eu sei como as outras pessoas se sentem, sem que elas tenham que me dizer.	
			E	I know how other people feel without them having to tell me.	
IS	Purposeful collection of information about a foreign country or another culture.	3 SA + 1 SJT	G	Bei der Planung einer Reise ins Ausland, nutze ich unterschiedliche Informationsquellen.	You work for a company and you are going to be sent abroad for six months. How are you most likely to prepare for this? I prepare myself professionally and will get to know the culture when I get there. (1) I read abou the basic rules of behavior on the internet. (2) I read a travel guide and look at a map in order to be able to cope when I get there. (3) I stock up on books about the culture, the country and the language and also do some research on the internet. (4)
			P	Durante o planejamento de uma viagem ao exterior, eu uso várias fontes de informação.	
			E	When planning a trip abroad I use various sources of information.	
SZ	Establishing and maintaining contact	4 SA + 1 SJT	G	Ich nutze einen großen Teil meiner Freizeit, um Kontakte zu pflegen.	You move to a new city and do not yet know anyone. How are you most likely to behave in this situation?
			P	Eu uso uma grande parte do meu tempo livre para manter contatos.	

	with people from other cultures quickly and easily.		E	I use a large part of my free time in order to cultivate contacts.	I concentrate fully on work. (1) So that I don't feel lonely I have long phone calls with friends or family from my hometown during my free time. (2) I make an effort to be friendly to everyone I meet and therefore signal my interest in getting to know new people. (3) I try to make contact with people through various free time activities. (4)
GS	Having clear goals and being able to implement them consistently.	4 SA + 1 SJT	G	Wenn ich mir etwas vornehme, realisiere ich dies gewöhnlich auch.	You have made it your aim to successfully complete the project by the end of the next month. However, after a short time you realize that you have barely made any progress with your projec How are you most likely to behave in such a situation? I delay completion of the project. (1) I concentrate on the parts of the project which are going well. (2) I work on eliminating the aspects which are responsible for the delay. (3) I define what must achieve and when I must have achieved it by in order to come closer to my aims. (4)
			P	Se planejo algo, normalmente também costumo realisar isso.	
			E	When I plan something I usually then go on to achieve my aim.	
MI	Mediating between parties in order to achieve the greatest possible benefit from different approaches.	4 SA + 1 SJT	G	Ich bin gut darin, zwischen Personen mit gegensätzlichen Interessen zu vermitteln.	You have been sent by company headquarters to a subsidiary abroad where you are to manage a project. You quickly notice that the company headquarters follow quite different interests to the subsidiary abroad. What are you most likely to do? I recommend that company headquarters give up on the project as soon as possible. (1) I use all my resources to implement the interests of company headquarters. (2) I analyze which interests have a higher priority and support the most important ones. (3) I put in a lot of effort to mediate between the interests of the subsidiary and company headquarters. (4)
			P	Eu sou bom na mediação entre pessoas com interesses opostos.	
			E	I am good at mediating between people with conflicting interests.	
CIR	Intensively and constantly reflecting upon ones own cultural character.	4 SA + 1 SJT	G	Ich bemühe mich zu verstehen, inwiefern mein Verhalten kulturell geprägt ist.	You begin a new position with a company based abroad. You find your new job very interesting and on the whole are satisfied. The only thing which bothers you are the new working hours which are much different to what you are used to in your working life. Your colleagues don't seem to be bothered. How are you most likely to behave? I simply accept the new working hours so that I am not viewed negatively. (1) I look for a way t get
			P	Tento entender como o meu comportamento é determinado culturalmente.	
			E	I make an effort to understand to what extent my behavior is shaped by culture.	

		as near as possible to the working hours I would like. (2) I consider why the working hours are s different. (3) I think about why the working hours bother me and how I can deal with this in the future. (4)

Note. F = Factor; SA = Self-appraisal, SJT = Situational Judgment Test; L = Language; G = German; P = Portuguese; E = English (the English items were only created for a better understanding of the content); SC = Sensitivity in communication; IS = Information seeking; SZ = Socializing; GS = Goal setting; MI = Mediation of interests; CIR = Cultural identity reflection.

7 Tables

Table 1

Overview of several intercultural instruments

Criteria	BASIC	CCAI	CQS	IDI	INCA
Operationalization of Intercultural Competence	Display of Respect, Interaction Posture, Orientation to Knowledge, Empathy, Individualistic Roles, Relational Role Orientation, Task-related Role Orientation, Interaction Management, Tolerance of Ambiguity.	Emotional Resilience, Flexibility and Openness, Perceptual Acuity, Personal Autonomy	Metacognition, Cognition, Motivation, Behavior	Denial/Defense, Reversal, Minimization, Acceptance/Adaptation, Encapsulated Marginality	Tolerance of Ambiguity, Behavioral Flexibility, Communicative Awareness, Knowledge Acquisition, Openness, Empathy
Type of Characteristics	Trait-Competence-Mix	Stable Traits	Cultural Intelligence	Intercultural Sensitivity	Trait-Competence-Mix
Number of Methods (Type of Method)	1 (external assessment)	1 (self-appraisal)	1 (self-appraisal)	1 (self-appraisal)	3 (biographic questionnaire, role play, scenario)
Cronbach's Alpha	.82	.68–.90	.77–.84	.80–.85	.82–.89

Note. BASIC = Behavioral Assessment Scale for Intercultural Communication (Koester & Olebe, 1988); CCAI = Cross-Cultural Adaptability Inventory (Kelley & Meyers, 1995); CQS = Cultural Intelligence Scale (Van Dyne et al., 2008) IDI = Intercultural Development Inventory (Hammer, 2008; Hammer et al., 2003); INCA = Intercultural Competence Assessment (Fantini & Tirmizi, 2006). Adapted from "Development and validation of a job-related multimethod test to measure intercultural competence," by D. Schnabel, A. Kelava, L. Seifert, and B. Kuhlbrodt, 2014, *Diagnostica,* Advance online publication, p. 4. Copyright 2014 by Hogrefe Verlag Göttingen. Adapted with permission.

Table 2

Intercultural Competence Model of Schnabel et al. (2014)

	Communication	**Learning**	**Social Interaction**	**Self-knowledge**	**Self-management**	**Creating Synergies**
Definition	In an international context it is particularly important to be responsive to the person you are talking to and to be able to actively direct the conversation. In intercultural communication verbal as well as non-verbal aspects play an important role.	During cooperation with people from other cultures or during a stay abroad individuals are often faced with unknown situations. This requires the motivation of a person to extend their own knowledge and to perform intercultural important behavioral patterns. Persons are seen as being capable of learning, if they recognize that they have gaps in their knowledge and, as a consequence, invest time in improving their knowledge.	The building of interpersonal relationships is of great importance, particularly during a stay abroad. Relationships with other people positively influence our own well-being and can reduce or prevent stress and avoid a culture shock. Furthermore, a well-functioning network can support when it comes to achieving aims and satisfying needs.	Actively reflecting and thus understanding of the own cultural identity increases self-knowledge and positively influences the awareness of and also the successful interaction with other cultures.	A stay abroad or cooperation with people from other cultures involves some challenges, which must be dealt with. Problems can arise, which have to be solved. Circumstances for the achievement of aims are more demanding and the new working and living environment can cause stress. The existence of strategies, which make dealing with these challenges easier is therefore of great importance.	Different ways of working as well as diverse interests and approaches come together during intercultural cooperation. When a joint aim is to be achieved it is of great importance to have the ability to realize potential miss-understandings and lead a group towards common solutions.
Related	*Sensitivity in communication,*	*Willingness to use a foreign*	*Building professional networks,*	*Cultural identity awareness,*	*Goal setting, strategic*	*Mediation of different interests,*

Facets	*clarity in communication, flexibility in communication, perspective-taking in communication*	*language, willingness to learn, information seeking*	*socializing, integration in groups, building trusting relationships*	*cultural identity reflection*	*problem-solving*	*enabling productive collaborations*

Table 3

Loadings, Descriptive Statistics, and Reliabilities for the German and Brazilian TMIC-S (CFA Model)

Competence	Item	Loadings		*SE*		*M*		*SD*		α	
		G	B	G	B	G	B	G	B	G	B
Sensitivity in communication	SC1	.70	.70	.023	.025	4.80	4.44	1.01	1.05	.86	.77
	SC2	.74	.61	.021	.025	4.10	3.94	0.96	1.10		
	SC3	.76	.66	.021	.024	4.53	4.42	0.86	1.02		
	SC4	.77	.64	.021	.025	4.63	4.40	0.95	1.06		
	SC5	.85	.71	.020	.022	4.48	4.44	0.96	0.99		
	SC6	.72	.64	.023	.026	4.43	4.44	0.91	1.07		
	SC-SJT	.20	.10	.045	.051	2.37	2.09	0.89	0.97		
Goal setting	GS1	.75	.80	.028	.030	5.00	4.56	0.88	1.00	.78	.70
	GS2	.76	.69	.030	.029	4.86	4.65	0.92	1.01		
	GS3	.82	.51	.028	.035	4.92	4.51	0.93	1.05		
	GS4	.67	.60	.032	.033	4.69	4.59	0.91	1.02		
	GS-SJT	.26	.11	.049	.052	3.36	3.34	0.82	0.86		
Information seeking	IS1	.69	.63	.031	.035	5.02	4.71	1.10	1.03	.81	.68
	IS2	.82	.79	.028	.029	4.37	4.54	1.21	1.17		
	IS3	.89	.77	.030	.033	4.97	5.12	1.02	0.99		

	IS-SJT	.53	.29	.044	.059	2.96	2.78	1.21	1.31		
	MI1	.74	.67	.028	.025	4.41	4.15	0.94	1.18	.72	.65
	MI2	.71	.63	.027	.027	4.56	4.32	1.00	1.17		
Mediation of interests	MI3	.63	.58	.029	.027	4.40	4.27	0.88	1.02		
	MI4	.62	.66	.029	.026	4.63	4.30	0.91	1.08		
	MI-SJT	.23	.14	.061	.054	3.63	3.49	0.53	0.73		
	CIR1	.83	.82	.025	.027	4.35	4.17	1.32	1.23	.84	.73
	CIR2	.65	.77	.033	.029	4.39	4.24	1.27	1.20		
Cultural identity reflection	CIR3	.68	.71	.028	.031	4.20	4.19	1.23	1.26		
	CIR4	.72	.62	.029	.034	4.20	4.37	1.34	1.34		
	CIR-SJT	.40	.17	.049	.055	3.10	2.82	1.08	1.23		
	SZ1	.81	.79	.027	.026	4.56	4.21	1.05	1.19	.78	.77
	SZ2	.85	.66	.024	.033	4.05	3.65	1.19	1.31		
Socializing	SZ3	.84	.65	.026	.030	4.69	4.22	0.98	1.24		
	SZ4	.85	.80	.024	.027	4.14	3.99	1.16	1.28		
	SZ-SJT	.27	.27	.027	.047	3.45	3.10	0.81	0.86		

Note. G = Germany; B = Brazil; SC = Sensitivity in communication; GS = Goal setting; IS = Information seeking; MI = Mediation of interests; CIR = Cultural identity reflection; SZ = Socializing.

Table 4

Comparing Three Levels of Measurement Invariance for the German and Brazilian TMIC-S Version

	χ^2	*df*	$\Delta\chi^2$	RMSEA [CI 90%]	ΔRMSEA	CFI	ΔCFI	TLI	ΔTLI
Configural invariance	1742.01*	838		.038 [.035; .040]		.908		.898	
Metric invariance	1678.93*	863	46.874(25)*	.036 [.033; .038]	.002	.917	.009	.910	.012
Scalar invariance	1802.26*	882	241.12(19)*	.037 [.035; .040]	.001	.906	.011	.901	.009

Note. $N = 1806$; *$p < .01$; cut-off values for measurement invariance are: ΔCFI ≤ .010, ΔRMSEA ≤ .015, and Bonferroni-Holm corrected $p < .01$ for $\Delta\chi^2$ (Chen, 2007; Holm, 1979).

Table 5

Correlation of the Latent TMIC-S Factors with the Four Dimensions of the Cultural Intelligence Scale in the German and Brazil Samples

	1	2	3	4	5	6	7	8	9	10
1. Metacognitive CQ (MC)	–	.64*	.81*	.62*	.49*	.30*	.38*	.58*	.51*	.40*
2. Cognitive CQ (COG)	.77*	–	.69*	.59*	.50*	.25*	.40*	.58*	.48*	.39*
3. Motivational CQ (MOT)	.83*	.81*	–	.76*	.56*	.42*	.43*	.60*	.64*	.51*
4. Behavioral CQ (BEH)	.73*	.73*	.76*	–	.56*	.33*	.32*	.59*	.37*	.55*
5. Sensitivity in communication (SC)	.57*	.31*	.43*	.34*	–	.44*	.38*	.73*	.37*	.39*
6. Goal setting (GS)	.30*	.35*	.37*	.40*	.32*	–	.50*	.42*	.22*	.31*
7. Information seeking (IS)	.18*	.20*	.32*	.18*	.09	.12	–	.34*	.35*	.32*
8. Mediation of interests (MI)	.57*	.32*	.44*	.41*	.89*	.24*	.07	–	.39*	.62*
9. Cultural identity reflection (CIR)	.44*	.56*	.78*	.50*	.26*	.20*	.36*	.26*	–	.43*
10. Socializing (SZ)	.34*	.31*	.37*	.51*	.25*	.24*	.14*	.31*	.31*	–

Note. Intercorrelations for the Brazilian participants (n = 769) are presented above the diagonal, and intercorrelations for the German participants (n = 313)
are presented below the diagonal. *p < .01.

Table 6

Multigroup SEM Comparisons for All Six Factors and the External Criteria Intercultural Training Participation, Intercultural Involvement, and Experience Abroad in the German and Brazilian Samples

		$\Delta\chi^2$		*d*		*Mean*		*Standard Error*	
Factor	Subgroup	G	B	G	B	G	B	G	B
SC	Training	0.96	9.58*	0.11	0.34	0.09	0.30	0.09	0.10
	Involvement	8.23*	10.17*	0.31	0.37	0.25	0.36	0.09	0.11
	Abroad > 3 months	1.08	15.64*	-0.02	0.37	-0.02	0.37	0.08	0.10
GS	Training	1.12	1.13	0.04	0.08	0.03	0.04	0.09	0.06
	Involvement	2.84	2.25	0.20	0.12	0.16	0.12	0.09	0.11
	Abroad > 3 months	0.05	0.20	0.02	-0.04	0.02	-0.02	0.08	0.05
IS	Training	4.40*	0.82	0.24	0.10	0.21	0.10	0.10	0.11
	Involvement	21.64*	6.50*	0.55	0.26	0.49	0.28	0.10	0.12
	Abroad > 3 months	6.34*	1.58	0.25	-0.07	0.25	-0.07	0.10	0.10
MI	Training	4.08*	16.23*	0.21	0.44	0.17	0.39	0.09	0.10
	Involvement	7.35*	13.63*	0.39	0.47	0.31	0.43	0.09	0.10
	Abroad > 3 months	0.29	17.07*	0.05	0.43	0.04	0.39	0.08	0.09
CIR	Training	80.50*	14.33*	1.03	0.42	0.91	0.37	0.10	0.10
	Involvement	372.69*	33.32*	1.24	0.68	0.99	0.60	0.09	0.10
	Abroad > 3 months	39.96*	15.17*	0.64	0.38	0.58	0.35	0.09	0.09
SZ	Training	15.86*	47.47*	0.44	0.76	0.39	0.74	0.10	0.11
	Involvement	35.44*	41.95*	0.64	0.72	0.57	0.64	0.10	0.10
	Abroad > 3 months	24.62*	30.56*	0.46	0.53	0.46	0.48	0.10	0.09

Note. Subgroups involve: Training versus no training, involvement versus no involvement, and experiences abroad more than 3 months versus less than 3 months. $\Delta\chi^2$ refers to the delta chi square test with one degree of freedom; *d* refers to Cohen's *d*. In the German sample SEM-comparisons were computed without item MI-SJT, because one of the subgroups did not contain value 1 of the categorical variable. German sample: $N = 724$, training = 237, no training = 487, involvement = 456, no involvement = 268, > 3 months = 420, < 3 months = 304; Brazilian sample: $N = 769$, training = 558, no training = 211, involvement = 604, no involvement = 164, > 3 months = 436, < 3 months = 333. G = German sample; B = Brazilian sample; SC = Sensitivity in communication; GS = Goal setting; IS = Information seeking; MI = Mediation of interests; CIR = Cultural identity reflection; SZ = Socializing. *$p < .05$.