



Contents lists available at ScienceDirect

International Journal of Intercultural Relations

journal homepage: www.elsevier.com/locate/ijintrel



Additional cross-cultural validity testing of the Intercultural Development Inventory[☆]

Mitchell R. Hammer^{a,b,*}

^a American University, United States

^b IDI, LLC and Hammer Consulting, LLC, P.O. Box 1388, Ocean Pines, MD 21811, United States

ARTICLE INFO

Article history:

Article now published: Hammer, M.R. (2011). Additional cross-cultural validity testing of the Intercultural Development Inventory. *International Journal of Intercultural Relations*. V35, 474–487.

ABSTRACT

Intercultural competence/sensitivity is increasingly recognized across the global spectrum of educational institutions, corporations, government agencies and non-government organizations as a central capability for the 21st century. The Intercultural Development Inventory (IDI) is an assessment tool that measures the level of intercultural competence/sensitivity across a developmental continuum for individuals, groups, and organizations and represents a theoretically grounded measure of this capability for perceiving cultural differences and commonalities and modifying behavior to cultural context. This study represents additional cross-cultural validity testing of the IDI, building on the previous work of Paige, Jacobs-Cassuto, Yerushova, and DeJaeghere (2003) and Hammer, Bennett, and Wiseman (2003). The 50 items from IDI v2 were administered to 4763 individuals from 11 distinct, cross-cultural samples. Confirmatory factor analysis confirms the following basic orientations toward cultural difference originally explicated by Bennett (1986, 1993) in the Developmental Model of Intercultural Sensitivity (DMIS): Denial, Defense, Reversal, Minimization, Acceptance, and Adaptation. In addition, it also identifies Cultural Disengagement as an additional scale within the IDI; but one that is not located along the developmental continuum. Second, the inter-scale correlations support the theoretically proposed, developmental formulation from Denial through Adaptation. Third, the current analysis offers strong support for an overall Developmental Orientation (DO) scale and an overall Perceived Orientation (PO) scale. Fourth, Minimization is found to be a transitional orientation toward cultural differences and commonalities, between the more monocultural (ethnocentric) orientations of Denial and Polarization (Defense, Reversal) and the more intercultural mindsets of Acceptance and Adaptation. Fifth, readability analysis indicates the IDI is appropriate for high (secondary) school students (age 15 or above) or individuals with a 10th grade reading level. Finally, criterion validity of the IDI was assessed. The results indicate that the IDI has strong predictive validity toward bottom-line goals within organizations; namely, the achievement of diversity and inclusion goals in the recruitment and staffing function. These findings complement previous results that demonstrated that the IDI also possesses strong content and construct validity across culture groups.

© 2011 Elsevier Ltd. All rights reserved.

[☆] The author would like to express his appreciation of Richard Wiseman, Ph.D. (deceased) for his long-term support of the IDI and life-long commitment to improving intercultural relations and for his valuable assistance in the analysis of the data and insights regarding the use of statistical analysis and interpretation of the empirical findings in this study. In addition, thanks to the 2400 (and growing) IDI Qualified Administrators in 42 countries who are using the IDI in their work in education, health care, government and corporations to improve the intercultural competence of individuals and institutions.

* Correspondence address: IDI, LLC and Hammer Consulting, LLC, P.O. Box 1388, Ocean Pines, MD 21811, United States.
E-mail address: mhammer@hammerconsulting.org

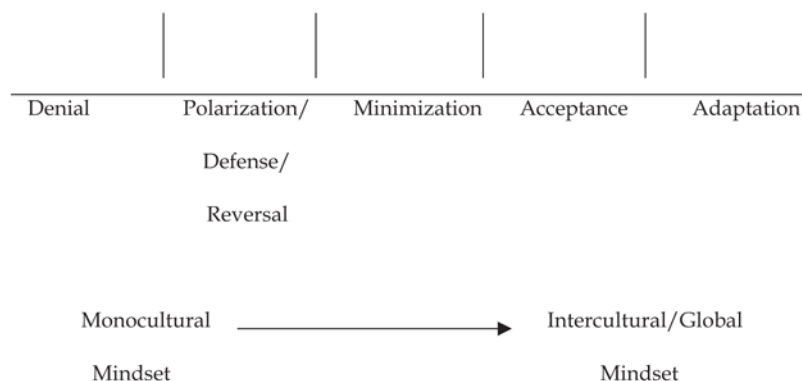


Fig. 1. Intercultural development continuum.

1. Introduction

The educational sector—from K-12 through universities as well corporations, government agencies and non-government organizations are increasingly recognizing the need for building intercultural competence in order to better prepare individuals to function more effectively in our global community (Hammer, 1989, 1999, 2009; Moodian, 2009). Bhawuk and Brislin (1992) posit that “people must be interested in other cultures, be sensitive enough to notice cultural differences, and then also be willing to modify their behavior as an indication of respect for the people of other cultures” in order to effectively bridge across cultural differences and commonalities (p. 416). The Intercultural Development Inventory (IDI) is an assessment tool that measures the level of intercultural competence/sensitivity across a developmental continuum for individuals, groups, and organizations and represents a theoretically grounded measure of this capability toward observing cultural differences and commonalities and modifying behavior to cultural context. The IDI is grounded in the Developmental Model of Intercultural Sensitivity (DMIS) originally proposed by Bennett (1986, 1993, 2004). The purpose of this study is to further test the cross-cultural validity of the IDI.

The IDI v1 was developed in 1998; IDI v2, based on continued research, was introduced in 2003; and IDI v3 is available for use in 2010 as a result of the research reported in this article. In its current version, the IDI v3 is a 50-item paper-and-pencil (and online) questionnaire, with selected demographic items. It has been back translated into multiple languages (Brislin, 1970, 1976, 1980) with selected demographics. The IDI can be completed in about 15–20 min. Accompanying the IDI are five open-ended “contexting” questions individual respondents may complete. These open-ended questions help further capture the experiences around cultural differences of the respondent. Once the IDI is completed, the IDI analytic structure generates an individual (or group) graphic profile of the respondent’s overall position on the intercultural development continuum (also produced is an Administrator’s IDI profile report that presents the statistical information for the various IDI scales). The intercultural competence/sensitivity developmental continuum is presented in Fig. 1, and ranges from more monocultural orientations to more intercultural/global orientations.

Table 1 presents a summary of each of these orientations toward cultural commonalities and differences (see Bennett, 2004; Hammer, 2009 for more information on these intercultural competence/sensitivity orientations).

This continuum represents a movement toward greater intercultural competence/sensitivity, from a less complex set of perceptions and behaviors around cultural commonalities and differences (monocultural mindset orientations) to a more complex set of perceptions and behaviors (intercultural/global mindset) (Bennett, 2004; Hammer, 2009). Perceiving cultural

Table 1

Summary of IDI developmental continuum orientations.

Denial	An orientation that likely recognizes more observable cultural differences (e.g., food) but, may not notice deeper cultural differences (e.g., conflict resolution styles), and may avoid or withdraw from cultural differences.
Polarization	A judgmental orientation that views cultural differences in terms of “us” and “them”. This can take the form of:
Defense	An uncritical view toward one’s own cultural values and practices and an overly critical view toward other cultural values and practices.
Reversal	An overly critical orientation toward one’s own cultural values and practices and an uncritical view toward other cultural values and practices.
Minimization	An orientation that highlights cultural commonality and universal values and principles that may also mask deeper recognition and appreciation of cultural differences.
Acceptance	An orientation that recognizes and appreciates patterns of cultural difference and commonality in one’s own and other cultures.
Adaptation	An orientation that is capable of shifting cultural perspective and changing behavior in culturally appropriate and authentic ways.
Cultural Disengagement	A sense of disconnection or detachment from a primary cultural group.

differences from one's own cultural perspective is indicative of a more monocultural mindset. In contrast, the capability of shifting cultural perspective and adapting behavior to cultural context represents an intercultural mindset.

As previously mentioned, this continuum is grounded in Bennett's (1986, 1993, 2004) Developmental Model of Intercultural Sensitivity. Based on the research conducted in developing and validating the IDI, the current intercultural competence/sensitivity continuum assessed by the IDI represents an adaptation from the original formulation of the DMIS in three main areas.

First, in the DMIS model, a final stage of intercultural development is proposed and is labeled as Integration, a stage that is concerned with the construction of an intercultural identity. However, the earlier stages posited in the DMIS are directly related to the developing of intercultural competence/sensitivity and only indirectly related to the formation of cultural identity (Bennett, 2004; Hammer, 2009). Because the IDI is a measure of the developmental continuum of intercultural competence/sensitivity and not a measure of identity development, it is appropriate to conceive of the developmental continuum as moving from Denial through Adaptation.

Second, the IDI assesses Cultural Disengagement—the degree to which an individual or group is experiencing a sense of alienation from their own cultural community. This is a separate dimension based on the current research presented below and is conceptually located (and empirically verified) outside of the developmental continuum. This dimension was termed encapsulated marginality in IDI v2 based on DMIS conceptualization. However, the data presented in this study indicates that this scale is better conceptualized as a measure of Cultural Disengagement rather than as an indicator of encapsulated marginality (a form of Integration as hypothesized in the DMIS).

Finally, Minimization in the original DMIS is viewed as an ethnocentric stage of development. However, the data presented below from the current research support the conclusion that Minimization is *not* ethnocentric in orientation toward cultural commonalities and differences. Rather, Minimization is conceived as a transitional orientation that is more effective around recognizing and responding to cultural commonalities but is challenged when complex cultural differences need to be adapted to through deeper understanding of the values and behavior patterns of the other cultural community.

2. Previous validation of the IDI

2.1. Phase 1 testing: preparation of IDI v1

Phase 1 testing began the development of an initial, 60-item IDI (IDI v1). The first step was to assess whether the observations of cultural commonalities and differences developed by M. Bennett incorporated in the DMIS theory could be replicated in more systematic ways. The solution was to examine how people construe their experiences around cultural commonalities and differences through qualitative interviewing of a culturally diverse sample of 40 individuals.

Following an interview guide, four trained interviewers conducted in-depth interviews with this sample. The interviews lasted between one to one and one half hours and were audio recorded. Verbatim transcripts of each interview were prepared and four members of the research team then independently reviewed the transcripts and rated the transcripts in terms of how consistent (or not) Denial, Defense/Reversal, Minimization, Acceptance, Adaptation and Integration orientations were expressed by each interviewee. Inter-rater reliabilities (Cohen's kappa) were calculated and resulted in high reliabilities, ranging from .66 to .86 (Hammer, Bennett, & Wiseman, 2003). These results indicated that people construe their cross-cultural experiences in ways described by the DMIS theory.

Because the "data set" was verbatim statements made by the interviewees, the actual statements made by this culturally diverse group were identified for each of the developmental orientations. Thus, unlike other instrument development approaches, the actual items of the IDI were originally generated in natural discourse by people from a wide range of cultures. This is in contrast to questionnaire items that are generated by the researchers themselves. This original innovative approach to generating the item pool provides evidence for the content validity of the items and counters possible criticism of systematic cultural "bias" being introduced by the researchers in the wording of the items themselves.

The items were then reviewed by culturally diverse pilot testing groups that reviewed the items for clarity. Following this, the pool of items was then submitted for expert panel review to seven acknowledged cross-cultural experts. These individuals independently classified each item statement in terms of whether the item reflected a specific DMIS orientation or were too difficult to categorize. Inter-rater reliabilities across the seven rater/experts were completed across the items, resulting in further winnowing of the item pool. This panel review protocol also demonstrated strong content validity support of the IDI.

The remaining items were then administered to a sample of 226 subjects, 70% from the United States and 30% from 28 different countries. Factor analysis was conducted on the items. This produced the following factors: Denial, Defense, Minimization, Acceptance, Cognitive Adaptation and Behavioral Adaptation. Reversal and Integration orientations identified in the DMIS model did not emerge as stable factors.

Subsequent research on the 60-item IDI by Paige, Jacobs-Cassuto, Yershova, and DeJaeghere (2003) factor analyzed the 60 items and identified six factors: Denial/Defense, Physical Universalism (a form of Minimization), Transcendent Universalism (a form of Minimization) Behavioral Adaptation, Cognitive Adaptation and a sixth factor not consistent with any identified DMIS orientation. While generally consistent with the original factor structure, some differences did arise. Further, their results suggested that Minimization (Physical Universalism, Transcendent Universalism) empirically clustered between the more ethnocentric (monocultural) orientations of Denial/Defense and the more ethnorelative (intercultural) orienta-

tion of Adaptation in ways that did not demonstrate the expected closer relationship to Denial/Defense (an ethnocentric orientation).

2.2. Phase 2 testing: preparation of IDI v2

The original set of items were again reviewed, and a subsequent set of 122 items were identified for further testing, with attention given to items that better reflect Reversal and Integration orientations (which were not identified in the Phase 1, IDI v1 analysis). Along with these IDI items, a worldmindedness scale, an intercultural anxiety measure, and a social desirability scale were administered to a new sample of 591 culturally diverse respondents. In this study, confirmatory factor analysis was employed. The results indicated that a two dimensional model (ethnocentrism, a monocultural mindset and ethnorelativism, an intercultural mindset) was not a good fit to the data. In contrast, both a seven dimension model (Denial, Defense/Reversal, Minimization, Acceptance, Adaptation and Integration (Encapsulated Marginality only)) and a five dimensional model (DD (Denial/Defense), R (Reversal), M (Minimization), AA (Acceptance/Adaptation) and EM (Encapsulated Marginality only)) were a satisfactory fit to the data. Based on the criteria of parsimony, the five-factor model was selected. Subsequent item analysis produced IDI v2, a 50-item measure. Coefficient alpha for each of these scales was .80 or higher (Hammer et al., 2003).

As mentioned earlier, content validity of the items in the IDI was addressed in the development of the initial, 60-item IDI v1. Construct validity testing of IDI v2 (50 items) was undertaken by comparing responses of the DD, R, M, AA and EM scales to two theoretically related constructs: worldmindedness and intercultural anxiety. The correlations were as theoretically predicted (e.g., higher scores on DD (Denial/Defense) were significantly related to lower worldmindedness and higher intercultural anxiety; higher AA (Acceptance/Adaptation) scores were significantly related to higher worldmindedness and lower intercultural anxiety). Further, there was no significant relationship of any of the IDI scales to social desirability. Additional testing also revealed no systematic, significant differences on the IDI scales by gender, age, or education level (Hammer et al., 2003).

2.2.1. Post Phase 2: additional testing of IDI v2

Following the publication of the Phase 1 and Phase 2 testing of the IDI by Hammer et al. (2003), further analyses of IDI v2 were conducted in order to provide additional insights based on the IDI and presented below. For this post Phase 2 research, another sample of respondents ($n = 766$) completed the IDI. The goal was to develop a "total IDI score" that would represent an individual's primary orientation along the developmental continuum that ranges from Denial/Defense through Acceptance Adaptation (the EM scale was not included in this effort because the EM scale was viewed as an incomplete measure of the larger Integration stage proposed in the DMIS model at the time this was undertaken).

Two distinct "total IDI scores" were computed, informed by the earlier work of Paige et al. (2003), and based on responses to the IDI from a separate sample of 766 respondents. A Developmental Score was calculated using a weighted formula that is theoretically consistent with the developmental continuum specified in the DMIS. This resulted in a standardized (z-score) with a mean of 100 and a standard deviation of 15. Using the same approach, a Perceived Score was also calculated, this time using an unweighted formula of the scales. This produced an overall score along the developmental continuum that reflects where an individual perceives his/her placement along the developmental continuum. The Developmental Score in contrast, identifies the individual's primary orientation by taking into consideration the degree to which earlier scale orientation scores are resolved compared to AA scale scores. The Developmental Score achieved reliability of .83 and the Perceived Score reliability was .82. Therefore, the result of Phase 2 testing was an expanded analytical capability with the Developmental Score and Perceived Score, each obtaining good reliability.

In order to continue validity testing of the IDI, the current (Phase 3) research was completed. Below is a summary of the various studies and analyses conducted in this Phase 3 effort, which resulted in IDI v3.

3. Current Phase 3 testing: development of IDI v3

The current 50 items from IDI v2 were administered to 4763 individuals from 11 distinct, cross-cultural samples. These are:

1. 213 managers from a wide range of countries who work in a non-government organization (NGO).
2. 150 members of a local church within the United States.
3. 2693 students from a major U.S. university.
 - 1850 high school students (10th–12th grade, US educational system) from the following 8 countries:
4. 114 from Austria.
5. 181 from Brazil.
6. 149 from Costa Rica and Ecuador.
7. 564 from Germany.
8. 94 from Hong Kong.
9. 295 from Italy.

10. 277 from Japan.
11. 175 from the United States.

This sample represents a greater number of participants compared to the IDI v1 validation sample of 226 and the IDI v2 sample of 591. In addition, the 4763 respondents in the current study are culturally diverse in composition. Further, all participants completed the IDI in their native language using rigorously back-translated versions (Brislin, 1970, 1976, 1980) of the IDI unless English was the language of the organization (e.g., managers from the non-government organization took the IDI in English due to exceptionally high English language fluency).

3.1. Sample characteristics

Demographic characteristics for the participants from the non-government organization (NGO) are as follows: (1) 46% were men, 54% were women; (2) the cultural background of the participants were: U.S. American (28%), Chinese (10%), Indian (India, 6%) and Canadian (4%); (3) 13% had a doctoral degree, 56% had a Master's degree, and 23% a Bachelor's degree; (4) the participants age was 31–40 (30%), 41–50 (35%), and 51–60 (31%); and (5) 43% speak two languages and 385 speak three or more languages.

Demographic characteristics for the participants from the Church are as follows: (1) 31% were women, 69% were men; (2) 14% were 31–40 years old, 31% were 41–50, 31% were 51–60 and 21% were 61 years of age and older; (3) 41% never lived in another culture; (4) 39% were college graduates and 32% had a Master's degree.

Demographic characteristics for the participants from College sample are as follows: (1) 34% were men, 65% were women; (2) 91% were 18–21 years of age; (3) 35% never lived in another culture while 33% lived 3 months or less in another culture; and (4) 96% were high school graduates (currently enrolled in college).

Demographic characteristics for the high school participants are as follows: overall, across the eight high school cohort groups, the following demographic characteristic emerge: (1) 63% men, 37% women; (2) 94% age 17 or under, 5% were 18–21 years of age; (3) 43% never previously lived in another country while another 45% lived less than 3 months in another country; (4) 76% did not complete high school (currently enrolled), 9% were high school graduates; (5) the participants were from the following countries: Austria (4%, 60), Brazil (8%, 129), Costa Rica and Ecuador (6%, 90), Germany (36%, 566), Hong Kong (3%, 50), Italy (15%, 228), Japan (15%, 227) and the U.S. (13%, 201).

3.2. Analytical protocols using confirmatory factor analysis

Before proceeding directly into the various tests undertaken to analyze responses, a brief discussion concerning some key aspects of the process of Structural Equation Modeling (SEM), which includes statistics such as confirmatory factor analysis, should be articulated. When attempting to test the “fit” of a model, you are assessing the degree to which the hypothesized model is consistent with the data. This is purely an empirical/statistical process. However, if you just rely on summary statistics to make your decisions, then you often get into the well-known situation of “garbage in, garbage out.” To avoid this situation, it is important to keep in mind the following points.

First, as Byrne (1998) suggests, “fit indices provide no guarantee whatsoever that a model is useful . . . Fit indices yield information bearing only on the model's lack of fit . . . They can in no way reflect the extent to which the model is plausible; this judgment rests squarely on the shoulders of the researcher” (p. 119). In this sense, fit indices are better for providing evidence concerning how poor a fit a model is to the data. Similar to hypothesis testing or theory construction, findings from the data are more suited for addressing whether the hypotheses/theories are weak, inconclusive, implausible, not whether the model reflects truth or reality.

Second, as starkly noted by Darden (1983), “a good fit for a model proves nothing. There are conceivably many models that could fit as well – or in some cases – better. In fact, a poor fit would tell you more; this would be more conclusive evidence that the model is not supported by the data” (p. 28). In this sense, a researcher can always “tinker” with the constraints/parameters of the model to make “a better fit”, but these changes may be meaningless from the standpoint of the theory being assessed.

Third, and perhaps most directly applicable to the purpose of the current study, Byrne (1998) further noted that, “evaluation of model fit should derive from a variety of sources and be based on several criteria that can assess model fit from a diversity of perspectives” (p. 103). This suggests that a number of criteria should be brought to bear on assessing the adequacies of different models. These criteria typically include parsimony, cross-sample consistency, interpretability, and *theoretical relevance*. In some cases, the application of these various criteria may result in equivocal recommendations. When this occurs, it is the researcher who ultimately determines what is best, given the empirical evidence and theoretical constructs being tested. This speaks directly to the initial validation study (Hammer et al., 2003) in which there was evidence that could have led to the choice of the seven-dimension model and evidence that led to choice of the five-dimension model. At that time, the criteria of parsimony led to our recommendation of the five-dimension solution rather than the seven factor model (the original DMIS conceptualization). However, research should be evolving and developmental; it should assist in refining and amending our theoretical notions of the phenomenon under study. With this in mind, the current Phase 3 validation study was undertaken.

Finally, and most paramount, “we cannot evaluate and interpret results as if they were divorced from the theory driving one’s study or from other conceptual and philosophical issues that bear upon the findings” (Bagozzi & Yi, 1988, p. 92). This process of interpreting results from ongoing research efforts within the theoretical context (to which the empirical findings are directed) is a back-and-forth dynamic, moving from social science methods testing specific theoretical propositions to interpreting those findings within the theoretical context and finally providing meta-level commentary on the viability of the theory itself. It is this process which is most directly addressed in the development of the IDI and is the grounding principle that guides subsequent empirical investigation of the IDI.

4. IDI v3 results

4.1. Analysis 1: confirmatory factor analysis: overall analysis across all groups

Confirmatory factor analysis was employed to test whether the constructs of the DMIS model would be consistent, or fit, the data collected from the IDI. A number of statistical procedures were employed to test the adequacy of this fit. First, the ratio of chi-square to degrees of freedom (χ^2/df) was used to assess the discrepancy between the proposed model and the data. Byrne (1998) suggested that the χ^2/df should be less than two for an adequate fit. However, a number of researchers have noted that χ^2 is sensitive to departures for multivariate normality and sample size (Long, 1983; Marsh, Balla, & McDonald, 1988). In fact, for large sample sizes, χ^2 becomes increasingly large, resulting in the false rejection of the model. Because we have such a large sample, we will treat the information from the χ^2/df statistics for comparison purposes among the competing models, not as absolute measures of fit.

Second, Jöreskog and Sörbom’s (1984) Goodness-of-Fit Index (GFI) is a generalized estimation criterion, which ranges from zero (no fit) to one (perfect fit). Typically, a GFI above .90 is thought to be an adequate fit.

Third, the Root Mean-square Residual (RMR) is an estimate obtained by comparing the values of variances and covariances predicted by the model with the actual variances and covariances ascertained from the data. The larger the RMR, the greater the discrepancy between the model and the data, with zero representing a perfect fit. An RMR in the .05 area is thought to reflect an adequate fit (Diamantopoulos & Siguaw, 2000).

Finally, the Root Mean Square Error of Approximation (RMSEA) provides a fit of the data taking into consideration the complexity of the model. Since there are many items involved in the IDI and multiple dimensions of the DMIS, it was felt that RMSEA should be employed, especially when models of different complexity are being compared. Browne and Cudeck (1989) recommend a criterion of .08 or less for the RMSEA for determining a good fit of the data.

Using the above criteria, three models were tested via confirmatory factor analysis: (1) based on the original theoretical model proposed by Bennett (1986, 1993), an elaborated seven-dimensional model which includes the dimensions of Denial, Defense, Reversal, Minimization, Acceptance, Adaptation, and Integration (Encapsulated Marginality) was examined; (2) based on the findings from Paige et al. (2003), a more parsimonious five-dimensional model which was comprised of the dimensions of Minimization, Reversal, and Integration, as well as the two merged dimensions of Denial/Defense and Acceptance/Adaptation was investigated; and (3) based on the theoretical supposition proposed by Bennett (1986, 1993) that ethnocentrism and ethnorelativism are two underlying dimensions of an individual’s orientation toward cultural differences, a very parsimonious two-dimensional model consisting of an ethnocentric orientation (the merging of Denial, Defense, Minimization, and Reversal) and an ethnorelative orientation (the merging of Acceptance, Adaptation, and Integration) was also tested.

As can be seen in Table 2, the confirmatory factor analysis of the seven-dimensional model indicated that it was the best fit among the three models compared ($\chi^2/\text{df}=9.45$, GFI=.91, RMR=.05, and RMSEA=.04). The confirmatory factor

Table 2
Summary of fit statistics for various cohorts.

Cohort (N)	Two-dimension (1174 df)			Five-dimension (1165 df)			Seven-dimension (1154 df)		
	χ^2	GFI	RMR	χ^2	GFI	RMR	χ^2	GFI	RMR
NGO (213)	2043.9	.714	.117	1795.1	.755	.097	1724.2	.767	.096
College (2579)	9994.1	.827	.101	4889.9	.925	.502	4205.6	.937	.049
Church (143)	2072.1	.664	.114	1834.9	.689	.127	1771.1	.698	.119
HS-Austria (114)	1835.2	.631	.120	1660.3	.711	.100	1651.2	.672	.100
HS-Brazil (180)	1708.3	.740	.113	1686.2	.746	.114	1634.7	.754	.112
HS-Costa Rica (149)	1690.9	.702	.126	1623.3	.719	.120	1579.8	.725	.118
HS-Germany (547)	2620.4	.822	.080	1930.5	.875	.058	1879.7	.879	.057
HS-Hong Kong (94)	1913.3	.587	.109	1881.3	.590	.106	1824.8	.600	.108
HS-Italy (294)	2009.0	.775	.096	1744.0	.811	.084	1680.3	.820	.081
HS-Japan (275)	1902.6	.778	.091	1672.8	.840	.080	1649.8	.866	.078
HS-USA (175)	1987.6	.687	.136	1680.5	.742	.103	1651.9	.748	.102
TOTALa (4763)	15,941.8	.849	.096	8746.7	.924	.056	7778.4	.933	.053
TOTALb ^a (2184)	7703.1	.847	.096	5253.3	.900	.068	4929.3	.906	.066

^a TOTALa includes all data from all cohorts; TOTALb includes all data except the College cohort.

Table 3

Differences in chi-squares to test adequacy of fits (two-dimensional solution used as baseline).

Cohort	Diff b/n 2- and 5-Dim df=9	Diff b/n 2- and 7-Dim df=20	Diff b/n 5- vs. 7-Dim df=11
NGO	$\chi^2 = 248.2, p < .0001$	$\chi^2 = 319.7, p < .0001$	$\chi^2 = 70.9, p < .001$
College	$\chi^2 = 5104.2, p < .00001$	$\chi^2 = 5788.5, p < .00001$	$\chi^2 = 684.3, p < .0001$
Church	$\chi^2 = 237.2, p < .0001$	$\chi^2 = 301.0, p < .0001$	$\chi^2 = 63.8, p < .001$
HS-Austria	$\chi^2 = 174.9, p < .0001$	$\chi^2 = 184.0, p < .0001$	$\chi^2 = 9.1, n.s.$
HS-Brazil	$\chi^2 = 22.1, p < .01$	$\chi^2 = 73.6, p < .001$	$\chi^2 = 51.5, p < .001$
HS-Costa Rica	$\chi^2 = 67.6, p < .001$	$\chi^2 = 111.1, p < .001$	$\chi^2 = 43.5, p < .001$
HS-Germany	$\chi^2 = 689.9, p < .0001$	$\chi^2 = 740.7, p < .0001$	$\chi^2 = 50.8, p < .001$
HS-Hong Kong	$\chi^2 = 32.0, p < .01$	$\chi^2 = 88.5, p < .0001$	$\chi^2 = 56.5, p < .001$
HS-Italy	$\chi^2 = 265.0, p < .0001$	$\chi^2 = 328.7, p < .0001$	$\chi^2 = 63.7, p < .001$
HS-Japan	$\chi^2 = 229.8, p < .0001$	$\chi^2 = 252.8, p < .0001$	$\chi^2 = 23.0, p < .05$
HS-USA	$\chi^2 = 307.1, p < .0001$	$\chi^2 = 335.7, p < .0001$	$\chi^2 = 28.6, p < .01$
TOTALa	$\chi^2 = 7195.1, p < .00001$	$\chi^2 = 8163.4, p < .00001$	$\chi^2 = 968.3, p < .0001$
TOTALb	$\chi^2 = 2449.8, p < .00001$	$\chi^2 = 2773.8, p < .00001$	$\chi^2 = 324.0, p < .0001$

^a TOTALa includes all data from all cohorts; TOTALb includes all data except the College cohort.

analysis of the five-dimensional DMIS was a lesser (although plausible) fit of the data ($\chi^2/\text{df} = 10.62$, GFI = .89, RMR = .06, and RMSEA = .05).

Table 3 presents both aggregate data analysis and individual cohort analysis. As reflected in Table 3 regarding the overall aggregate data, further evidence of the significantly better fit for the seven-dimensional model was found by comparing the chi-square decrease vs. the decrease in degrees of freedom for the two models. For the seven-dimensional model, chi-square was 7778.4 with 1154 degrees of freedom, while for the five-dimensional model, the chi-square was 8746.7 with 1165 degrees of freedom. The resultant differences in chi-square and degrees of freedom indicate a significantly better fit for the seven-dimensional solution ($\chi^2_{\text{diff}} = 968.3$ and $\text{df}_{\text{diff}} = 11, p < .0001$). Overall, these results provide persuasive support for the seven-dimensional model over the five-dimensional model.

It was also decided to also test the most parsimonious version of the DMIS model, namely, a two-dimensional representation (ethnocentrism and ethnorelativism). The fit for this simpler model ($\chi^2/\text{df} = 25.16$, GFI = .72, RMR = .12, and RMSEA = .07) did not meet acceptable statistical standards.

Further evidence for the better fit of the seven- and five-dimensional models over the two-dimensional one was obtained by comparing the differences in chi-squares (see Table 3). The chi-squares (and their dfs) for the two-, five-, and seven-dimensional solutions are, respectively, 15,941.8 (1174), 8746.7 (1165), and 7778.4 (1154). As noted in Table 3, for the two- and five-dimensional comparison, the $\chi^2_{\text{diff}} = 7195.1$ and $\text{df}_{\text{diff}} = 9$ was highly significant ($p < .00001$); similarly, for the two- and seven-dimensional comparison, the $\chi^2_{\text{diff}} = 8163.4$ and $\text{df}_{\text{diff}} = 20$ was also highly significant ($p < .00001$). These results suggest that the two-dimensional model is the worst fit of the IDI data.

4.2. Analysis 2: cross-validation of the results: individual cohort analysis

The above results reflect an analysis of an aggregate of 11 different cohorts. It is possible that there may be differences from the aggregate results when individual cohort samples are examined. While the overall analysis speaks to the cross-cultural generalizability of the seven factor model, individual cohort analysis speaks more directly to the culture-specific applicability of the seven factor model (i.e., across specific cultural communities).

Table 3 also presents summaries of the individual cohort analyses along with Table 4 (which also presents reliability analysis discussed below). As can be seen from these tables, for 10 of the 11 cohorts, the seven-dimensional solution was the best fit to the data. For the sample of Austrian high school students, both the seven- and five-factor models were equally good fit to the data. That is, there was no significant difference between the five and seven factor solutions ($\chi^2_{\text{diff}} = 9.1$ and $\text{df}_{\text{diff}} = 11, n.s.$). In all the other ten cohort groups, the seven factor solution was a significantly better fit to the specific cultural data. Stated another way, the seven factor solution fit equally well to the data (compared to the five factor model) for the Austrian sample and was significantly superior to the five dimensional solution for the remaining ten cohort groups. Overall,

Table 4

Reliabilities (coefficient alpha) of IDI scales by cohort groups.

Scales	Total	College	HS1	HS2	HS3	HS4	HS5	HS6	HS7	HS8	Church	NGO
Denial	0.66	0.66	0.58	0.49	0.40	0.57	0.38	0.65	0.46	0.65	0.76	0.68
Defense	0.72	0.73	0.54	0.70	0.54	0.60	0.51	0.69	0.49	0.67	0.77	0.77
Reversal	0.78	0.81	0.77	0.73	0.69	0.72	0.68	0.72	0.63	0.80	0.74	0.76
Minimi-zation	0.74	0.76	0.65	0.48	0.62	0.70	0.62	0.66	0.67	0.67	0.78	0.79
Accept-ance	0.69	0.72	0.74	0.54	0.64	0.62	0.54	0.61	0.62	0.73	0.71	0.74
Adapta-tion	0.71	0.75	0.53	0.66	0.51	0.63	0.68	0.62	0.63	0.74	0.67	0.69
Cultural Disengag	0.79	0.83	0.73	0.52	0.50	0.73	0.59	0.73	0.65	0.73	0.74	0.72
DO	0.83	0.84	0.74	0.81	0.76	0.76	0.77	0.80	0.73	0.82	0.83	0.87
PO	0.82	0.84	0.76	0.77	0.75	0.75	0.78	0.76	0.74	0.82	0.80	0.86

Table 5

Correlations among seven latent variables.

	Denial	Defense	Reversal	Minim	Accept	Adapt	Cul. Disengage
Denial	1.000						
Defense	.830	1.000					
Reversal	.358	.367	1.000				
Minim	-.033	.062	.036	1.000			
Accept	-.169	-.111	.012	.014	1.000		
Adapt	-.185	-.084	.124	.144	.638	1.000	
Cul. Disengage	.227	.110	.433	.007	-.073	.033	1.000

there is clear evidence across specific cultural groups that the seven-dimensional solution is quite robust and provides the best fit to the data.

4.3. Scale reliabilities for seven-dimension solution

Given the weight of the aggregate and the individual cohort results, it was felt that the seven-dimensional model was the best fit of the present data. In light of this conclusion, reliability coefficients were computed for the seven dimensions (Table 4). The alpha coefficients for the seven scales were: $\alpha = .66$ for Denial, $\alpha = .72$ for Defense, $\alpha = .78$ for Reversal, $\alpha = .74$ for Minimization, $\alpha = .69$ for Acceptance, $\alpha = .71$ for Adaptation and $\alpha = .79$ for Cultural Disengagement. In addition, Table 4 presents reliabilities (coefficient alpha) for the overall Developmental Orientation (DO) scale (.83) and for the overall Perceived Orientation (PO) scale (.82). Given the cross-cultural, linguistically diverse samples, it may be concluded that all of the scales had good levels of inter-item reliabilities.

4.4. Correlations among the seven dimensions

Table 5 presents the intercorrelations among the seven dimensions of the 50-item IDI (v3). Noteworthy about this correlation matrix are the following. First, there is a strong correlation between Defense and Denial ($r = .83$) and a strong correlation between Acceptance and Adaptation ($r = .64$). Second, Reversal is positively correlated with Denial (.36) and with Defense (.37) and not substantially correlated with Acceptance (.01) or Adaptation (.12), consistent with the finding and explanation from the initial validation study (Hammer et al., 2003).

Third, the very weak correlations of Minimization with both the more monocultural orientations (Denial, Defense, Reversal) and the more intercultural orientations (Acceptance, Adaptation), suggest Minimization exists as a transitional orientation between the more monocultural and intercultural orientations. As a transitional orientation, Minimization is thus more accurately conceptualized as being not ethnocentric (monocultural) yet also not as interculturally competent/sensitive as Acceptance or Adaptation.

Fourth, there are negative correlations between Defense and Denial scales and the Acceptance and Adaptation scales which provide support for the developmental continuum.

Fifth, the scale of Encapsulated Marginality is correlated most strongly with Reversal but not significantly correlated with Acceptance or Adaptation. This is a finding that is not hypothesized by the original DMIS model. That is, Encapsulated Marginality (EM) is theoretically a dimension of Integration, which follows Adaptation according to DMIS propositions. Therefore, there should be a strong correlation to Adaptation, slightly less to Acceptance, and so on down the developmental continuum. In fact, DMIS theory would predict a negative correlation with Denial, Defense and Reversal. The results indicate that Encapsulated Marginality, however, is positively correlated with Reversal but not correlated with Acceptance or Adaptation. Based on these results, a review of the meaning of the items in the EM scale reveals that the measure is more of a measure of Cultural Disengagement—a sense of disconnection from one's cultural community. This suggests that the scale is not a measure of EM—or an assessment of a dimension of cultural identity formation/transformation. When understood in this manner, it makes sense that the Cultural Disengagement (CD) scale would have some correlation to Reversal as it is likely capturing a shared sense of emotional distance from one's own cultural group. Cultural Disengagement is also then clearly understood as a separate measure within the IDI, not a scale that is conceptually found on the intercultural competence/sensitivity continuum.

Finally, in Hammer et al. (2003), both the seven factor model (Denial, Defense, Reversal, Minimization, Acceptance, Adaptation and Encapsulated Marginality) and the five factor solution (DD (denial/defense), Reversal, Minimization, AA (acceptance/adaptation) and Encapsulated Marginality) were both a good fit to the data. The empirical overlap between Defense and Denial, and between Acceptance and Adaptation, the smaller and more homogeneous sample in the Hammer et al. (2003) study, and the reliance on the criteria of parsimony may likely have contributed to a conclusion that the five-dimensional solution was the best fit to the data. The current study, using a larger, cross-culturally diverse sample, strongly indicates the seven factor solution (Denial, Defense, Reversal, Minimization, Acceptance, Adaptation) along with the factor of Cultural Disengagement (formally labeled as encapsulated marginality) is clearly the best fit to the data. The results from this study are consistent across all respondents as well as for each of the cohort groups when analyzed separately.

While some modifications to the original DMIS orientations arise as a result of the collective Phase 1, 2, post Phase 2 testing, and the current Phase 3 validation studies of the IDI, an overall conclusion from these various efforts is that the main theoretical insights offered by the DMIS are consistently confirmed. These studies provide overall support for the DMIS as a fundamentally sound theoretical framework and simultaneously support the modifications to the DMIS framework presented in this article.

5. Additional validity tests of the IDI

5.1. IDI results and population estimation

An assumption is often made that a particular construct examined through social science research methods is normally distributed within the population. This is both a statistical assumption employed when using parametric statistics. When a researcher does not wish to assume normal distribution of a phenomena (construct) under investigation, then non-parametric statistics are relied upon. Recently, [Bennett \(2009\)](#) questions this assumption of “normal distribution of scores” of the IDI and claims that the orientation of Minimization is overestimated in the IDI. As Bennett comments:

The disadvantage of the IDI is that it sacrifices ideographic data in favor of nomothetic data necessary for group comparisons. What this means is that the instrument is not very sensitive to individual differences; it tends to overestimate the “normative condition—Minimization—and underestimate the extent of more ethnocentric or more ethnorelative positions. For individuals, this tendency can be counteracted through individual interpretations taking into account the distribution of responses on the instrument. But for a group profile, individual variations are summarized as group data, and it is therefore not possible to counteract the over-attribution of Minimization with the IDI data. Consequently, the IDI should be used cautiously and only with other measures, such as the qualitative data reported in descriptive studies, to discover the overall intercultural sensitivity of groups (p. 8–9).

Unfortunately, [Bennett \(2009\)](#) provides no empirical evidence to support such an assertion. In fact, there is evidence that [Bennett's \(2009\)](#) unsupported observation that the IDI overestimates the normative condition is not accurate.

The DO scores of the IDI are “normalized” on a scale with a mean of 100 and a standard deviation of 15. This then, based on a normal distribution, produces the following standard deviation percentages: 2.28% (Denial), 13.59% (Polarization, i.e., Defense/Reversal), 68.26% (Minimization), 13.59% (Acceptance), and 2.28% (Adaptation). It is puzzling why Dr. Bennett would make such an unsubstantiated claim insofar as he is well aware that the “normed” cross-cultural sample for IDI v2 of 1000 individuals reflected the normal distribution of IDI DO scores.

Further, a review of the current IDI v3 validation study data (this article) from 4763 individuals provides empirical support for the normal distribution of the IDI. Specifically, the distribution of IDI DO scores within this large, international, cross-cultural sample reveals the following distribution: Denial: 2.6%; Polarization: 14%; Minimization: 67%; Acceptance: 14.9%; and Adaptation: 1.6%.

Additionally, a review of IDI v3 data collected with a separate sample of 4654 subjects from the United States reveals the following distribution: Denial: 3.5%; Polarization: 17.1%; Minimization: 63.5%; Acceptance: 14.4%; and Adaptation: 1.5%. Combining both sample distribution of IDI scores of the 9417 respondents indicates the following: Denial: 3.05%; Polarization: 15.55%; Minimization: 65.25%; Acceptance: 14.65%; and Adaptation: 1.55%. These results from the initial, IDI v2 sample of 1000 subjects along with the two studies completed, clearly suggest a normal distribution, indicating there is no overestimation of Minimization. Further, these results clearly indicate the IDI is very sensitive to individual differences, and the IDI does not underestimate the more ethnocentric orientations. In short, the IDI is an equally robust and valid assessment for both individuals and groups.

5.2. Criterion validity testing: study 1

An additional study was conducted by the author with Peter Bye of MDB Group, an intercultural consultancy, to assess the criterion (i.e., concurrent) validity of the IDI. In this study, a total of 71 recruiters of the U.S. staffing function within a high tech, multinational organization participated. They worked in six teams, located throughout the United States (team 1, $n = 15$; team 2, $n = 8$; team 3, $n = 13$; team 4, $n = 10$; team 5, $n = 13$; and team 6, $n = 12$). This study examines the criterion validity of the IDI by assessing the relationship between the team IDI Developmental Orientation scores and separately, individual IDI results of each member of the six teams to the degree to which cultural diversity was reflected in each of the six the team's U.S. hiring results in the organization during 2008. The IDI was administered to the 71 recruiters near the time the hiring results were calculated.

Achievement of the Diversity and Inclusion (D & I) goals by each of the staffing teams was determined in the following method. First appropriate external benchmarks were selected for each level in the organization. These were a combination of two sets of U.S. Government workforce data:

- U.S. Equal Employment Opportunity Commission 2006 EEO-1 reporting data for the appropriate industry sector in which the organization does business and seeks talent. The benchmark population consisted of over 1.2 million employees in several thousand reporting units. This EEO-1 data is provided by race and gender in the nine primary EEO-1 job categories

Team #	N	Team DO score	Team Benchmarks Missed
1	13	89.42	12
2	8	91.30	11
3	15	98.96	8
4	13	99.43	8
5	10	108.63	5
6	12	108.62	5

Fig. 2. Staffing team IDI developmental orientation scores and total diversity hiring benchmarks missed.

of Officials & Managers, Professionals, Technicians, Sales Workers, Office & Clerical, Craft Workers, Operatives, Laborer, and Service Workers.

- U.S. Census 2000 data for representation by race and gender in the CEO and General Manager part of the U.S. workforce.

Representation of each race group in the 2008 hires at each level was compared to the relevant benchmark. Results were categorized as either exceeding the benchmark, almost making the benchmark by being within 20% of the benchmark, or missing the benchmark by at least 20%. This analysis was conducted separately for the part of the U.S.-based organization served by each of the six staffing teams. A rating was developed for each of the six teams that was equal to one plus the number of individual *benchmarks missed*. A rating of one would indicate no benchmarks were missed.

Two separate correlational analyses are presented below. The first uses the overall Developmental Orientation score (DO) for each of the six teams and correlates this with each team's number of benchmarks missed. Fig. 2 presents the mean team DO score and their number of benchmarks missed. Spearman's Rho correlational statistic was used due to the small sample size ($n = 6$). Results indicate Spearman's Rho correlation is $-.96$ ($p = .000$; $n = 6$), a very strong, negative correlation between overall team DO scores and team benchmarks missed. The coefficient of determination is quite robust (.92); suggesting that 92% of the variance in the number of benchmarks missed is accounted for by the IDI team DO score. These results indicate that the *lower* the team's DO score (lower intercultural competence/sensitivity), the *more* the team's Diversity and Inclusion benchmarks were missed.

While this clearly indicates that overall team level of intercultural competence is related to team effectiveness in meeting diversity and inclusion staffing goals, this team level analysis does not directly indicate the degree to which individual team member DO scores were also reflective of this relationship. That is, to what degree are individual team member scores related to overall team effectiveness in meeting diversity and inclusion staffing goals?

Pearson product-moment correlations (Pearson's r) were run between *individual* IDI Developmental Orientation scores ($n = 71$) and each individual's overall team Diversity goal achievement rating. The resulting correlation was $-.43$ ($p = .000$; $n = 71$), indicating a moderately strong, negative correlation between the individual's overall Developmental Orientation and lack of success in meeting Diversity goals for recruitment. The coefficient of determination is .19, indicating 19% of the variance in the overall team success in meeting D & I benchmarking goals is accounted for by individual member DO scores. These results indicate that the *lower* the score on the IDI (the lower the level of intercultural competence), the *more* the team "missed" meeting diversity and inclusion goals. Said another way, higher levels of intercultural competence are associated with greater results in meeting diversity and inclusion objectives.

This fairly strong correlation is somewhat surprising as it was not possible to assign individual missing benchmarking data to each individual. This was due to the fact that in a number of cases, there were multiple individuals working together to produce the team result in terms of benchmarking misses. We suspect, therefore, that the *team level* benchmarking data tends to "water down" the correlation between the individual and his/her own actual productivity and accomplishment around D & I goals. Therefore, it is possible that the correlation between the individual DO scores and the team benchmark misses of $-.45$ is conservative.

Overall, both correlational analysis results (i.e., team and individual IDI DO score analysis) lend strong, predictive validity support to the IDI instrument.

5.3. Criterion validity testing: study 2

Hammer (2005) conducted one of the largest examinations of the development of intercultural competence using the IDI and various outcomes of a ten-month, study abroad experience for high school students throughout the world. The IDI was "back translated" from English into the following languages: Spanish, Portuguese, Italian, German, Chinese, and Japanese. Students who were going on an AFS study abroad program and control group participants from Austria, Brazil, Costa Rica, Ecuador, Germany, Hong Kong, Italy, Japan and the United States were identified. Control group participants were identified using the "best friend" methodology. The study used a pretest, posttest, post-posttest, control group design. A total

Table 6

IDI developmental orientation scores for pax and control groups.

Pax/control group	DO score: pretest	DO score: posttest	DO score: post-posttest
<i>Monocultural orientation</i>			
Pax	76 (<i>n</i> = 471)	84 (<i>n</i> = 262)	84 (<i>n</i> = 180)
Control	73 (<i>n</i> = 243)	76 (<i>n</i> = 97)	78 (<i>n</i> = 82)
<i>Transitional (M) orientation</i>			
Pax	96 (<i>n</i> = 630)	96 (<i>n</i> = 464)	95 (<i>n</i> = 328)
Control	94 (<i>n</i> = 178)	91 (<i>n</i> = 85)	89 (<i>n</i> = 53)

of 1500 students and 638 control group students completed the pretest and therefore participated in the study. In addition to data gathered from international study abroad organization and control group students, the student's own family, and the student's host family also completed questionnaires. In addition, E-journals were also kept, but were not included in the analysis of this study.

Data from this effort is analyzed to test the criterion validity of the IDI. Specifically, in addition to the IDI overall DO score, the following "outcome" measures were examined: (1) knowledge of the host culture, (2) intercultural anxiety (degree of discomfort interacting with people from other cultures), (3) intercultural friends (percentage of friends individual maintains with people from other cultures), and (4) intercultural satisfaction (degree of satisfaction/enjoyment with study abroad experience; assessed only at the posttest). In addition, a measure of cultural similarity (degree to which cultural differences were perceived to exist between participants own culture and host culture) was used initially to confirm the need for intercultural competence among participants. Reliabilities of all measures ranged from satisfactory (.66) to very good (.92).

Results from the *cultural similarity* measure indicate both the students (pax) and the host families compare their countries as not similar, supporting the conclusion that the cultural experiences the students encountered during their study abroad did involve challenges in adapting to cultural differences. In other words, intercultural competence was a salient aspect of the study abroad experience for both the students and their host families.

In order to examine the criterion validity of the IDI, subjects IDI overall pretest Developmental Orientation (DO) scores were categorized into (1) a more monocultural orientation (Denial/Polarization DO scores), (2) a more transitional (Minimization) orientation (Minimization DO scores), and (3) a more intercultural orientation (Acceptance/Adaptation DO scores). Subsequent use of the more intercultural orientation scores had to be dropped due to an insufficient number of subjects in posttest and post-posttest data gathering.

Table 6 indicates the mean IDI DO scores for the pax and control group participants by monocultural orientation and transitional (Minimization) orientation.

Statistically significant differences were found. Overall patterns reveal that the pax moved from a monocultural orientation to a transitional orientation while pax who had a transitional (Minimization) orientation prior to their study abroad experience did not move at all in their intercultural competence (overall DO score). Control group participants did not significantly move in their overall DO scores.

Therefore, significant gains in intercultural competence as assessed by the IDI was obtained for those pax students who started their study abroad program with a monocultural orientation and developed their capabilities to the transitional (Minimization) orientation—achieving orientation similarity to those pax students who started the program with a transitional (Minimization) orientation but remained at the same intercultural competence level at the posttest and post-posttest administration.

Do the gains obtained by those students who moved from a pretest monocultural orientation to posttest transitional (Minimization) orientation and maintained this transitional orientation through the post-posttest reflect significant changes in outcome measures compared to those pax students who started the program with a transitional orientation and remained in this Minimization orientation at the posttest and post-posttest assessments (criterion validity focus)?

T-tests were run between group 1 (monocultural orientation, pretest pax students) and group 2 (transitional (Minimization) orientation, pretest pax students) on the students pretest outcome measures. Significant differences, predictive of IDI development, are found: pax DO scores are significantly different between group 1 (monocultural orientation group) and group 2 (transitional orientation group). Specifically, *T*-test results comparing group 1 (monocultural orientation group), and group 2 (more intercultural competent, transitional orientation group) reveals the following (1) significantly less cultural knowledge of the host culture, (2) significantly greater intercultural anxiety, and (3) significantly fewer intercultural friends.

T-tests run between groups 1 and 2 on the students posttest outcome measures indicate no significant differences in IDI DO posttest scores between group 1 and 2 pax participants. In other words, the initially more monocultural students (group 1) developed (gained 8 points on the IDI) to move their orientation into the transitional (Minimization orientation); the same orientation as group 2 maintained between the pretest and posttest. Specifically, group 1 (monocultural group) scores in the outcomes below substantially increased and essentially matched group 2 outcome scores so that no significant differences between groups 1 and 2 on the following posttest measures resulted: (1) knowledge of the host culture, (2) intercultural anxiety, (3) satisfaction with the study abroad experience (only measured at posttest), and (4) intercultural friendships.

Finally, *T*-tests were run between groups 1 and 2 scores obtained after the pax students returned to their home culture (3–6 months). Results indicate that the posttest gains in IDI DO scores for group 1 were maintained through the post-posttest, mirroring the same relationship in IDI DO scores found in the posttest administration. The following post-posttest

findings emerged: no significant differences are found between groups 1 and 2 on the post-posttest outcome measures: (1) knowledge of host culture, (2) intercultural anxiety, and (3) intercultural friendships.

Overall, these research results provide strong criterion validity to the IDI. Specifically, the observed increases in study abroad outcomes of knowledge of the host culture, intercultural anxiety, intercultural friendships, and satisfaction with the study abroad experience are significantly associated with increases in intercultural competence as assessed by the IDI.

5.4. Readability validation

Examination of readability of written documents has been conducted notably for school and university textbooks for decades (Gibson, Gruner, Brooks, & Petrie, 1970; Sellnow, Child, & Ahlfeldt, 2005; Schneider & Walter-Reed, 2009) and across disciplines, including communication (Schneider & Walter-Reed, 2009), psychology (Torres & Roig, 2005) and science and health care (Goldbort, 2001; White, 2000). Readability refers to “the ease of comprehension because of style of writing” (Harris & Hodges, 1995, p. 203). Comprehension is concerned with “understanding and applying knowledge from a written test (Schneider & Walter-Reed, 2009), clearly an important aspect of questionnaire completion.

The overall DO and PO IDI scales along with the individual subscale reliabilities summarized in Table 3 suggest that the culturally diverse high school samples (HS1 through HS8) maintain satisfactory reliability similar to that found with the slightly older college sample and with the older, adult samples of the NGO and church organization. These results support the applicability of the IDI to a wide range of age groups, from high school years through older adults.

In this readability analysis, the applicability of the IDI to various age groupings based on an examination of the readability level of the IDI was completed. This is important to determine, as readability may not always correlate with the age of respondent. For example, adults may possess a “reading comprehension level” below college or even high school grade levels. Readability analysis, therefore, provides an objective indicator of the level of reading literacy needed to comprehend the information presented in the IDI. This readability analysis was conducted using the English IDI version (results would not be expected to differ across other language versions).

The protocols for conducting the readability analysis of the IDI involved submitting the instructions page along with the 50 questionnaire items for analysis to a readability calculation software program developed by Micro Power & Light Company (version 7.8, 1995). While there are a number of readability formula's identified as particularly reliable (Schneider & Walter-Reed, 2009), the Fry (1977) formula and the Flesch Reading Ease Formula were used because they both cover a wide range of reading levels, from early elementary grades through college. The Fry formula is based on multisyllabic words and sentence length; two of the strongest predictor variables of reading difficulty (Schneider & Walter-Reed, 2009) while the Flesch Reading Ease Formula (Flesch, 1948) considers the number of words, syllables and sentences in adult reading materials (Micro Power & Light Company, 1995). The mean number of words surveyed was 1142 (IDI instructions and 50 items). This sample size is considered quite adequate compared to the generally accepted standard of 300–400 words identified by reading specialists (Micro Power & Light Company, 1995).

The results indicate that the mean grade level (United States educational system) for the IDI for the Flesch Reading Ease is 56, which is indicative of 10th–12th grade reading comprehension levels. The results from the Fry formula indicate an 11th grade reading level. These results indicate the IDI, is appropriate for students and/or adults who possess a 10th grade high school reading comprehension reading level. This would translate roughly into respondents who are approximately 14–15 years of age or above, providing their reading difficulty level is comparable to the 10th grade level.

The IDI instruction page was also analyzed separately from the actual IDI 50 items. These results were comparable to the overall reading sample results. Specifically, for the instructions page, the mean number of words is 450 with a Flesch Reading Ease score of 58 (10th–12th grade range) and the Fry formula results indicating an 11th–12th grade range. For the 50 item sample, the mean word length was 662 and the Flesch Reading Ease score was 55 (10th–12th grade range) and Fry formula results were at the 11th grade range.

Overall, the results from the reliability analysis support the applicability of the IDI to 10th grade (roughly 15 years of age) and above educational levels.

6. Conclusion

A number of concluding observations arise from the additional validity testing of the Intercultural Development Inventory (IDI v3) presented in this article. First, the confirmatory factor analysis confirms the basic orientations toward cultural difference originally explicated by Bennett (1986, 1993) in the Developmental Model of Intercultural Sensitivity. Second, the inter-scale correlations support the theoretically proposed, developmental formulation from Denial through Adaptation, termed the intercultural competence continuum. Third, the current analyses offers strong support for the cross-cultural generalizability, validity and reliability of the IDI v3 measure, to include the individual sub-scales of Denial, Defense, Reversal, Minimization, Acceptance and Adaptation as well as the overall Developmental Orientation (DO) and overall Perceived Orientation (PO) scales.

This additional testing has also suggested modifications to some of the DMIS orientation descriptions. First, this testing has resulted in a re-interpretation of what was initially termed, the Encapsulated Marginality (EM) scale. Encapsulated Marginality was initially conceived as one of two dimensions of the larger Orientation of Integration, which is theorized in the DMIS model, to represent the next developmental stage of intercultural competence beyond Adaptation.

However, the results from this current study indicate that the correlations of this scale to the other scales were not in the theorized direction. This led to a re-interpretation of the EM scale and renaming it the Cultural Disengagement (CD) scale. This scale appears to be a more general measure of the disconnection one can experience from one's primary cultural community(ies). As such, this is not a measure of encapsulated marginality. Further, this construct and measure of Cultural Disengagement is viewed in the current research as neither a sub-dimension of Integration nor as an Orientation located along the Developmental Continuum (Denial through Adaptation).

In addition, the original Integration items (Hammer et al., 2003) did not possess sufficient factor stability to produce a valid and reliable scale. In reviewing the DMIS conceptual formulations, it appears that Integration is primarily understood as a dimension of identity development. Yet the DMIS model is essentially a model of the development of intercultural competence/sensitivity; not a model of cultural/ethnic identity development. The research indirectly supports this distinction as the findings confirm a coherent, stable set of developmental orientations focused on intercultural competence/sensitivity. This research has not been able to confirm Integration (an identity orientation) within this developmental continuum. This is an area of further research.

Second, Minimization is found to be a transitional orientation toward cultural differences and commonalities, between the more monocultural (ethnocentric) orientations of Denial and Polarization (Defense, Reversal) and the more intercultural mindsets of Acceptance and Adaptation. This suggests that Minimization may be more interculturally competent than originally characterized in the DMIS. Nevertheless, Minimization, consistent with the DMIS, is also not sufficiently capable of understanding as deeply as needed various patterns of cultural difference nor is it able to easily identify and implement cross-culturally adaptive behavior or solutions to complex intercultural problems.

Third, IDI v3 results clearly posit Denial, Defense, Reversal, Minimization, Acceptance and Adaptation as the core concepts and scales arrayed along the developmental competence continuum (it also identifies Cultural Disengagement as an additional scale within the IDI; but one that is not located along the developmental continuum).

Fourth, countering unsubstantiated claims by Bennett (2009), analysis from three distinct validation samples (cumulatively over 10,000 subjects) provides empirical evidence of the normal distribution of IDI DO scores, supporting the application of the IDI results to both individual and group profile analysis.

Fifth, criterion validity testing from two additional studies reveals the predictive validity of IDI DO scores: (1) results indicate that the IDI has strong predictive validity toward bottom-line goals within organizations; namely, the achievement of diversity and inclusion goals in the recruitment and staffing function and (2) the IDI has predictive validity as tested within study abroad in terms of knowledge of host culture, intercultural anxiety, intercultural friendships, and post sojourn overall satisfaction with the study abroad experience.

Sixth, readability analysis of the IDI supports the cognitive development recommendation that the IDI is appropriate for high (secondary) school students (age 15 or above) or individuals with a 10th grade reading level.

Overall, these findings complement previous results demonstrating strong content and construct validity of the IDI across culture groups.

References

- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16, 74–94.
- Bennett, M. J. (1986). Towards ethnocentrism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Cross-cultural orientation: New conceptualizations and applications* (pp. 27–70). New York, NY: University Press of America.
- Bennett, M. J. (1993). Towards ethnocentrism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (pp. 21–71). Yarmouth, ME: Intercultural Press.
- Bennett, M. J. (2004). Becoming intercultural competent. In J. Wurzel (Ed.), *Toward multiculturalism: A reader in multicultural education* (2nd ed., pp. 62–77). Newton, MA: Intercultural Resource Corporation.
- Bennett, M. J. (2009). Defining, measuring, and facilitating intercultural learning: A conceptual introduction the Intercultural Education double supplement. *Intercultural Education*, 20(4 (Suppl. 1)), 1–13.
- Bhawuk, D. P. S., & Brislin, R. (1992). The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International Journal of Intercultural Relations*, 16(4), 413–436.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3), 185–216.
- Brislin, R. W. (1976). Comparative research methodology: Cross-cultural studies. *International Journal of Psychology*, 11(3), 215–229.
- Brislin, R. W. (1980). Translation and content analysis of oral and written materials. In H. C. Triandis, & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (pp. 389–444). Boston, MA: Allyn and Bacon.
- Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, 24, 445–455.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Darden, W. R. (1983). Review of behavioral modeling in marketing. In W. R. Darden, K. B. Monroe, & W. R. Dillon (Eds.), *Research methods and causal modeling in marketing*. Chicago: American Marketing Association.
- Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL: A guide for the uninitiated*. Thousand Oaks, CA: Sage Publications.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221–233.
- Fry, E. B. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of Reading*, 21, 242–252.
- Gibson, J. W., Gruner, C. R., Brooks, W. D., & Petrie, C. R. J. R. (1970). The first course in speech: A survey of U.S. colleges and universities. *Speech Teacher*, 19, 13–20.
- Goldbort, R. (2001). Readable writing by scientists and researchers. *Journal of Environmental health*, 63(8), 40–42.
- Hammer, M. R. (1989). Intercultural communication competence. In M. K. Asante, & W. B. Gudykunst (Eds.), *The handbook of intercultural communication*. Newbury Park, CA: Sage.
- Hammer, M. R. (1999). A measure of intercultural sensitivity: The Intercultural Development Inventory. In S. M. Fowler, & M. G. Mumford (Eds.), *The intercultural sourcebook*. Yarmouth, ME: Intercultural Press.
- Hammer, M. R. (2005). *Assessment of the impact of the AFS study abroad experience*. New York: AFS, International.

- Hammer, M. R. (2009). The Intercultural Development Inventory. In M. A. Moodian (Ed.), *Contemporary leadership and intercultural competence* (pp. 203–217). 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.
- Harris, R. L., & Hodges, R. E. (Eds.). (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.
- Jöreskog, K. G., & Sörbom, D. (1984). *LISREL-VI user's guide* (3rd ed.). Mooresville, IN: Scientific Software.
- Long, J. S. (1983). *Confirmatory factor analysis*. Beverly Hills, CA: Sage.
- Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391–410.
- Micro Power & Light Company. (1995). Dallas, TX.
- Moodian, M. A. (Ed.). (2009). *Contemporary leadership and intercultural competence*. Thousand Oaks, CA: Sage.
- 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. In R. M. Paige (Ed.) (Special issue: Intercultural development). *International Journal of Intercultural Relations*, 27, 467–486.
- Schneider, D. E., & Walter-Reed, M. (2009). Readability and basic course hybrid textbooks in human communication: A disparity of reading difficulty. *Communication Research Reports*, 26(4), 361–366.
- Sellnow, D., Child, J. T., & Ahlfeldt, S. L. (2005). Textbook technology supplements: What are they good for? *Communication Education*, 54, 243–253.
- Torres, M., & Roig, M. (2005). The Cloze Test Procedure as a test of plagiarism: The influence of text readability. *The Journal of Psychology*, 139, 221–231.
- White, F. (2000). Grading the readability of articles. *Canadian Journal of Public Health*, 9(1), 73.