

Robustness and cross-cultural equivalence of the Cultural Intelligence Scale (CQS)

Joost Bueker

*Institute for Management Research,
Radboud University, Nijmegen, The Netherlands*

Olivier Furrer

*Department of Marketing,
Universite de Fribourg Faculte des sciences economiques et sociales,
Fribourg, Switzerland, and*

Tanja Peeters Weem

*Department of HR Transformation, Deloitte Netherlands,
Amsterdam, The Netherlands*

Abstract

Purpose – The purpose of this paper is to assess the cross-cultural equivalence of the four-dimensional 20-item Cultural Intelligence Scale (CQS) and the two-dimensional 12-item cultural intelligence (CQ) short scale. Furthermore, the study elaborates on the results by discussing the differences between culturally equivalent and culturally non-equivalent items.

Design/methodology/approach – Data gathered from 607 students with a Chinese or Dutch background and mature international experience serve to test the cross-cultural equivalence of the CQS.

Findings – This study addresses the lack of clarity concerning the cross-cultural equivalence of the CQS in the extended domain of empirical research involving CQ. Furthermore, the consequences of the cultural equivalence tests are discussed.

Practical implications – Comparing CQ scores across cultures is only meaningful with the use of the adjusted, two-dimensional scale. Practitioners must be aware of the emic-etic character of the measurement instrument they use.

Originality/value – This study addresses the lack of clarity concerning the cross-cultural equivalence of the CQS in the extended domain of empirical research involving CQ. Furthermore, the consequences of the cultural equivalence tests are discussed.

Keywords Culture, Cultural intelligence, CQS, Emic, Etic, Measurement equivalence

Paper type Research paper

Introduction

The globalization and liberalization of trade and service have induced growth among multinational corporations (MNCs) (Bueker and Poutsma, 2010), increasing foreign direct investment and global mobility, such that employees with varied cultural backgrounds come in closer contact through working relationships (Ang *et al.*, 2006). Working with people from different cultural backgrounds imposes new demands on employees and managers and requires new cross-cultural competencies (Chong, 2008; Johnson *et al.*, 2006; Morley and Cerdin, 2010). Because people vary in their ability to develop cross-cultural competencies (Ang and Van Dyne, 2008), MNCs need ways to identify employees and managers who meet the job requirements of increasingly diverse cultural contexts. Globally competent managers are scarce, and effective personnel selection and training demands a cross-culturally valid instrument to measure the relevant competences.



Accordingly, Ang *et al.* (2007) proposed a Cultural Intelligence Scale (CQS) to measure people's cross-cultural competencies. For such a measurement instrument to be useful for MNCs' managers in their selection and training of employees across the world, it needs to be robust and cross-culturally equivalent.

Ang *et al.* (2007) conducted initial tests of the cross-cultural invariance of the CQS. However, they seemed to have neglected some important aspects of invariance. For example, they did not report any assessment of scalar equivalence, which is critical for comparing means across country samples (Steenkamp and Baumgartner, 1998).

Furthermore, while the convergent validity of the scale has been repeatedly validated in recent empirical studies, discriminant validity between the four dimensions of the CQS scale appears to be problematic (see Bückner *et al.*, 2015, for a review). To remedy this issue, Bückner *et al.* (2015) proposed and found empirical support for a two-dimensional model of cultural intelligence (CQ). Although the CQS, as a measurement tool for CQ, might seem promising (see Ang *et al.*, 2007), in several studies the four dimensions of the construct do not appear to be distinguishable (Ward *et al.*, 2009; Lee and Sukoco, 2010; Fischer, 2011).

Despite the limited assessment of its cross-cultural measurement equivalence and potential lack of discriminant validity between its dimensions, the CQS has been used in studies that compare CQ across countries and cultures (e.g. Imai and Gelfand, 2010) or use pooled samples from multiple countries or cultures (e.g. Groves and Feyerherm, 2011). In single-country studies, no assessments confirm the validity of the CQS (e.g. Vedadi *et al.*, 2010). Without formal tests of the cross-cultural measurement invariance of the CQS, the validity of these results is questionable (Steenkamp and Baumgartner, 1998). Therefore, this study seeks to test the cross-cultural measurement equivalence of the four-dimensional and two-dimensional versions of the CQS, including their scalar equivalence and discriminant validity, across respondents from the Netherlands and China. As far as we know, this is the first study that tests all aspects of cultural equivalence of the CQS.

Assessing the cross-cultural measurement invariance of the CQS should contribute to CQ literature, as well as help establishing less ambiguous interpretations of the outcomes of empirical research on CQ. Several studies (e.g. Ward *et al.*, 2009) were unable to demonstrate the incremental validity of CQ to predict adaptation outcomes above personality and cognitive ability. There is no clear support for the mediating role of motivational CQ between personality and general adjustment (Ward and Fischer, 2008). Ward *et al.* (2011) found non-significant effects of CQ on performance and Lee and Sukoco (2010) found no direct effect of CQ on performance. These non-significant results might be due to measurement issues or a lack of measurement equivalence.

By determining the cross-cultural validity of the CQS, our study also contributes to the emic-etic debate in international management literature (Ng and Earley, 2006). Etic research seeks to make generalizations across cultures; while emic research documents principles of behavior in a specific culture, without requiring cross-cultural measurement invariance. As Triandis and Marin (1983) observe, most construct measures adopt an etic approach and assume item content to be relevant in every culture. However, some constructs miss this etic condition and as a result may be interpreted differently across cultures, such as the meaning of general intelligence (IQ), which varies across cultures (Triandis, 2006), suggesting emic characteristics.

This study reveals that some CQS items vary across cultures and their content is emic, then proposes a cross-culturally equivalent and valid scale with etic properties. The study results offer MNCs insights into selecting and training MNC managers to

develop cross-cultural competencies. By establishing the measurement equivalence of a two-dimensional version of CQS, this investigation enhances the trustworthiness of selection processes and enables comparisons of candidates across countries.

Literature review

CQ

Earley and Ang (2003) propose CQ, to capture a person's capability to adjust effectively across cultural contexts. This conceptualization builds on the idea of "multiple intelligences" (Gardner, 1983). However, most types of intelligence are culturally specific and lose meaning when extended beyond their original context (Ng and Earley, 2006). For example, western cultures link intelligence to rapid, correct judgments; in many African cultures, intelligence relates to behavior that supports the demands of the elders (Triandis, 2006), and in traditional Asian cultures, intelligence implies some degree of self-insight and self-control (Yang and Sternberg, 1997). However, CQ, by definition, is an etic construct (Ng and Earley, 2006), to indicate people's capabilities to grasp, reason, and behave in situations characterized by cultural diversity (Ang *et al.*, 2007), with seemingly the same meaning across cultures. Two different studies on CQ were developed: Earley and Ang's (2003) study defined CQ as a four-dimensional model and developed the CQS (Ang *et al.*, 2006, 2007). Thomas *et al.*'s (2008) study defined a three dimensional CQ (leaving out the motivational component) construct and developed the SFCQ (Thomas *et al.*, 2015).

The so far most often used four-dimensional conceptualization of CQ (Ang *et al.*, 2006) includes a metacognitive form that refers to a person's cultural consciousness and awareness of cultural cues during interactions with people from other cultural backgrounds, such that people with metacognitive CQ question cultural assumptions, reflect on assumptions, and develop cultural knowledge and skills during intercultural interactions (Ang and Van Dyne, 2008). Cognitive CQ is based on knowledge of norms, practices, and conventions in different cultural settings, acquired through education and personal experience (Ang and Van Dyne, 2008), including knowledge of economic, legal, and social systems, as well as value systems. In the motivational form, learning centers on effective functioning in situations characterized by cultural differences, so people with high-motivational CQ are interested in cross-cultural situations and confident of their personal cross-cultural effectiveness (Ang *et al.*, 2007), related to a high level of self-efficacy (Ng and Earley, 2006). Finally, behavioral CQ refers to the capability to exhibit appropriate verbal and non-verbal behavior during interactions, such as culturally appropriate words, tones, gestures, and facial expressions (Ang *et al.*, 2006, 2007).

The CQS has appeared in an increasing number of empirical studies in various cultures across the world, as detailed in the Appendix. However, several studies use only selected dimensions of this scale though (e.g. Lee and Sukoco, 2010; MacNab and Worthley, 2012) and other studies only use CQ as a second order construct (Lee and Sukoco, 2010; Bückner *et al.*, 2014). As research about the dimensional structure of CQ still shows inconsistencies, it is not far-fetched to perceive different broader dimensions in the CQ construct, e.g. a cognitive one and an action-focussed one.

Bückner *et al.* (2015) found that the CQ construct is most likely composed of two dimensions: internalized cultural knowledge intelligence (ICK intelligence), which regroups items from the original metacognitive (MC) and cognitive (COG) dimensions, and effective cultural flexibility intelligence (ECF intelligence), which regroups items from the original motivational (MOT) and behavioral (BEH) dimensions. The two

dimensions, which exhibit discriminant validity, reflect the two sides of CQ: on the one side, the cognitive character and on the other side, the dynamic, action-oriented character of CQ.

Measuring CQ across cultures

In developing the CQS, Ang *et al.* (2007) tested the scale for cross-cultural psychometric properties and for the validity and reliability of the underlying construct but did not assess measurement equivalence at a level that would permit cross-cultural comparisons of means or allow for pooling data across samples. An assessment of cross-cultural invariance of any scale used across cultures is critical (Van de Vijver and Leung, 2000), because biases might occur if item scores for a particular construct do not correspond to differences in the underlying construct dimensions (Van de Vijver and Tanzer, 2004). This is part of Riordan and Vandenberg's (1994) two concerns: first, using different frames of reference by the diverse cultural groups make cultural comparisons impossible because scores on the instrument refer to different constructs for each group; and second, to what extent do respondents calibrate the intervals anchoring the measurement continuum in the same manner; put differently, are scale intervals perceived identically across groups. As an example of the latter, a 5 on a five-point Likert scale may be perceived as an extreme score in China and preferably not used while this same 5 may give no problems in use in the Netherlands. As a result, data will be interpreted inappropriately. Although cross-cultural invariance is required when using a construct across cultures (Steenkamp and Baumgartner, 1998; Van de Vijver and Leung, 2000), many empirical studies have already used the CQS in different cultures (see the Appendix). As Vandenberg and Lance (2000, p. 6) caution: "If not tested, violations of measurement equivalence assumptions are as threatening to substantive interpretations as is an inability to demonstrate reliability and validity."

Measurement invariance is a matter of degree, which varies with the study purpose (Van de Vijver and Leung, 2000). Three levels of cross-cultural invariance are particularly critical for applied cross-cultural research (Steenkamp and Baumgartner, 1998):

- (1) Configural, or whether the same items measure the same factors across cultures; a study that only aims to confirm the basic structure of a construct may use this confirmation.
- (2) Metric, or whether the factor loadings are equal across cultures due to the equivalence of the slopes in the regression of an item on the latent variable. Metric invariance can establish that a correlational analysis is valid, because the items measure the latent variables equally well across cultures. A study that examines relationships between constructs across cultures requires metric invariance.
- (3) Scalar, which indicates that in addition to equal factor loadings, the intercepts of the regression of an item on the latent variable are equal across cultures, and people across cultures use scale magnitudes in the same way (Robert *et al.*, 2006). Differences or similarities in factor means result from differences in item scores and intercepts or factor loadings.

Reaching a level of scalar invariance is necessary if the research objective is to compare means across cultures and pool raw data (Cheung and Rensvold, 2002), and is a prerequisite for interpreting differences in CQ. However, full scalar invariance might

not be necessary for the further tests of invariance to be meaningful, provided at least one item is invariant (i.e. partial scalar invariance) (Steenkamp and Baumgartner, 1998). Higher level of invariance, including latent mean invariance is not required, because cross-cultural differences in the level of CQ can be expected. If partial scalar invariance is not achieved, it is still possible to compare means across cultures and pool data, but only after that data have been standardized (Fischer, 2004). However, standardization procedures are often problematic, as they might mask some meaningful differences and results might be difficult to interpret (Fischer, 2004).

Review of empirical studies regarding CQ

Prior CQ literature features three main types of empirical studies (see the Appendix).

Comparative studies. Many CQ studies seek to measure and compare CQ scores across two or more countries. For example, Ang *et al.* (2007) measure CQ scores among Singaporean and US respondents, and Imai and Gelfand (2010) compare scores between Caucasian/white and Asian-American respondents, as well as between US and East-Asian respondents. Meaningful comparisons of mean scores across cultures demand at least partial scalar equivalence (Steenkamp and Baumgartner, 1998). However, many comparative studies do not assess scalar equivalence. Ang *et al.*'s (2007) initial study only reports tests of configural, metric, and covariance equivalence, not scalar equivalence. Imai and Gelfand (2010) do not report any cross-cultural measurement invariance tests, arguing that Ang *et al.* (2007) already demonstrated the psychometric properties of the CQS. However, this application of CQS occurred in cultures for which no tests have been conducted. Without proper measurement tests, these results are questionable; without scalar invariance, we cannot know if differences in CQ scores are actual cross-cultural differences or due to variations in scale uses, such as acquiescence or extreme response style biases (Van de Vijver and Leung, 2000).

Single-country/culture studies. A second group of studies uses the CQS in single-country research (see the Appendix). They investigate CQ in new countries but do not test if the meaning of the CQS items remains the same. Even configural invariance might not be sufficient to use an existing scale in a new country (Van de Vijver and Leung, 2000), because the scale must be equally reliable in the new country. For example, Vedadi *et al.* (2010) use the CQS to test the relationship between CQ and need for achievement among a sample of Iranian managers, without assessing whether CQ has the same meaning in Iran as in Singapore or the USA, the two original countries from which the scale was developed. Chen *et al.* (2011) study the role of CQ and its effect on the performance of Philippine laborers working in Taiwan, without assessing whether the content of the CQS items has the same meaning as in the original studies. Using a survey in a new culture or country implies a pseudo- or imposed etic approach (Triandis and Marin, 1983) and the assumption that the item content is relevant in every culture (Robert *et al.*, 2006). However, if the construct relevance of an item differs across cultures, the measure refers to something conceptually different than the original study (Van de Vijver and Leung, 2000). Therefore, the equivalence of the scale needs to be assessed every time it is used in a new cultural context.

Mixed/pooled sample studies. A third category of empirical CQS studies analyzes data from pooled samples containing respondents from multiple cultures or countries (see the Appendix). The results might be biased if respondents from different cultures use the CQS differently (Van de Vijver and Tanzer, 2004). One response bias features a systematic tendency to distort responses, such that observed scores do not relate to the respondent's

true score, because the respondent selects either extreme or overly modest answers or shifts responses to either end of the scale. Such response biases are influenced by cultural values and therefore can exert important cultural influences on data (Smith, 2004). Thus, pooling together culturally biased data is likely to distort results. Measurement equivalence and the absence of bias must exist to pool data from multiple cultures. If (even partial) scalar equivalence is not achieved, but configural and metric equivalence are, the data can be standardized to control for response bias and provide meaningful comparisons (Fischer, 2004). However, no mixed or pooled sample studies report cross-cultural equivalence tests or standardization procedures. For example, Ward *et al.* (2011) examine cross-cultural adaptation problems of students from 25 countries. Similarly, Groves and Feyerherm (2011) pool data from a diverse sample of respondents, without any assessment of measurement equivalence or response bias. By pooling data from multiple cultures, these authors implicitly assume high levels of cross-cultural measurement equivalence and uniform bias across cultures, yet such an analysis demands confirmation of scalar invariance (Steenkamp and Baumgartner, 1998; Van de Vijver and Leung, 2000).

Method

Sample selection and description

This study meets the challenge to assess the psychometric properties of the CQS in two countries, China and the Netherlands. These two countries exhibit large cultural distance (Hofstede *et al.*, 2010), have different institutions (Huang, 2006), and belong to distinct cultural clusters (Gupta *et al.*, 2002), which provides a stronger test of measurement invariance, because cultural distance is likely to render cross-cultural equivalence more difficult to achieve (Van de Vijver and Hambleton, 1996).

Data were collected in the same way (online) and at the same time across cultures to ensure data collection equivalence (Hult *et al.*, 2008). Collection of Chinese respondents with work or living experience outside China started from a Dutch University by two Chinese (PhD) students. A Dutch student with study experience abroad found Dutch respondents who had also been abroad for study or work (internship). What followed was a mixed sampling process making use of snowball sampling on one hand and using personal networks and various communities of (Chinese) students in different countries on the other hand within the same time period. Snowball sampling has often been used to enable access to previously hidden populations which in this study was partly true for the Chinese students. Also calls for respondents were placed on several formal networks of master students and PhD students at the Dutch University. Earlier research has shown that for gaining understanding and knowledge about behavioral processes student samples can be useful sources (Bernstein *et al.*, 1975; Barr and Hitt, 1986; Greenberg, 1987).

Participants included 299 Dutch and 308 Chinese students with experience living overseas for at least three months, which should ensure sufficient knowledge of a foreign culture (Crowne, 2013) and enhances their maturity (Frisch, 1990; Van Hoof and Verbeeten, 2005). To minimize language biases, both the Chinese and Dutch versions were translated, from the original English, following recommended translation-back translation procedures (Brislin, 1986). The Chinese translation was available from Ang *et al.* (2007), and the Dutch version was translated by two bilingual speakers. As the descriptive characteristics in Table I reveal, participants averaged 27.69 years of age ($SD = 6.03$) in the Chinese sample and 24.06 years of age ($SD = 4.18$) in the Dutch sample. Female respondents represented 65.2 percent of the Dutch respondents and 51.3 percent of the Chinese sample. Both samples consisted of respondents with

Table I.
Sample
characteristics

	China		The Netherlands	
	Mean	SD	Mean	SD
Age	27.69	6.03	24.06	4.18
Gender (% female) ^a	51.3		65.2	
Education	3.94	0.95	3.73	0.63
1 = high school/vocational	3.9		1.0	
2 = bachelor	28.2		32.8	
3 = master/MBA	34.7		58.1	
4 = PhD	33.1		8.1	
Time abroad	4.45	1.49	2.50	1.04
1 = less than 3 months	7.5		8.8	
2 = 3-6 months	3.9		54.4	
3 = 6-12 months	7.8		23.3	
4 = 1-3 years	31.5		8.1	
5 = 3-5 years	15.9		2.7	
6 = more than 5 years	33.4		2.7	
Contact frequency	2.90	0.74	3.76	0.49
1 = seldom	2.6		0	
2 = occasionally	25.3		2.7	
3 = often	51.9		18.6	
4 = all the time	20.1		78.7	

Notes: $n_{\text{(China)}} = 308$, $n_{\text{(The Netherlands)}} = 299$. ^aThere are three missing values for gender in the Netherlands

academic profiles. The samples differed in time spent abroad though: whereas only 13.5 percent of Dutch respondents stayed abroad for one year or more, in the Chinese sample, 80.8 percent did so. Contact frequency also varied, such that in the Dutch sample, 97.35 percent of the respondents had intensive contact with foreigners (“often” or “all the time”), but only 72 percent of the Chinese respondents noted this frequency.

Respondents could drop out of the survey at any time without any penalty for incompleteness. Thus, not all respondents provided completed responses, and only data from complete questionnaires were analyzed. The completion of the questionnaire suggests the self-selection of more motivated respondents, so inefficient effort is unlikely (Huang *et al.*, 2012). In addition, respondents took between 12 and 17 minutes to complete the survey, indicating satisfactory attention levels. Despite some differences in the composition of the two samples, data collection equivalence is likely sufficiently strong (Hult *et al.*, 2008).

Measures

The measure of CQ relied on the 20-item scale developed by Ang *et al.* (2007), with four dimensions: MC (four items; e.g. “I check the accuracy of my cultural knowledge as I interact with people from different cultures”), COG (six items; e.g. “I know the rules for expressing non-verbal behaviors in other cultures”), MOT (five items; e.g. “I am confident that I can get accustomed to the shopping conditions in a different culture”), and BEH (five items; e.g. “I alter my facial expressions when a cross-cultural interaction requires it”).

Data analysis procedure

To assess the psychometric properties and cross-cultural equivalence of the Ang *et al.* (2007) four-dimensional CQS (20 items) and the two-dimensional adapted version (12 items) proposed by Bückner *et al.* (2015). The analytical strategy began with the

assessment of the scales' reliability, using Cronbach's α (> 0.70). After that, confirmatory factor analyses (CFA), using AMOS was used to assess the proposed four- and two-factor models. Following common practice (e.g. Byrne, 2001; Hu and Bentler, 1999), we used the multiple indicators to assess model fit. Important indices are root mean square error of approximation (RMSEA), the non-normed fit index (NNFI), and the comparative fit index (CFI). The CFA test featured the multiple-factor models, rather than the dimensions separately, to assess discriminant validity (Hult *et al.*, 2008). Maximum likelihood estimation procedures were appropriate, because the data did not strongly violate multivariate normality assumptions (McDonald and Ho, 2002). To assess convergent validity, factor loadings were examined (> 0.50) for each country separately. The discriminant validity test began with Fornell and Larcker's (1981) procedure, comparing the square root of the average variance extracted (AVE) with the correlations of the CQS dimensions (Voorhees *et al.*, 2016). The test of the measurement equivalence of the CQS used multigroup CFA (Byrne, 2001), one of the most popular techniques (De Beuckelaer, 2005). In addition to the four-step procedure to test configural, metric, scalar, and latent mean invariance (Steenkamp and Baumgartner, 1998), two CFIs evaluated the difference between the nested models: the χ^2 difference test ($\Delta\chi^2$) and the changes in the Δ CFI, which is relatively unaffected by sample size (Cheung and Rensvold, 2002).

Results

The assessment of the reliability of the four CQ dimensions first addressed the two samples. In the Dutch sample, the acceptable Cronbach's α ranges from 0.75 to 0.81; in the Chinese sample, three α 's fall below the 0.70 threshold: 0.68 for BEH and 0.69 for MC and MOT (see Table II). The CFA results in both samples indicate poor fit for the four-dimensional CQ. For the Chinese and Dutch samples, respectively, the RMSEA scores are 0.070 (90 percent confidence interval (CI): 0.062-0.079) and 0.072 (90 percent CI: 0.064-0.081). The other fit indexes also are less than satisfactory ($\chi^2/\text{df} = 2.51$ and 2.57; standardized root mean square residual (SRMR) = 0.061 and 0.061; (NNFI) = 0.847 and 0.869; CFI = 0.868 and 0.887 for the Chinese and Dutch models, respectively; Table III). In terms of convergent validity, the factor loadings of four items in the Chinese sample and one item in the Dutch sample do not reach the required 0.50 values. To assess discriminant validity, we used Fornell and Larcker's (1981) test. The result of the test indicates a lack of discriminant validity in both samples. The comparison of the square root of the AVE with the correlations between the respective CQS dimensions demonstrates a lack of discriminant validity for all four dimensions in the Chinese sample, as well as for MC and COG and for MC and BEH in the Dutch sample (Table IV).

To assess if this lack of discriminant validity could be attributed to social desirability responding (Johnson and Van de Vijver, 2003), we looked at the correlations between social desirability and the CQS items. To measure social desirability, we used the MC2 version of the Marlowe-Crowne social desirability scale (Strahan and Gerbasi, 1972). The Chinese data show correlations with CQS items that are significant but smaller than 0.20 and thus sufficiently low (Watkins, 1996), with the exception of MC3, which reaches a value of 0.21. For the Dutch data, all correlations are smaller than 0.20. Thus, we do not expect that social desirability responding caused the lack of discriminant validity between the CQS dimensions.

Alternatively, the lack of discriminant validity might indicate a lack of configural invariance of the four-factor CQ model. Given this finding, we tested the four-dimensional model and compared it to the two-dimensional model as proposed by Bückner *et al.* (2015)

Table II.
CFA standardized
factor loadings and
cronbach's α

	China				The Netherlands			
	Four-dimensional model		Two- dimensional CQS model		Four-dimensional model		Two- dimensional CQS model	
	MC COG	MOT BEH	ICK ECF	ECF ICK	MC COG	MOT BEH	ICK ECF	ICK ECF
MC1	0.66		0.52		0.74		0.50	
MC2	0.39		na		0.61		na	
MC3	0.85		0.75		0.80		0.60	
MC4	0.55		na		0.49		na	
COG1	0.69		0.67		0.54		0.55	
COG2	0.51		0.52		0.50		0.49	
COG3	0.76		0.79		0.75		0.75	
COG4	0.70		0.67		0.68		0.67	
COG5	0.66		0.62		0.69		0.70	
COG6	0.62		na		0.76		na	
MOT1	0.46		na		0.69		na	
MOT2	0.70		0.77		0.66		0.50	
MOT3	0.65		0.65		0.63		0.56	
MOT4	0.46		na		0.76		na	
MOT5	0.50		na		0.67		na	
BEH1		0.43				0.63		na
BEH2		0.51				0.68		0.65
BEH3		0.59				0.58		0.56
BEH4		0.64				0.76		0.64
BEH5		0.63				0.71		na
Cronbach's α	0.69	0.81	0.68	0.83	0.71	0.81	0.81	0.72

Model	χ^2	df	χ^2/df	RMSEA	SRMR	NNFI	CFI	ΔCFI	$\Delta\chi^2$	Δdf	p-value
China (original four dimensions)	411.0	164	2.51	0.070 (0.062–0.079)	0.061	0.847	0.868	–	–	–	–
The Netherlands (original four dimensions)	420.7	164	2.57	0.072 (0.064–0.081)	0.061	0.869	0.887	–	–	–	–
China (two dimensions)	88.3	50	1.77	0.050 (0.032–0.065)	0.045	0.956	0.966	–	–	–	–
The Netherlands (two dimensions)	124.9	50	2.50	0.071 (0.055–0.087)	0.056	0.909	0.931	–	–	–	–
Configural invariance (without equality constraints)	213.3	100	2.13	0.043 (0.035–0.051)	0.045	0.933	0.949	–	–	–	–
Full metric invariance (equal factor loadings)	244.0	109	2.24	0.045 (0.038–0.053)	0.052	0.926	0.939	0.010	30.7	9	0.000
Partial metric invariance ^a	221.6	106	2.09	0.042 (0.035–0.050)	0.046	0.935	0.948	–0.001	8.3	6	0.216
Full scalar invariance (equal factor loadings and intercepts)	285.1	114	2.50	0.050 (0.043–0.057)	0.047	0.911	0.923	0.025	63.5	8	0.000
Partial scalar invariance ^b	232.0	112	2.07	0.042 (0.034–0.050)	0.047	0.936	0.946	0.002	10.4	6	0.109
Latent mean invariance (equal latent means)	268.1	114	2.35	0.047 (0.040–0.055)	0.049	0.920	0.931	0.015	36.1	2	0.000

Notes: ^aThe factor loadings of COG5, BEH2, and BEH4 were not invariant across countries and were released; ^bconsistent with the partial metric invariant model, the intercepts of the items with variant factor loadings and the intercepts of COG4 and MC3 that were not invariant were released

Table III.
Measurement
models' fit indices

JGM 4,3	MC	COG	MOT	BEH
<i>China</i>				
MC	0.64			
COG	0.82***	0.66		
MOT	0.69***	0.69***	0.56	
BEH	0.64***	0.54***	0.71***	0.56
<i>The Netherlands</i>				
MC	0.67			
COG	0.71***	0.66		
MOT	0.57***	0.54***	0.69	
BEH	0.73***	0.63***	0.55***	0.68

Table IV. Correlation matrix

Notes: The square root of the average variance extracted is on the diagonal. *** $p < 0.001$

to achieve discriminant validity and configural equivalence (Caramellia and Van de Vijver, 2013; Hult *et al.*, 2008; Steenkamp and Baumgartner, 1998).

Then, we assessed the psychometric properties and cross-cultural invariance of the 12 items, two-dimensional model proposed by Bückner *et al.* (2015). In the Dutch sample, the acceptable Cronbach's α are 0.81 and 0.72, for ICK and ECF, respectively, and in the Chinese sample α 's are 0.83 and 0.71, respectively (see Table II). The two-dimensional CFA model provides improved fit indexes: RMSEA of 0.050 (90 percent CI: 0.032-0.065) for the Chinese sample and 0.071 (90 percent CI: 0.055-0.087) for the Dutch sample. The other fit indexes also improved in both samples, reaching acceptable levels ($\chi^2/df = 1.77$ and 2.50; SRMR = 0.045 and 0.056; NNFI = 0.956 and 0.909, CFI = 0.966 and 0.931 for the Chinese and Dutch samples, respectively). In terms of convergent validity, the factor loadings of all items reach the required 0.50 values, except one in the Chinese sample one item in the Dutch sample. Fornell and Larcker's (1981) test supports the discriminant validity of the two dimensions in both country samples. For the Chinese sample, the correlation of ICK and ECF is 0.57, less than the square root of the respective AVEs (i.e. 0.66 and 0.61). Similarly, for the Dutch sample, the correlation is 0.59, less than or equal to the square roots of the respective AVEs (0.61 and 0.59).

The next test entailed assessing the measurement equivalence of the two-dimensional model, in four steps using multigroup CFA. First, regarding configural invariance with an unconstrained model, the model fit the data relatively well ($\chi^2/df = 2.11$; RMSEA = 0.043 (90 percent CI: 0.035-0.051); SRMR = 0.045; NNFI = 0.933; CFI = 0.949; Table IV). Second, to assess metric invariance, the factor loadings were constrained to be equal across the two samples. The constrained model was still satisfactory but statistically worse than the unconstrained model ($\Delta\chi^2 = 30.7$, $p < 0.000$, $\Delta CFI = 0.01$), rejecting full metric invariance. Releasing the equality constraint of the loadings of three items (COG5, BEH2, BEH4) produced partial metric invariance ($\Delta\chi^2 = 8.3$, $p = 0.216$, $\Delta CFI = -0.001$). Third, the test for scalar invariance constrained the intercepts to be equal across samples. The full scalar invariance model was statistically worse than the partial metric invariance model ($\Delta\chi^2 = 63.5$, $p \leq 0.000$, $\Delta CFI = 0.025$); the model did not achieve full scalar invariance. Releasing the equality constraint of the intercepts of two items (COG4 and MC3) led to partial scalar invariance ($\Delta\chi^2 = 10.4$, $p = 0.109$, $\Delta CFI = 0.002$), and the fit indexes of the partial scalar invariance model were satisfactory ($\chi^2/df = 2.07$; RMSEA = 0.042 (90 percent CI: 0.034-0.050); SRMR = 0.047; NNFI = 0.936; CFI = 0.946; Table IV).

To assess if the five non-invariant items exerted significant impacts on the remainder of the analysis, we conducted a differential item functioning (DIF) analysis with CFA (Ferrando, 1996). As recommended by Chan (2000), a comparison of the standardized mean difference (d) across groups considered the values both with and without the five items. The difference between the two-country samples yielded an index of the practical significance of the DIF. The average pairwise d difference was 0.004 for ICK and 0.012 for ECF. These low values suggested that the impact of retaining the DIF items was not substantial. That is, DIF was inconsequential at the scale level of mean scores (Chan, 2000). The two-dimensional model of CQ possessed partial scalar equivalence. This relatively high level of equivalence is necessary to conduct cross-country comparisons and pool the data (Caramellia and Van de Vijver, 2013).

Partial scalar equivalence having been reached, it was possible to assess mean differences in terms of ICK and ECF across the Chinese and Dutch samples. The latent mean invariance assessment showed that model fit indexes declined significantly ($\Delta\chi^2 = 36.1, p < 0.000, \Delta CFI = 0.015$), indicating that the mean level of the two new CQ dimensions significantly varied across samples. The latent means for both dimensions were significantly higher for the Dutch than for the Chinese sample: $ICK_{Dutch} = 0.516$ ($p \leq 0.000$) and $ECF_{Dutch} = 0.358$ ($p \leq 0.000$).

Discussion

This study investigates the cross-cultural invariance of the CQS. A survey of 607 Dutch and Chinese students demonstrates that, contrary to the claim of Ang *et al.* (2007), the original four-dimensional CQS lacks discriminant validity and does not possess cross-cultural measurement invariance. However, the two-dimensional CQS proposed by Bückner *et al.* (2015) possesses both discriminant and partial scalar invariance, as required to allow meaningful comparisons across cultures. These results are consistent with those of recent empirical studies using either an overall CQ measure (Groves and Feyerherm, 2011) or a limited number of dimensions (MacNab and Worthley, 2012), likely because of the lack of discriminant validity of the four CQ dimensions. The present study offers a valid, cross-culturally equivalent, two-dimensional scale to measure CQ. Finding no good fit of the four-dimensional CQS in our study on China and the Netherlands may be due to the fact that these countries are different from Singapore and the USA, included in the Ang *et al.* (2007) study.

In addition, this study contributes to the emic-etic debate in cross-cultural research (Ng and Earley, 2006). The study showed that some items demonstrated equivalence across two cultures represented in the study, but the generalization to cultures (generally) should probably be withheld until more evidence has been accrued. For example, "I know the rules for expressing non-verbal behaviors in other cultures" is not invariant across cultures and therefore did not appear in the final scale. As Hall (1976) argues, some low-context cultures (e.g. the Netherlands) make a clear distinction between verbal and non-verbal communication, whereas high-context cultures (e.g. China) blur this distinction. Therefore, this item might be emic, with different meanings across cultures. Except for linguistic differences (phraseology), it is also possible that other systematic factors are responsible for the in-equivalent findings in this study. For example, the Chinese sample had a significantly higher average age (almost 28 versus 24 years), more experience living abroad, and less interaction with foreigners than the respondents in the Dutch sample (see Table I). Hence, further research that may look into the sources of in-equivalence at multiple levels of the measurement model of CQ is needed.

Ryan *et al.* (1999) claim that as a result of urgency, more often researchers but also practitioners prefer to use etic instruments once they need to roll out a global survey on culturally sensitive topics, such as job satisfaction or work stress. For example, a US-based HR manager may take an attitude survey developed for US employees based on theories of employee attitudes from the USA and uses it worldwide. However, culture-specific measures of constructs of interest using emic items may be preferred as they deliver more detailed culture-specific insights. Despite this preference, the rapid globalization of organizations drives HR managers (but also academics) toward the use of etic scales with universal meaning (Ryan *et al.*, 1999). Academics and practitioners should recognize that any comparison of cultural groups is appropriate only if measurement equivalence exists (Steenkamp and Baumgartner, 1998; Vandenberg and Lance, 2000). Despite warnings about the failure to establish measurement equivalence, recent empirical, CQ literature rarely offers evidence of measurement equivalence. Such equivalence cannot be assumed but must be tested, using different strategies according to the type of study. Comparative studies can assess scalar invariance (Steenkamp and Baumgartner, 1998); single-country studies must ensure that the new instrument measures the same constructs with the same degree of accuracy as in the original cultural setting (Morales and Ladhari, 2011); and studies that pool data from respondents from various cultures must also assess construct and measurement invariance first. If partial scalar invariance exists, pooling the data is acceptable, though pooled data samples are often unbalanced, so the bias of dominant representation is a concern, which warrants tests for measurement equivalence.

Similar to rational intelligence, CQ affects employees' performance (Johnson *et al.*, 2006), and culturally sensitive people show greater commitment to intercultural relationships (Earley and Ang, 2003). The current study thus has notable implications for managers. The two-dimensional CQS offers a high level of measurement equivalence across China and the Netherlands and can apply across these cultures to select employees for international assignments or identify training needs. For MNCs that must recruit large cohorts of new employees simultaneously, the two-dimensional CQ instrument can support comparisons across countries, because the scale possesses measurement invariance.

This study also has limitations. First, the initial goal was to find support for the cross-cultural validity of the original four-dimensional CQS, but the finding of a valid two-dimensional scale, initially developed by Bückner *et al.* (2015), requires further theoretical development and empirical replications should validate this two-dimensional CQS. Second, the refinement of the scale deleted some items that were emic, which raises concerns about the face validity of the remaining items (Thomas, 2006). Additional studies should investigate if the items of the two-dimensional CQS are broad enough to cover CQ; new etic items might expand the scale and improve face validity and reliability. In further development of new items sources of invariance can be minimized by "excluding local language expressions in wording and by using scaling categories and anchors that are comparable in terms of responding patterns across a majority of cultures" (Wernsing, 2014). Third, this study used an existing Chinese CQS (Ang *et al.*, 2007), but intra-country differences in meaning and interpretations could challenge the validity of the Chinese CQS (Roy *et al.*, 2001). Fourth, the data referred to only two cultures. Although China and the Netherlands are different enough to ensure the power of the tests, two-country data might limit the generalizability of the findings. New studies should extend the number of countries, to enable further generalization.

The CQS enjoys great popularity among management scholars. In addition, as business becomes more international, a cross-culturally valid instrument for measuring competencies to cope with cultural diversity is in great demand. This study contributes

to the development of such an instrument by testing the measurement equivalence of the CQS. The assessed two-dimensional CQS possesses measurement equivalence across the two cultures in this study and thus represents a step forward.

Future research could collect more evidence about the emic-etic character of the items by further testing the four-dimensional model and the two-dimensional model in various other less and more distant cultures.

References

- Ang, S. and Van Dyne, L. (2008), *Handbook of Cultural Intelligence: Theory, Measures and Applications*, M.E. Sharpe, Armonk, NY.
- Ang, S., Van Dyne, L. and Koh, C. (2006), "Personality correlates of the four-factor model of cultural intelligence", *Group & Organization Management*, Vol. 31 No. 1, pp. 20-26.
- Ang, S., Van Dyne, L., Koh, C., Ng, K.Y., Templer, K.J., Tay, C. and Chandrasekar, N.A. (2007), "Cultural intelligence: its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance", *Management and Organization Review*, Vol. 3 No. 3, pp. 335-371.
- Barr, S.H. and Hitt, M.A. (1986), "A comparison of selection decision models in manager versus student samples", *Personnel Psychology*, Vol. 39 No. 3, pp. 599-617.
- Bernstein, V., Hakel, M.D. and Harlan, A. (1975), "The college student as interviewer: a threat to generalizability?", *Journal of Applied Psychology*, Vol. 60 No. 2, pp. 266-268.
- Brislin, R.W. (1986), "The wording and translation of research instruments", in Lonner, W.J. and Berry, J.W. (Eds), *Field Methods in Cross-Cultural Research*, Sage, Newbury Park, CA, pp. 137-164.
- Bücker, J. and Poutsma, E. (2010), "Global management competencies: a theoretical foundation", *Journal of Managerial Psychology*, Vol. 25 No. 8, pp. 829-844.
- Bücker, J., Furrer, O. and Lin, Y. (2015), "Measuring cultural intelligence: a new test of the CQ scale", *International Journal of Cross-Cultural Management*, Vol. 15 No. 3, pp. 259-284.
- Bücker, J.J., Furrer, O., Poutsma, E. and Buyens, D. (2014), "The impact of cultural intelligence on communication effectiveness, job satisfaction and anxiety for Chinese host country managers working for foreign multinationals", *The International Journal of Human Resource Management*, Vol. 25 No. 14, pp. 2068-2087.
- Byrne, B.M. (2001), *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*, Lawrence Erlbaum Associates, Mahway, NJ.
- Caramelli, M. and Van de Vijver, F.J. (2013), "Towards a comprehensive procedure for developing measurement scales for cross-cultural management research", *Management international/International Management/Gestión Internacional*, Vol. 7 No. 2, pp. 150-163.
- Chan, D. (2000), "Detection of differential item functioning on the Kirton adaptation-innovation inventory using multiple-group mean and covariance structure analysis", *Multivariate Behavioral Research*, Vol. 35 No. 2, pp. 169-199.
- Chen, A.S.-Y., Lin, Y.-C. and Sawangpattanakul, A. (2011), "The relationship between cultural intelligence and performance with the mediating effect of culture shock: a case of Philippine laborers in Taiwan", *International Journal of Intercultural Relations*, Vol. 35 No. 2, pp. 246-258.
- Cheung, G.W. and Rensvold, R.B. (2002), "Evaluating goodness-of-fit indexes for testing measurement invariance", *Structural Equation Modeling*, Vol. 9 No. 2, pp. 233-255.
- Chong, E. (2008), "Managerial competency appraisal: a cross-cultural study of American and East-Asian managers", *Journal of Business Research*, Vol. 61 No. 3, pp. 191-200.

- Crowne, K.A. (2013), "Cultural exposure, emotional intelligence, and cultural intelligence: an exploratory study", *International Journal of Cross Cultural Management*, Vol. 13 No. 1, pp. 5-22.
- De Beuckelaer, A. (2005), "Measurement invariance issues in international management research", unpublished dissertation, Limburg.
- Earley, P.C. and Ang, S. (2003), *Cultural Intelligence: Individual Interactions Across Cultures*, Stanford University Press, Stanford, CA.
- Ferrando, P.J. (1996), "Calibration of invariant item parameters in a continuous item response model using the extended Lisrel measurement submodel", *Multivariate Behavioral Research*, Vol. 31 No. 4, pp. 419-439.
- Fischer, R. (2004), "Standardization to account for cross-cultural response bias: a classification of score adjustment procedures and review of research in JCCP", *Journal of Cross-Cultural Psychology*, Vol. 35 No. 3, pp. 263-282.
- Fischer, R. (2011), "Cross-cultural training effects on cultural essentialism beliefs and cultural intelligence", *International Journal of Intercultural Relations*, Vol. 35 No. 6, pp. 767-775.
- Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 48 No. 1, pp. 39-50.
- Frisch, N.C. (1990), "An international nursing student exchange program: an educational experience that enhanced student cognitive development", *Journal of Nursing Education*, Vol. 29 No. 1, pp. 10-12.
- Gardner, H. (1983), *Frames of Mind*, Basic Books, New York, NY.
- Greenberg, J. (1987), "The College Sophomore as Guinea Pig: setting the record straight", *The Academy of Management Review*, Vol. 12 No. 1, pp. 157-159.
- Groves, K.S. and Feyerherm, A.E. (2011), "Leader cultural intelligence in context: testing the moderating effects of team cultural diversity on leader and team performance", *Group & Organization Management*, Vol. 36 No. 5, pp. 535-566.
- Gupta, V., Hanges, P.J. and Dorfman, P. (2002), "Cultural clusters: methodology and findings", *Journal of World Business*, Vol. 37 No. 1, pp. 11-15.
- Hall, E.T. (1976), *Beyond Culture*, Doubleday, New York, NY.
- Hofstede, G., Hofstede, G.J. and Minkov, M. (2010), *Cultures and Organizations: Software of the Mind. Intercultural Cooperation and its Importance for Survival*, 3rd ed., McGraw-Hill, New York, NY.
- Hu, L.T. and Bentler, P.M. (1999), "Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives", *Structural Equation Modeling*, Vol. 6 No. 1, pp. 1-55.
- Huang, F. (2006), "Internationalization of curricula in higher education institutions in comparative perspectives: case studies of China, Japan and the Netherlands", *Higher Education*, Vol. 51 No. 4, pp. 521-539.
- Huang, J.L., Curran, P.G., Keeney, J., Poposki, E.M. and DeShon, R.P. (2012), "Detecting and deterring insufficient effort responding to surveys", *Journal of Business Psychology*, Vol. 27 No. 1, pp. 99-114.
- Hult, G.T.M., Ketchen, D.J. Jr, Griffith, D.A., Finnegan, C.A., Gonzalez-Padron, T., Harmancioglu, N., Huang, Y., Berk Talay, M. and Tamer Cavusgil, S. (2008), "Data equivalence in cross-cultural international business research: assessment and guidelines", *Journal of International Business Studies*, Vol. 39 No. 6, pp. 1027-1044.
- Imai, L. and Gelfand, M.J. (2010), "The cultural intelligent negotiator: the impact of cultural intelligence (CQ) on negotiation sequences and outcomes", *Organizational Behavior and Human Decision Processes*, Vol. 112 No. 2, pp. 83-98.

-
- Johnson, J.P., Lenartowicz, T. and Apud, S. (2006), "Cross-cultural competence in international business: toward a definition and a model", *Journal of International Business Studies*, Vol. 37 No. 4, pp. 525-543.
- Johnson, T.P. and Van de Vijver, F.J.R. (2003), "Social desirability in cross-cultural research", in Harkness, J.A., Van de Vijver, F.J.R. and Mohler, P.P. (Eds), *Cross-Cultural Survey Methods*, John Wiley & Sons, Hoboken, NJ, pp. 195-204.
- Lee, L.-Y. and Sukoco, B.M. (2010), "The effect of cultural intelligence on expatriate performance: the moderating effects of international experience", *The International Journal of Human Resource Management*, Vol. 21 No. 7, pp. 963-981.
- McDonald, R.P. and Ho, M.-H.R. (2002), "Principles and practice in reporting structural equation analyses", *Psychological Methods*, Vol. 7 No. 1, pp. 64-82.
- MacNab, B.R. and Worthley, R. (2012), "Individual characteristics as predictors of cultural intelligence development: the relevance of self-efficacy", *International Journal of Intercultural Relations*, Vol. 36 No. 1, pp. 62-71.
- Morales, M. and Ladhari, R. (2011), "Comparative cross-cultural service quality: an assessment of research methodology", *Journal of Service Management*, Vol. 22 No. 2, pp. 241-265.
- Morley, M. and Cerdin, J.-L. (2010), "Intercultural competence in the international business arena", *Journal of Managerial Psychology*, Vol. 25 No. 8, pp. 805-809.
- Ng, K.Y. and Earley, C.P. (2006), "Culture+intelligence: old constructs, new frontiers", *Group & Organization Management*, Vol. 31 No. 1, pp. 4-19.
- Oolders, T., Chernyshenko, O.S. and Stark, S. (2008), "Cultural intelligence as a mediator of relationships between openness to experience and adaptive performance", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence: Theory, Measurement, and Applications*, pp. 145-158.
- Riordan, C.M. and Vandenberg, R.J. (1994), "A central question in cross-cultural research: do employees of different cultures interpret work-related measures in an equivalent manner?", *Journal of Management*, Vol. 20 No. 3, pp. 643-671.
- Robert, C., Lee, W.C. and Chan, K.-Y. (2006), "An empirical analysis of measurement equivalence with the INCOL measure of individualism and collectivism: implications for valid cross-cultural inference", *Personnel Psychology*, Vol. 59 No. 1, pp. 65-99.
- Roy, A., Walters, P.G.P. and Luk, S.T.K. (2001), "Chinese puzzles and paradoxes: conducting business research in China", *Journal of Business Research*, Vol. 52 No. 2, pp. 203-210.
- Ryan, A., Chan, D., Ployhart, R.E. and Slade, L.A. (1999), "Employee attitude surveys in a multinational organization: considering language and culture in assessing measurement equivalence", *Personnel Psychology*, Vol. 52 No. 1, pp. 37-58.
- Smith, P.B. (2004), "Acquiescent response bias as an aspect of cultural communication style", *Journal of Cross-Cultural Psychology*, Vol. 35 No. 1, pp. 50-61.
- Steenkamp, J.-B.E.M. and Baumgartner, H. (1998), "Assessing measurement invariance in cross-national consumer research", *Journal of Consumer Research*, Vol. 25 No. 1, pp. 78-90.
- Strahan, R. and Gerbasi, K.C. (1972), "Short, homogeneous versions of Marlow-Crowne social desirability scale", *Journal of Clinical Psychology*, Vol. 28 No. 2, pp. 191-193.
- Thomas, D.C. (2006), "Domain and development of cultural intelligence: the importance of mindfulness", *Group & Organization Management*, Vol. 31 No. 1, pp. 78-99.
- Thomas, D.C., Elron, E., Stahl, G., Ekelund, B.Z., Ravelin, E.C., Cerdin, J.-L., Poelmans, S., Brislin, R., Pekerti, R., Aycan, Z., Maznevski, M., Au, K. and Lazarova, M.B. (2008), "Cultural intelligence domain and assessment", *International Journal of Cross Cultural Management*, Vol. 8 No. 2, pp. 123-143.

- Thomas, D.C., Liao, Y., Aycan, Z., Cerdin, J.L., Pekerti, A.A., Ravlin, E.C. and Moeller, M. (2015), "Cultural intelligence: a theory-based, short form measure", *Journal of International Business Studies*, Vol. 46 No. 9, pp. 1099-1118.
- Triandis, H.C. (2006), "Cultural intelligence in organizations", *Group & Organization Management*, Vol. 31 No. 1, pp. 20-26.
- Triandis, H.C. and Marin, G. (1983), "Etic plus emic versus pseudoetic: a test of basic assumption of contemporary cross-cultural psychology", *Journal of Cross Cultural Psychology*, Vol. 14 No. 4, pp. 489-500.
- Van de Vijver, F. and Leung, K. (2000), "Methodological issues in psychological research on culture", *Journal of Cross-Cultural Psychology*, Vol. 31 No. 1, pp. 33-51.
- Van de Vijver, F. and Tanzer, N.K. (2004), "Bias and equivalence in cross-cultural assessment: an overview", *Revue Européenne de Psychologie Appliquée*, Vol. 54 No. 2, pp. 119-135.
- Van de Vijver, F.J.R. and Hambleton, R.K. (1996), "Translating tests: some practical guidelines", *European Psychologist*, Vol. 1 No. 2, pp. 89-99.
- Van Hoof, H.B. and Verbeeten, M.J. (2005), "Wine is for drinking, water is for washing: student opinions about international exchange programs", *Journal of Studies in International Education*, Vol. 9 No. 1, pp. 42-61.
- Vandenberg, R.J. and Lance, C.E. (2000), "A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research", *Organizational Research Methods*, Vol. 3 No. 1, pp. 4-70.
- Vedadi, A., Kheiri, B. and Abbasalizadeh, M. (2010), "The relationship between cultural intelligence and achievement: a case study in an Iranian company", *Iranian Journal of Management Studies*, Vol. 3 No. 3, pp. 25-40.
- Voorhees, C.M., Brady, M.K., Calantone, R. and Ramirez, E. (2016), "Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies", *Journal of the Academy of Marketing Science*, Vol. 44 No. 1, pp. 119-134.
- Ward, C. and Fischer, R. (2008), "Personality, cultural intelligence and cross-cultural adaptation", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence: Theory, Measurements, and Applications*, Routledge, Oxon, pp. 159-173.
- Ward, C., Wilson, J. and Fischer, R. (2011), "Assessing the predictive validity of cultural intelligence over time", *Personality and Individual Differences*, Vol. 51 No. 2, pp. 138-142.
- Ward, C., Fischer, R., Lam, F.S.Z. and Hall, L. (2009), "The convergent, discriminant and incremental validity of the scores on a self-report measure of cultural intelligence", *Educational and Psychological Measurement*, Vol. 69 No. 1, pp. 85-105.
- Watkins, D. (1996), "The influence of social desirability on learning process questions: a neglected possibility?", *Contemporary Educational Psychology*, Vol. 21 No. 1, pp. 80-82.
- Wernsing, T. (2014), "Psychological capital a test of measurement invariance across 12 national cultures", *Journal of Leadership & Organizational Studies*, Vol. 21 No. 2, pp. 179-190.
- Yang, S. and Sternberg, R.J. (1997), "Taiwanese Chinese people's conceptions of intelligence", *Intelligence*, Vol. 25 No. 1, pp. 21-36.

Further reading

- Allport, G.W. (1954), *The Nature of Prejudice*, Addison-Wesley, Reading, MA.
- Ramsey, J.R., Nassif Leonel, J., Zoccal Gomes, G. and Rafael Reis Monteiro, P. (2011), "Cultural intelligence's influence on international business travelers' stress", *Cross Cultural Management: An International Journal*, Vol. 18 No. 1, pp. 21-37.

Authors	Sample	Reliability and validity test	Remarks
<i>Type 1: comparative studies</i>			
Ang <i>et al.</i> (2007), Singapore and USA scale development;	Singapore and USA	Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural equivalence: configural, metric, covariance	No test for scalar invariance; means cannot be compared across countries
Ang and Van Dyne (2008)			
Ang <i>et al.</i> (2007), study 1	Singapore and USA	Study 1 Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported	Configural invariance assumed; the 20 CQS items are used for the same dimensions, but no cross-cultural equivalence test, so results cannot be compared across cultural groups
Imai and Gelfand, (2010), study 1	Caucasian/white (58%), Asian-American/Pacific Islander (31%), others (11%)	Reliability: yes Convergent validity: yes Discriminant validity: not reported Cross-cultural equivalence: not reported	Four-dimension structure of the CQS established with pooled data. No cross-cultural equivalence test, so results cannot be compared across cultural groups
Imai and Gelfand (2010), study 2	USA and East-Asia (China, Japan, and Korea)	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural equivalence: not reported	No cross-cultural equivalence test, so results (including means) cannot be compared across cultural groups
<i>Type 2: single-country/culture studies</i>			
Oolders <i>et al.</i> (2008) ^a	New Zealand	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: no	No cross-cultural equivalence test, so results cannot be compared to results from other countries
Amiri <i>et al.</i> (2010)	Iran	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: no	Configural equivalence is assumed but not reported for the 20 CQS items. No cross-cultural equivalence test, so results cannot be compared to results from other countries
Vedadi <i>et al.</i> (2010)	Iran	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: no	Configural equivalence is assumed but not reported for the 20 CQS items. No cross-cultural equivalence test, so results cannot be compared to results from other countries

(continued)

Table AI.
Overview of
empirical studies
using CQS

Authors	Sample	Reliability and validity test	Remarks
Lee and Sukoco (2010)	Taiwan	Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported	No measure of metacognitive CQ and limited number of items (9), so configural equivalence is not achieved. No cross-cultural equivalence test, so results cannot be compared to other countries
Moon (2010)	Korea	Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported	No cross-cultural equivalence test, so results cannot be compared to results from other countries
Chen <i>et al.</i> (2011)	Philippines	Reliability: yes Convergent validity: yes Discriminant validity: not reported Cross-cultural validity: no	Configural equivalence is assumed but not reported for the 20 CQS items. No cross-cultural equivalence test, so results cannot be compared to results from other countries
Eisenberg <i>et al.</i> (2013), study 1	Australian students	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: no	Configural equivalence is assumed but not reported for the 20 CQS items. No cross-cultural equivalence test, so results cannot be compared to results from other countries
<i>Type 3: mixed (or pooled) sample studies</i>			
Adair <i>et al.</i> (2013)	Culturally homogeneous teams and culturally heterogeneous teams	Reliability: yes Convergent validity: no Discriminant validity: no Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Ang <i>et al.</i> (2006)	Multietnic and multicultural group of students in Singapore	Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported	Configural equivalence is achieved for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Templer <i>et al.</i> (2006)	India, South-East Asia, other Asian countries, Europe, America and Canada, Australia and New Zealand, other nations	Reliability: yes (MOT only) Convergent validity: not reported Discriminant validity: no Cross-cultural validity: not reported	Configural equivalence is achieved for the five MOT items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Ang <i>et al.</i> (2007), study 2	International managers from 17 different countries	Reliability: yes Convergent validity: yes	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization

(continued)

Authors	Sample	Reliability and validity test	Remarks
Ang <i>et al.</i> (2007), study 3	International professionals from 12 different countries	Discriminant validity: yes Cross-cultural validity: not reported Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural equivalence: configural, metric, covariance	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization No test for scalar invariance; the sample cannot be pooled without standardization
Ang and Van Dyne (2008), study 6 ^a	Singapore and US pooled	Reliability: yes Convergent validity: yes Discriminant validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Kim <i>et al.</i> (2008) ^a	White American, Latino American, African American, and others	Reliability: yes Convergent validity: yes Discriminant validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Rockstuhl and Ng (2008)	Respondents from 19 countries	Reliability: yes Convergent validity: not reported Discriminant validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Shannon and Begley (2008) ^a	Respondents from 24 different countries	Reliability: yes Convergent validity: yes Discriminant validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Shokef and Erez (2008) ^a	Israel, Hong Kong, Spain, South Korea, and USA	Reliability: yes Convergent validity: yes Discriminant validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Tarique and Takeuchi (2008) ^a	Mixed group of respondents	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization

(continued)

Table AI.

Authors	Sample	Reliability and validity test	Remarks
Tay <i>et al.</i> (2008) ^a	Singapore, Israel, and Brazil	Reliability: limited due to small number of items Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported Reliability: yes Convergent validity: yes Divergent validity: not relevant, one factor only Cross-cultural validity: not reported	2 items for COG, 1 item for MC, 3 items for MOT, and 2 items for BEH; configural invariance is not achieved. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization Only motivational CQ was used. Cross-cultural equivalence is assumed but not tested
Ward and Fischer (2008) ^a	Respondents from 30 different countries	Reliability: yes Convergent validity: yes Divergent validity: not relevant, one factor only Cross-cultural validity: not reported	Configural equivalence is assumed but not reported for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Elenkov and Manev (2009)	Respondents 27 European countries	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported	Configural equivalence is assumed but not reported for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Ward <i>et al.</i> (2009), study 1	Respondents from 130 countries	Reliability: yes Convergent validity: yes Discriminant validity: not reported Cross-cultural validity: not reported Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported	Configural equivalence is assumed but not reported for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Ward <i>et al.</i> (2009), studies 2 and 3	Respondents from Asia, Europe, North-America, Pacific, and other countries	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported	Discriminant validity might not be achieved (high intercorrelations between CQ dimensions). Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Ramula <i>et al.</i> (2010)	Respondents from India, UK, Australia, and 42 other countries	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported	Configural equivalence is assumed but not reported for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Rose <i>et al.</i> (2010)	Respondents from India, UK, Australia, and 42 other countries	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported	Configural equivalence is assumed but not for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization

(continued)

(continued)

Table AI.

Authors	Sample	Reliability and validity test	Remarks
Groves and Feyerherm (2011)	Respondents from the USA, China, Mexico, Philippines, Salvador, Saudi Arabia, Vietnam, Armenia	Reliability: yes Convergent validity: yes Divergent validity: yes Cross-cultural validity: not reported	Configural equivalence is assumed but not reported for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
MacNab and Worthley (2012)	Respondents from 30 different countries	Reliability: yes Convergent validity: yes Discriminant validity: not reported Cross-cultural validity: not reported	14 items of three CQ factors (MET, MOT, and BEH) show configural equivalence for these dimensions only. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Ward <i>et al.</i> (2011)	Respondents from 25 different countries	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported	Configural equivalence seems to be achieved for the 20 CQS items. Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Wu and Ang (2011)	Respondents from European countries, Asian countries, North-America, and Australia	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Rockstuhl <i>et al.</i> (2011)	Swiss respondents from multiple linguistic regions	Reliability: yes Convergent validity: yes Discriminant validity: not reported Cross-cultural validity: no	Four factors of CQ were averaged to form one overall CQ, but configural equivalence cannot be assumed
Eisenberg <i>et al.</i> (2013), study 2	Group of mixed students from Europe and other countries	Reliability: yes Convergent validity: no Discriminant validity: no Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Erez <i>et al.</i> (2013)	Respondents from 12 countries	Reliability: yes Convergent validity: yes Discriminant validity: not reported Cross-cultural validity: no	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Huff (2013)	Expatriates in Japan from English-speaking countries	Reliability: yes Convergent validity: no Discriminant validity: no Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization

Table AI.

Authors	Sample	Reliability and validity test	Remarks
Malek and Budhwar (2013)	Asian and non-Asian expatriates in Malaysia	Reliability: yes Convergent validity: yes Discriminant validity: yes Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Rosenblatt <i>et al.</i> (2013)	Respondents from 32 countries	Reliability: yes Convergent validity: not reported Discriminant validity: not reported Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization
Bücker and Korzilius (2015)	French and Dutch students	Reliability: yes Convergent validity: no Discriminant validity: no Cross-cultural validity: not reported	Cross-cultural equivalence is assumed but not tested, and the sample cannot be pooled without standardization

Note: ^aStudies published in the *Handbook of Cultural Intelligence* (Ang and Van Dyne, 2008)

References

- Adair, W.L., Hideg, I. and Spence, J.R. (2013), "The culturally intelligent team the impact of team cultural intelligence and cultural heterogeneity on team shared values", *Journal of Cross-Cultural Psychology*, Vol. 44 No. 6, pp. 941-962.
- Amiri, A.N., Moghimi, S.M. and Kazemi, M. (2010), "Studying the relationship between cultural intelligence and employees' performance", *European Journal of Scientific Research*, Vol. 42, pp. 432-441.
- Ang, S. and Van Dyne, L. (2008), *Handbook of Cultural Intelligence: Theory, Measures and Applications*, M.E. Sharpe, Armonk, NY.
- Ang, S., Van Dyne, L. and Koh, C. (2006), "Personality correlates of the four-factor model of cultural intelligence", *Group & Organization Management*, Vol. 31, pp. 20-26.
- Ang, S., Van Dyne, L., Koh, C., Ng, K.Y., Templer, K.J., Tay, C. and Chandrasekar, N.A. (2007), "Cultural intelligence: its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance", *Management and Organization Review*, Vol. 3, pp. 335-371.
- Bücker, J.J. and Korzilius, H. (2015), "Developing cultural intelligence: assessing the effect of the ecotonos cultural simulation game for international business students", *The International Journal of Human Resource Management*, Vol. 26 No. 15, pp. 1995-2014.
- Chen, A.S.-Y., Lin, Y.-C. and Sawangpattanakul, A. (2011), "The relationship between cultural intelligence and performance with the mediating effect of culture shock: a case of Philippine laborers in Taiwan", *International Journal of Intercultural Relations*, Vol. 35, pp. 246-258.
- Eisenberg, J., Lee, H.J., Brück, F., Brenner, B., Claes, M.T., Mironski, J. and Bell, R. (2013), "Can business schools make students culturally competent? Effects of cross-cultural management courses on cultural intelligence", *Academy of Management Learning & Education*, Vol. 12 No. 4, pp. 603-621.
- Elenkov, D.S. and Manev, I.M. (2009), "Senior expatriate leadership's effects on innovation and the role of cultural intelligence", *Journal of World Business*, Vol. 44, pp. 357-369.
- Erez, M., Lisak, A., Harush, R., Glikson, E., Nouri, R. and Shokef, E. (2013), "Going global: developing management's students' cultural intelligence and global identity in culturally diverse virtual teams", *Academy of Management Learning & Education*, Vol. 12 No. 3, pp. 330-355.
- Groves, K.S. and Feyerherm, A.E. (2011), "Leader cultural intelligence in context: testing the moderating effects of team cultural diversity on leader and team performance", *Group & Organization Management*, Vol. 36 No. 5, pp. 535-566.
- Huff, K.C. (2013), "Language, cultural intelligence and expatriate success", *Management Research Review*, Vol 36 No. 6, pp. 596-612.
- Imai, L. and Gelfand, M.J. (2010), "The cultural intelligent negotiator: the impact of cultural intelligence (CQ) on negotiation sequences and outcomes", *Organizational Behavior and Human Decision Processes*, Vol. 112, pp. 83-98.
- Kim, K., Kirkman, B.L. and Chen, G. (2008), "Cultural intelligence and international assignment effectiveness", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 71-90.
- Lee, L.-Y. and Sukoco, B.M. (2010), "The effect of cultural intelligence on expatriate performance: the moderating effects of international experience", *The International Journal of Human Resource Management*, Vol. 21 No. 7, pp. 963-981.

MacNab, B.R. and Worthley, R. (2012), "Individual characteristics as predictors of cultural intelligence development: the relevance of self-efficacy", *International Journal of Intercultural Relations*, Vol. 36 No. 1, pp. 62-71.

Malek, M.A. and Budhwar, P. (2013), "Cultural intelligence as a predictor of expatriate adjustment and performance in Malaysia", *Journal of World Business*, Vol. 48 No. 2, pp. 222-231.

Moon, T. (2010), "Emotional intelligence correlates of the four-factor model of cultural intelligence", *Journal of Managerial Psychology*, Vol. 25 No. 8, pp. 876-898.

Oolders, T., Chernyshenko, O.S. and Stark, S. (2008), "Cultural intelligence as a mediator of relationships between openness to experience and adaptive performance", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 145-158.

Ramula, S.S., Rose, R.C., Kumar, N. and Uli, J. (2010), "Doing business in global arena: an examination of the relationship between cultural intelligence and cross-cultural adjustment", *Asian Academy of Management Journal*, Vol. 15, pp. 79-97.

Rockstuhl, T. and Ng, K.-Y. (2008), "The effects of cultural intelligence on interpersonal trust in multicultural teams", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 206-220.

Rockstuhl, T., Seiler, S., Ang, S., Van Dyne, L. and Annen, A. (2011), "Beyond cultural intelligence (IQ) and emotional intelligence (EQ): the role of cultural intelligence (CQ) on cross-border leadership effectiveness in a globalized world", *Journal of Social Issues*, Vol. 67 No. 4, pp. 825-840.

Rose, R.C., Ramalu, S.S., Uli, J. and Kumar, N. (2010), "Expatriate performance in international assignments: the role of cultural intelligence as dynamic intercultural competency", *International Journal of Business and Management*, Vol. 5, pp. 76-85.

Rosenblatt, V., Worthley, R. and MacNab, B. (2013), "From contact to development in experiential cultural intelligence education: the mediating influence of expectancy disconfirmation", *Academy of Management Learning & Education*, Vol. 12 No. 3, pp. 356-379.

Shannon, L.M. and Begley, T.M. (2008), "Antecedents of the four-factor model of cultural intelligence", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 41-55.

Shokef, E. and Erez, M. (2008), "Cultural intelligence and global identity in multicultural teams", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 171-191.

Tarique, I. and Takeuchi, R. (2008), "Developing cultural intelligence: the role of international nonwork experiences", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 56-70.

Tay, C., Westman, M. and Chia, A. (2008), in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 126-144.

Templer, K.J., Tay, C. and Chandrasekar, N.A. (2006), "Motivational cultural intelligence, realistic job preview, realistic living conditions preview, and cross-cultural adjustment", *Group & Organization Management*, Vol. 31, pp. 154-173.

Vedadi, A., Kheiri, B. and Abbasalizadeh, M. (2010), "The relationship between cultural intelligence and achievement: a case study in an Iranian company", *Iranian Journal of Management Studies*, Vol. 3, pp. 25-40.

Ward, C. and Fischer, R. (2008), "Personality, cultural intelligence, and cross-cultural adaptation", in Ang, S. and Van Dyne, L. (Eds), *Handbook of Cultural Intelligence*, M.E. Sharpe, Armonk, NY, pp. 159-173.

Ward, C., Wilson, J. and Fischer, R. (2011), "Assessing the predictive validity of cultural intelligence over time", *Personality and Individual Differences*, Vol. 51, pp. 138-142.

Ward, C., Fischer, R., Lam, F.S.Z. and Hall, L. (2009), "The convergent, discriminant, and incremental validity of scores on a self-report measure of cultural intelligence", *Educational and Psychological Measurement*, Vol. 69, pp. 85-105.

Wu, P.-C. and Ang, S.H. (2011), "The impact of expatriate supporting practices and cultural intelligence on cross-cultural adjustment and performance of expatriates in Singapore", *International Journal of Human Resource Management*, Vol. 22 No. 13, pp. 2683-2702.

Corresponding author

Joost Bucker can be contacted at: J.Bucker@fm.ru.nl