

Parallel Session I

Tuesday, 12 Nov 2013

9.15—10.30am

Venue Theme		ID	Title	Presenter	Affiliation		
		A09	Paradigm Shift On Islamic Educa- tion In Indonesia	Aries Musnandar Munirul Abidin	UIN Maliki Malang, Indonesia		
A: Grand Ballroom	1: Teacher Education Philoso- phy, Poli- cies & Pro-	A12	Exploring Teachers' Perceptions In Managing Interracial Interactions in a Multicultural Setting: A Case Study of Five Malaysian Secondary Schools in Klang Valley	Azam Othman Ismail Sheikh Ahmad Hamidon Abdul Rahman Shahrul Fhaizal Shabhu	International Islamic University Malaysia		
		A16	Instilling Murabbic Values Amongst Teachers: Ikram-Musleh's Experi- ence	Noor Azlan Ahmad Zanzali Ust Megat Mo- hamed Amin	UTM, Malaysia		
	grammes	A18	المنهج والأيديولوجيا: انقلاب صلاح الدين الأيوبي الأبيض	Tahraoui Ramdane Merah Souad	International Islamic University Malaysia		
B: Royale 2	2:	B01	دراسة تحليلية لمستوى الكفايات التدريسية اللازمة لطالبات التربية العملية فى آلية التربية بالمجمعة فى ضوء معايير الجودة والإعتماد الأكاديمى	Amany Ibrahim Hamad bin Abdullah Alqumaizi	Majmaah Univ, Saudi Arabia		
	Accredita- tion and Certifica- tion of Teacher Qualifica- tion	B07	تصور مقترح لإعداد المعلم في ضوء معايير الاعتماد الأكاديمي	Fahd bin Abdul Aziz Aldakhil	جامعة الإمام محمد بن سعود الإسلامية - الرياض Saudi Arabia		
		B06	دور المشرف في تنمية الجودة عند المعلم: مدى اهتمام المشرفين والمعلمين بالجودة الشاملة في العملية التعليمية	Al-Shareef Salman Refan	International Islamic University Malaysia		
		B05	التوجهات المعاصرة في إعداد المعلم	Iman Aldoughan	King Faisal Univer- sity, Saudi Arabia		
C: Royale 4		C12	Evaluating Measurement Model Of Student Teacher Motivation	Ridwan Harris	International Islamic University Malaysia		
		C15	Guiding life skills teacher in con- ducting action research to improve interactive teaching strategies	Ziyana Al-Harthy	Ministry of Educa- tion, Oman		
	3: Teacher & teacher education	C09	Teacher's Worth: Wisdom Lessons from an Exemplary Teacher	Ibrahim Ali Allafiaji Siti Rafiah Abd Hamid Sharifah Sariah Syed Hasan Suhailah Hussein Nik Muhammad Saif	International Islamic University Malaysia		
		C06	Factorial Invariance of the Moral Character of Muslim Personality Scale: Confirmatory Factor Analysis of Yemeni and Malaysian Teachers	Fawziah K. Al-Ammar A.P. Ismail H. Ahmed Mohamed S. Nordin Khamsiah Binti Ismail	Ammran Universi- ty, Yemen		



Factorial Invariance of the Moral Character of Muslim Personality Scale: Confirmatory Factor Analysis of Yemeni and Malaysian Teachers

1. Fawziah K. Al-Ammar, Ph.D, Dept. of Educational Psychology, Ammran University, Yemen,

2. Ismail, H. Ahmed, Ph.D, AsssociateProfessor, Institute of Education, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia,

3. Mohamed S.Nordin, PhD, Professor, Institute of Education, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia,

4. KhamsiahBinti Ismail, PhD, Institute of Education, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia.

Conference Key areas: Teachers and teacher educators

Abstract:

The study aimed to assess evidence of measurement equivalence of Moral Character of Muslim Personality (MCMP). Teachers from two targeted countries were called to participate in answering two versions of the instrument: Arabic version of MCMP used to collect data from Yemeni teachers (n= 321) and Malaysian teachers (n= 213) answered on *Bahasa Malaysia* version of the MCMP. Precisely, the study examines the factorial invariance of the MCMP on the data collected from two samples. Multi-group of Confirmatory Factor Analysis (MGCFA) used to examine data across two samples, using the software of SEM technique (AMOS). The results of multi-group of CFA support of the configural invariance (second-order) for the hypothesized MCMP model. The results of CFA show that factors of cooperation, striving with wealth and self, reconciliation, voluntary charity, truth, fulfilling promises, forgivingness, and steadfastness are applicable in two cultures based on data collected from different samples. The goodness-of fit statistics related to constrained two-groups for the hypothesized model of MCMP show good- fit indices, supporting validity of the MCMP model across two samples. The value of Chi-square was insignificant (p=0.001), Root Mean Square Error Appreciation (RMSEA) = 04, Comparative Fit Index (CFI) = 0.92, and Incremental Fit Index (IFI) = 0.92. The findings suggest many implications for educators and policy makers in two countries to recruit teachers who exemplify good Islamic moral.



Introduction

Teachers have significant roles on the lives of young students, through building students' capacities in reading, mathematics, technology, and science in order to prepare them to be contributing members of society. Parent and general public expected from teachers to be morally upright individuals who display good moral character, because of the expectation from teachers to teach knowledge and science and display a character that reflect moral virtues such as fairness, honest, truth; to adhere professional codes of conducts (Lumpkin. 2008). In this regards, Lickona (1991, cited in Lumpkin, 2008) indicated that schools and teachers should educated for character, and play a role model of character via their interaction with students and display a behavior of moral virtues. A person of moral character has wisdom to know right and wrong, is honest, trustful, fair, respectful, and responsible, and commit to living according to these principles (Lumpkin, 2008). Moreover, teachers model integrity by choosing to do the right thing even when no one is looking. Lumpkin (2008) describes a person of integrity who consistently doing what is right, even when it would be easier to do something that is personally more beneficial. A person of integrity behaves in virtuous ways such as keeping promise and refraining from lying, cheating, and stealing. Lickona (1991) indicated that moral character is a universal phenomenon (Lumpkin, 2008), all over the world and religions generated shared values, shape peoples' behavior (Simither & Khoresandi, 2009) who possess the courage and conviction to live by moral virtue, and moral character.



1. An Islamic Moral Character

In this regards, Islamic teachings encourage Muslims to be honest, trustworthy and truthful, and to fulfil their promises, besides others moral character as influenced by faith factors. There are many Qura'anic verses determined the character of pious people (*muttaqin*). Pious people characterized by ten moral characters were emerged from the Qur'anic verses based on the analysis of the framework of *taqwa* (see Al-Amar, 2008). The model of *al-muttaqin* is an ideal model of Muslim personality. The relationship of a Muslim with others set a wonderful example of humanity based on love and seeking Allah. The willingness to spend property (*Infāq*) for the sake of Allah is a major characteristic of good believers.

By no means shall you attain Al-Birr (piety, righteousness, etc., paradise), unless you spend (in Allah's cause) of that which you love; and whatever of good you spend, Allah knows it well.

(Qur'an, *Al-'Imran*: 92)

Islam inspires helping others especially those who are in need. Muslim should offer their help and cooperate with others in doing good deeds to gain Allah's reward ($ta^c awn$). Spending money in for the sake of Allah ($Jih\bar{a}d$) and supporting others financially is a basic quality of Islamic ethics. This type of people has willingness to reconcile between two parties of believers who are in dispute (*sulh*). If those people do so, then they should fear Allah by justly carrying out the reconciliation, and the reward of their efforts will come from Allah's mercy as reported in the following verse:

"The believers are nothing but brothers (in Islamic religion). So, make reconciliation between your brothers' and fear Allah, that you may receive mercy."

(Qur'an, A-Hujurat: 10)

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

Islamic teachings emphasize the purification of the self. This is because without such purification, external good behaviour can only be hypocrisy. The Muslim is truthful (*sidq*), he does not cheat others and is sincere with all people in his saying and doing. A good Muslim behave is fair in dealing with others, always giving them their rights (*adl*). He fulfills his promises (*al-wfa' bill al- wa id*). A good Muslim should be a trustful person (*'amānah*), such as those portrayed in Sūrah al-Baqarah 2: 283. It is their faith (*taqwa*) that causes them to do whatever it takes to fulfill their responsibilities.

And if you are on a journey and cannot find a scribe, then let there be a pledge taken (mortgaging); then if one of you discharges his trust (faithfully), let him be afraid of Allah, his Lord. And conceal not evidence, for he who hides it, surely his heart is sinful. And Allah is All-Knower of what you do.

(Qur'an, *Al-Baqarah*: 283)

This fact was reported in a *hadith* narrated by Anas, who said that whenever the Prophet preached a sermon, he invariably repeated this sentence: "The man has no faith who cannot keep trust and the man who does not respect his promises has no religion" (Ahmed). Muslim is patience during hardship and when he faces difficulties. This is because he believes that he is examined by Allah and that he will obtain the reward from Him. The Model of al-*muttaqin* as Allah described them, being steadfast (*al-şabr fi al-ba'sā' wa al-dharā'*) and willing to forgive (*al-a'fw*), strengthens their faith, and purifies their hearts from illnesses. Such values strengthen one's capability to endure difficulties. They also guide believers in preventing themselves from doing harm to others when seeking Allah's way. A good Muslim never sustains any evil or enmity towards those who do hurt him, forgives those who oppress him, gives to those who do not give to him, and visits those who do not visit him. In fact, these are the traits of people who



seek Allah's way. In the following verses the Qur'an sets two criteria i.e patience and forgiveness for entering paradise:

Good deeds and evil deeds cannot be equal. Repel (evil) with one which is better (i.e., Allah exhorts believers to be patient in times of anger, and to excuse those who treat them badly). Then, verily, the enemy (will become) as though he was a close friend. But it (the above quality) is not granted except to those who are patient, and none is granted it except the owner of the great portion (of the happiness in the Hereafter and in this world of a high moral character).

(Qur.an, Fussilat: 34-35).

2. Challenges of cross-culture research

Cross- culture research is a scholarlydiscipline and increasingly has grown, to study racial-ethnic, religious, and other forms of cultural diversity either in cross-nationality or crosscultures. Majority of cultural projects involve tests and related assessment devices borrowed or copied instruments from techniques that widely available in countries where psychology is highly developed (Lonner, 1990). Cheung (2004) states that major instruments of personality such as the MinnesotaMultiphase Personality Inventory (MMPI), and the Eysenck Personality Questionnaire (EPQ) were translated and used in clinical setting. Also, others inventories related to several issues in psychology are borrowed and used in industrial and organizational settings in Hong Kong, India, Japan, Taiwan, and China. He reported that cross-cultural research on major personality measures in Asian countries faced several challenges related to research and applications of personality assessment (Cheung, 2004). This is almost the situation among many countries where psychology is not highly developed. The growth of borrowing Western tests to another culture cause significant problems. Sinha (1983, cited in Lonner, 1990, P.56) indicates that many concepts that are important in a borrowed country for tests may misscausing by IC ,

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

imperfect translation or adaptation of Western tests. In this regard, researchers addressed two an important issues: (1) the cross- languages differences that related to linguistics and translation effects on the measurement equivalence, (2) the effects of multinational surveys on measurement equivalence have been examined by researchers (Lui, Borg, & Spector, 2004; Ghorpade, Hattrup, Lackriz, 1999).

A greater attention should be paid to the quality of translations, cultural relevance, psychometric equivalence, and cross-cultural validity of instruments; and this causing by researchers who are often using an instrument that has been shown an adequate psychometric property in one culture, and they assume that the practical instrument that targets to measure the same psychological construct in another group in different culture (Milfont& Fischer, 2010; Cheung, 2004; Lonner, 1990; Church & Lonner, 1998). Moreover, Lui, Borg, and Spector (2004) stated that researchers in cross-cultural research commonly speaking about instability of the scale meanings caused by culture would make comparisons between cultural groups impossible. In general, cross-cultural adaptation of tests from different countries has been complicated by knowledge of words, verbal, and translations of vocabulary tests (Naglieri, Taddei, & Williams, 2012).

The sources of bias lower the test results and raise questions about the suitability of direct applications of instruments to different cultures (Cheung, Kwong, & Zhang, 2003; Cheung, 2004). Psychologists in cross-culture research have highlighted the fact that the construct bias is caused when the structure of the specific construct operates differently across different ethical



groups or cultures. Hambleton, Merenda, and Spielberger (2005) show that the sources of construct bias are a result of partial overlap in the definitions of the construct across different cultures. And poor sampling (i.e., short instrument) of the behaviours' elements which are not covered all aspects that are relevant to the domain of the specific constructs (Van de Vijver, 2003).

Multigroups of confirmatory factor analysis (MCFA) is one of the most common approaches to testing for measurement equivalence, within the framework of confirmatory factor analysis (CFA) model in the use of structural equation modelling (Comsa, 2010; Byrne, 2001). The invariant CFA validates the equivalence of the latent valuables of the measures via more than one test versions that are obtained from different populations by analyzing multiple groups simultaneously.

3. Measurement Equivalence (ME)

Measurement equivalence (or measurement invariance) has been an essential issue in psychological assessment and cross- culture research (e.g., Blievens & Anseel, 2004; Liu, Borg & Spector, 2004; Nguyen, Kitnertriolo, Evans, & Zonderman, 2004; *Wu Amery, Li Zhen, Zumbo Bruno, 2007*; Milfont & Fischer, 2010; Naglieri, Taddei & Williams, 2012; Olver, Neumann, Wong, Hare, 2012). Testing for equivalence allows for researchers to test different groups either across ethnic groups (e.g., Yen & Lan, 2013), nation (e.g., Nguyen, Kitner-triolo, Evans, & Zonderman, 2004) or cultures (e.g., Olver, Neumann, Wong, & Hare, 2012).

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

Raju, Laffitte, and Byrne (2002) states that when ME/MI presents, the relationship between the latent variable and observed variables in remains invariant as reflecting only the true difference between populations. Measurement invariance lets to compare scales, constructs, and surveys from different countries and cultures. Researchers reported that if ME is present, then confidently that:(a) the measured construct can be generalized to each cultural context; (b) respondents from different cultures interpret the meaning of the scale in a conceptually similar manner;(c) respondents from different cultures standardize the rating continuum in the same way; (d) the observed difference between groups means reflects the true differences between cultural groups on the constructs; (e) sources of cultural bias, translation, administration, errors, and administration errors are minimal. In contrast, without ME, the comparisons of the results made among countries would likely lead to incorrect conclusions (Alwin & Jackson, 1981; Drasgow, 1984, 1987, cited in Lui, Borg, and Spector, 2004).

4. Testing for Measurement Equivalence:

Any data collected based on psychological assessment may capitulate unreliable results due to measurement biases (Milfont, Fischer, 2011). And the biased findings may increase when the data collected using self-report from two different groups. Indeed, comparing groups in crosscultural research is more liability than others assessments. Scholars in cross- cultural literature have been highlighted methodological aspects important to cross-cultural psychological research. Structural and measurement equivalence issues are very common methodological concerns in cross-cultural research (Byrne, Oakland, Leong, van de Vijver, Hambleton, Cheung, & Bartram, 2009). They mentioned that when psychological construct use to measure in different cultural IC "

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

groups, it is difficult to assume that the meanings of the scores will be identical and that the scores subsequently can be compared across groups (Byrne et. al., 2009). Certainly, van de Vijver & Leung (1997) state that instrument must be test statistically for the equivalence across-cultural groups, and in order to interpret the scores in cross-cultural comparisons in the light of evidence that meaning of the dimensional structure of the measuring construct as well as the measuring instrument are group – equivalent. Measuring equivalence is present through acting invariance analyses that allow researchers to examine the variables of interest represent the same theoretical constructs across groups (Siedlecki, Manly, Brickman, Schupf, Tang, & Stem, 2010). Measurement equivalence (ME) provides that a degree of similarity in terms of psychometric or measurement properties of the survey instrument as observed in each group (Schaffer & Riordan, 2003). There are many types of equivalences or invariance analyses(Vandenberg & Lance, 2000), and Multi-group of Confirmatory Factor Analysis (MGCFA) is method often used to evaluate different groups, using Structural Equation Modeling (SEM). Measurement invariance is currently is increased among assessment literature, and there are different levels of invariance, which are explained below.

4.1. Configural Invariance:

It requires that the same indicators measure the same theoretical constructs under different groups (Yen & Lan, 2013; Byrne, 2001), explaining the participants from different groups conceptualize the constructs in the same way (Milfont & Fischer, 2010; Lee, 2009; Schmitt & Kuljanin, 2008). Configural invariance can be tested by running individual CFAs in each group, and necessary to run as Multi-group Confirmatory Factor Analysis (MGCFA) as well (Milfont &



Fischer, 2010). Hu & Blenter (1998) indicate that there are multiple fit indices that should be checked in order to evaluate the overall good-fit indexes for MGCFA. The Satora-Bentler chisquare (χ 2) statistic (Satora & Bentler, 1999) to account for non- normalityof the data, the normed –fit index (NFI: *Bentler, 1990*), the comparative fit index (CFI: Bentler, 1990), the No normed fit index or Tucker –Lewis index (NNFI:), and the root mean Square error of approximation (RMSEA: Browne & Cudeck, *1993*). RMEA scores range from 0 to 1.0, where 0.5 indicates a close fit, scores under 0.08 suggest an adequate fit, and values greater than 0.10 refers to a poor fir (Byrne, Shavelson, & Muthén, 1989). The values of NFI, CFI, NNFI also range from 0 to 1.0, where values close to 1.0 signify the best fit (Chen, souse, &West, 2005 Bryne, 2004).

4.2. Metric invariance:

Measurement equivalence is generally focuses solely on the invariant operation of the items and, in particular, on the factor loadings (Byrne, 2001). As such, interest centers on the extent to which the content of each item is being perceived and interpreted in exactly the same way across the samples. It is a precondition for using values as predictors of the factor loadings of each variable on each factor are the same across group (Yen &Lan, 2013; Schmitt &Kuljanin, 2008; Vandenberg & Lance, 2000), meaning that besides the configural invariance, the slope or factors loadings should be identical across all groups(Comsa, 2010). Then, when metric invariance is established, the values of factors loading can be comparable (Byrne, 2001). This level of equivalence is described as "strong invariance" whereas the configural invariance is sometimes called weak invariance (e.g., Siedlecki et. al., 2010; Schmitt & Kulijanin, 2008; Chen, Souse,



&West, 2005). However, standardized effects may not be compared, and to reach this level of invariance, it is necessary to guarantee that the variances of the values and their predictors are equal across groups (Yen &Lan, 2013). The fit indices that support metric invariance are presented when the factor loadings (λ) for all compared groups are constrained equal $\lambda g1 = \lambda g2 = \lambda g3$. Failure to reject the null hypothesis as supported by the small reduction of Comparative - Fit –index(Δ CFI) ≤0.01 (Cheung & Rensvold, 2002), when all variances and covariances in the model are freely estimated, as well as the indicators of good-ft indexes (e.g., $\chi 2$, p ≤.00, TLI, CFI, RMEA as indicated by statisticians in SEM techniques) (Hu & Blenter, 1998).

4.3. Structural invariance:

Testing for structural equivalence is centered on the unobserved (or latent) variables and this type of equivalence is of critical important to construct validity researchers whose interested in testing the dimensionality of a construct validity as defined by theory, valid across groups (e.g., Byrne & Wtkins, 2003). In testing for multigroup equivalence of a measuring instrument, interest can focus on both the factor variances and factors' covariance (Byrne, 2001). In practical, there are three types of structural invariance tests, which are factor invariance, factor covariance, and factor mean invariance (Milfont & Ficher, 2011). However, the latter are typically of most interest and therefore constitute the focus in most tests for instrument equivalence. For this study will test factor variance invariant, when the range of scores do not be vary across targeted groups, and the variance of all factors should be constrained across all groups (Byrne, Shavelson, & Muthén, 1989). Rejecting the hypothesis that constructs variances is invariance across groups and is invoked by placing constraints between construct variances across groups (Yen & Lan,



2013; Johnson & Mead, 2007). Moreover, good model-data fit indices and a small CFI decrease ($\Delta CFI \leq 0.01$) compared to the metric model implies the existence of variance invariance (Cheung & Rensvold, 2002).

5. Aim of the study

The purpose of the paper is to give empirical evidence that shows the construct equivalence of the Moral Character of Muslim Personality (MCMP) by examining configural invariance, metric invariance, and factorial structural invariant across two cultural groups in Yemen and Malaysia.

6. Method

6.1. Participant

The ample for the study includes participants from Yemen and Malaysia. The Yemeni sample consist of 312 teachers from 22 secondary schools; 198 female and 114 male, with a mean age of 32.39 (SD =4.8). Some of them (202) were married, while 98 were single. The sample also included six divorced and three widows. The majority of respondents (279) had bachelor's degrees, 10 persons held diploma, and 23 had a master's degree. The data was collected from respondents who have teaching experience: 41% had more than 10 years of teaching experience, 31% of them had between more than 5 years and less than 10 years of teaching experience, and 28% had less than 5 years of teaching experience. The respondents taught different subject matters. The response rate of the collecting data was very high (91.3%). While, Malaysian samples is 213 teachers, majority of them who were receiving training courses in the Institute of Education (INSTED), IIUM at Malaysia, 20% were male and 79.3% were female. The mean age

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

of Malaysian participants was 30.5 (SD= 6.0). The majority of respondents were Malay (84%), 12.2% were Indians while the percentage of Chinese was 3.3%. About 75% held bachelor's degrees, 14.6 % held diploma, while a small percentage (8.5%) held master or doctorate. The majority of Malaysian participants were single (59.6%) and the percentage of married was 38 %, while a very small number were divorced or widows. The background of the two samples summarized in the Table 1.

Table 1: Summary of Background Characteristics of Respondents in Both Samples

-	N	Age		Geno	Educational Level			evel	Marital status			
		М	SD	М	F	D	В	Μ	Ph	Married	Single	others
									D			
Yemeni	312	32.39	4.8	114	198	10	279	21	2	202	98	12
sample												
Malaysian sample	213	30.5	6	44	169	37	158	14	4	81	127	5

Note: M: mean and SD: standard deviation.

D: means diploma, B: means bachelor, M: master, PhD: means doctorate.

6.2. Instrument

Moral Character of Muslim Personality(MCMP) is a multidimensional measure, consists of 28item. It was developed by the first author, to assess the theoretical background of Islamic moral as stated in the Qura'anic context (see Al-Ammar, 2008). It builds up on 7- point Likert type scale as follows: (1) never, (2) rarely, (3) slightly, (4) occasionally, (5) usually, (6) often, and (7) always. The MCMP has been validated based on data collected from Yemeni secondary school

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

teachers in Sana'a (Al-Ammar, Ahmed, & Nordin, 2012), and the indicators of good-fit indices (Chi-square= 493.478, p= .00, CMINDF= 1.44, CFI= .951, IFI=. 95, and RMEA=. 04) show good fit, supporting the eight dimensions of the second-order of MCMP (cooperation, striving with wealth and self, reconciliation, voluntary charity, truth, fulfilling promises, forgivingness, and steadfastness). Also, the structural validity of MCMP instrument was advanced in multiculture context (e.g., International Islamic University Malaysia, IIUM's students), and the data collected from heterogeneous sample within the main campus of IIUM, Gombak (Al-Ammar & Ahamed, in press). The results advocate the eight factors of second-order MCMP model, using Confirmatory of Factor Analysis. But, the findings of later study show approximately the same results expect the dimension of *truth* (3 items) was disappeared, and instead of the dimension of *trust* (3 items) is emerged. Thus, the differences in the structural validity of MCMP between two studies suggest that to examine the invariance of structural validity of MCMP based on data collected from two different cultural contexts (e.g., languages) such as Yemeni and Malaysian schools.

6.3. Translation Procedures

According to the recommendation of Berry (1996 cited in He & van de Vijver, 2012), which advises researchers to explore the structure of the construct and adequacy of the samples items, in order to maintain the same conceptualizing meaning and construct equivalence. Simultaneously, the the origin version of MCMP instrument was written in both languages : English and Arabic language (see He & van de Vijver, 2012; see Al-Ammar, 2008). Then, multiple processes were implemented to avoid poor items translation (van de Vijver, 2003; van IC "

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

de Vijver & Poortingga, 2005), adopting the suggestions of Brislin, Lonner, and Thorndike (1973) as follows:

- The MCMP was given to two bilingual individuals, who translated the instrument from English to Malay language. The two translations were done by the two bilinguals separately. The third author compared the translation from English to Malay language and chose the better appropriate translation.
- 2. The back translation of MCMP from Malay language (*Bahasa Melayu*) to English to was achieved by another bilingual person who is familiar with the psychological constructs of MCMP. Then, the first author compared the back translation of MCMP (from Malay language to English) with the original version of MCMP. Accordingly, the majority of the items maintain their aims but were worded differently. Meanwhile, the translation of only a three items was different from the original version. A discussion was conducted with two bilingual individuals to maintain the goals of these items in Malay language.
- 3. Both the original version of MCMP and the *BahasaMelayu* version were given to a fourth bilingual individual who is familiar with the content of items in order to investigate the appropriateness of each item in the *Bahasa Melayu* version and to reflect a clear meaning of English version. Indeed, this checking included whether the translations of demographic variables and instructions of the test are appropriate from English to Malay language. A few changes were made by the bilingual individual in terms of demographic data and test instructions.
- 4. The final version of test was written in two languages (Malay and English) including demographic data and the instructions of the test. In this point, a final version of test was

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

checked by the third author and a few changes were made, and in turn the MCMP was used to collect data from the Malaysian sample.

7. Data Analysis:

The factorial invariance of MCMP was assessed through Multi Group Confirmatory Factor Analysis (MGCFA), using AMOS 16.0.1(Arbuckle, 1983-2007). In order to evaluate the adequacy of factorial invariance of MCMP across two samples (Yemeni and Malaysian teachers), series of CFAs with the maximum likelihood estimation was used to conduct CFA for each group separately. The factorial invariance or equivalence (MI/E) was evaluated using MGCFA procedures (see Byrne, 2001; Naglieri, Taddei, & Williams, 2012; Kline, 2005; Liu, Borg, & Spector, 2004; Nguyen, Kitner- Triolo, Evans, & Zonderman, 2004; Ghorpade, Hattrup, & Lackriz, 1999). The M/E technique includes establishing cross-group constraints and evaluating the models with increasingly restricted parameters. Precisely, an unconstrained model is first tested to establish configural invariance. Subsequent tests constrain various factors to be equal , including measurement weights, measurement intercepts, and measurement residuals to examine metric invariance, strong invariance, and strict invariance, respectively (Naglieri, Taddei, & Williams, 2012).

There are several well-known goodness-of-fit indices were used to evaluate model fit: the chi-square $\chi 2$, the comparative fit index (CFI), the incremental index of fit (IFI), the root mean square error of approximation (RMSEA) and the goodness-of-fit indices provide "rules of thumb" for the recommended cutoff values to evaluate data-model fit. Statisticians of SEM



analysis recommend that researchers to use combinations of good-fit- indices to obtain a robust evaluation of model fit. The criterion values they list for a model with good fit are CFI > 0.95, IFI> 0.95, RMSEA < 0.06 (e.g., Hu & Bentler, 1995; Bollen, 1989; Schumacker & Lomax, 1996; Schermelleh-Engel, Moosbrugger, & Müler, 2003) for assessing fit in structural equation modeling. Hu and Bentler offer cautions about the use of good-fit indices. In terms of for RMSEA index, researchers reported that the value of .05 or less indicates the best fit, while the value up to .08 represents reasonable fit , and the indexes CFI, TLI, IFI values vary between 0 and 1.00, and the value near from 0 indicates poor fit while, the value near 1.00 indicates a good fit (e.g., Liu, Borg, & Spector, 2004; Schumacker & Lomax, 1996; Hu & Bentler, 1995).

Coovert and Craiger (2000 cites in Liu, Borg, & Spector, 2004) indicates that the root mean square approximation(RMSEA, Steiger & Lind, 1980), and comparative fit index (CFI, Bolllen, 1989) were considered the two most important from the list of fit- indices. Changes in goodness-of-fit statistics have been examined to detect differences in structure parameters. In terms of the statistical versus practical significance, Cheung and Rensvald (2002) recommended researchers to consult the Bentler's (1990) comparative fit index (CFI) to identify items lacking measurement invariance, the changes in the CFI (Δ CFI < 0.1) between the constrained and freely estimated models (Wesi, Tan, Brower, Önder, 2007; Byrne, 2001). Additionally, Chi-square tatistic considers good fit (Hair et al., 2006), and the changes between the values of chi-square ($\Delta\chi 2$) between the constrained and freely estimated models should be statistically insignificant (p > 0.05)

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)

8. Results:

The findings of CFA showed that the second-order of MCMP model indicate a goddess- fitindexes for Yemeni sample as well as Malaysian sample as displayed in Table2. For Yemeni sample χ^2 (342) = 493.5, p=0.00, CFI=0.95, IFI= 0.95, RMSEA= 0.04, meanwhile Malaysian sample demonstrated reasonable fit indexes χ^2 (342) = 512.2, p=0.00, CFI=0.92, IFI= 0.92, RMSEA= 0.06. Given the adequate fit's findings of factorial structure for second-order MCMP model, we proposed to examine series of models to establish measurement invariance and comparing with configural model of MCMP simultaneously as recommended by Cheung and Rensvold (2001) and Byrne (2001).

8.1. Results of Configural Invariance

In testing for this form of invariance, an unrestricted baseline model was specified in which each group have the same structure (eight sub-constructs), and the results of configural model (baseline model)generated a set of good-fit –indexes for MCMP as displayed in table 2: (χ 2 (684) = 1085.95, p=0.00, CFI=0.94, IFI= 0.94, RMSEA= 0.03), and we turn to test the invariance of factorial measurement and structure across two groups.

8.2. Results of Metric Invariance test

The factors loadings were constraint equal across two groups (Yemeni and Malaysian). Overall, the MCMP indices produce an adequate fitting model, with an important criterion to claim that invariance was holds, the minimum differences of $\Delta CFI=(-0.01)$, Besides, goodness of fit indices such as the change in $\Delta \chi 2$ was statistically insignificant and the values of fit



(CFI=0.92, IFI= 0.92, RMSEA= 0.04) were reasonable, supporting invariance measurement for

MCMP across two groups as shown in Table 2.

Model	χ2	df	Δχ2	∆df	CMIND/DF	CFI	∆CFI	RMSEA	IFI
Yemeni sample	493.478	342			1.44	0.95		0.04	0.95
Malaysian	592.243	342	1		1.73	0.92		0.06	0.93
sample									
Blooded model	605.012	342			1.76	0.95		0.04	0.95
Configural	1085.95	684			1.59	0.94	••••	0.034	0.94
model(baseline model)			-		-				
Equal factor loadings	1202.56	711.	616.61	27	1.69	0.92	-0.01	0.036	0.92
as 2 nd order									
(constrained model)									
Fully invariant	1202.56	711	616.61	27	1.69	.923	-0.01	0.036	0.924
(constrained errors) for									
the structural									
invariance as 2 nd order							1. N		
model)									

Table 2: Good-Fit- Indicators for Testing Measurement Invariance across Yemeni and Malaysian Groups

8.3. Results of Structural invariance test

In the context of structural Equation modeling (SEM), testing for the invariance on particular parameters (two regression paths) for the second-order MCMP model across Yemeni teachers and Malaysian teachers(see Figure 1. 2). The results of constraining factor loadings, and



unobserved (or latent) variables confirmed the equality of the second-order MCMP model across two samples. Following the criterion of the change in Δ CFI \leq 0.01and the $\Delta\chi^2$ (711)=616.6, p<0.05 was statistically insignificant which were within the criteria recommended by Cheung and Rensvold (2001) and Byrne (2001) to suggest invariance hold, besides fit indices of (CFI=0.92, IFI= 0.92, RMSEA= 0.04), supporting good fit indices.

9. DISCUSSION

This article exemplified the method of testing measurement invariance in a second-order model for the scale of Moral Character of Muslim personality. The second-order factor model hypothesized that the responses to the measurement of Moral character could be explained by eight first-order factors which are cooperation, striving with wealth and self, reconciliation, voluntary charity, truth, fulfilling promises, forgivingness, and steadfastness. Moreover, there is one second-order factor (moral character of Muslim personality) that underlies the eight firstorder factors. Compared to a correlated eight-factor model, the second- order factor model is more parsimonious and provides theoretically error-free estimates of both the general factor and each specific factor. This latter advantage is particularly important when researchers are interested in understanding whether the specific factors can predict external criteria over and above the general factor.

Measurement invariance provides strong evidence that the same construct has been measured across two groups, and it is an important issue in comparing results across groups. Our illustration tested measurement invariance across groups for the second-order factor model at



different levels that are tested in sequence: configural invariance, factor loadings of the firstsecond order factors, and structure invariance were supported. Configural invariance requires the same indicators that measure the same theoretical constructs in different groups. Configural invariance does not guarantee that the relationships between indicators and factors are the same across two samples. Metric invariance which is of a higher level of invariance or "strong factorial invariance" (Cheung and Rensvold,20001;Vandenberg & Lance, 2000) is required to establish this relationship. If metric invariance test holds, it guarantees that the factor loadings between factors and indicators are similar across groups. In General, the findings of measuring invariance for MCMP model sustain the power of Islamic theory in morality to be empirically usable in selecting teachers within Muslim countries.

10. Limitations and conclusion

Indeed, the invariance of the differences in means for variance and covariance of the measured variables was not achieved with the data reported here. It is necessary to be test using different samples from different populations.

Despite the evidence of establishing measurement invariance for MCMP across two groups in Yemen and Malaysia, it is recommended that conducting longitudinal studies as well as cross –sectional studies to be carried out in order to test and establish the measurement invariance stability for the scale of MCMP. For further research, it is suggested that the study be replicated using MCMP across gender, age, and different groups from different cultures and nations.

International Conference on Teacher Education in the Muslim World 2013 Institute of Education (INSTED), International Islamic University Malaysia (IIUM)



Figure1 displayed the factorial measurement and structure invariant for second –order MCMP model across Yemeni and Malaysian teachers after constraining factor loadings. The hypothesized model of MCM (t2-t50,e1-e36 represent error variances, single headed arrows from factors depict factor loading, coper_m means cooperation, stri_m means *jihād*, rec_m means reconciliation, cha_m means voluntary charity, tru_m means truth, ful_m means fulfilling a promise, for_m means forgiveness, and sted_m means steadfastness)



References

- Al-Amar, Fawzia K., & Mohammad S. Nordin (2008, March). Moral character of an intellectual Muslim: Scale Validation. Paper presented at international seminar on higher education in the Muslim world: challenges and prospects at 24th-25the March 2008, International Institute for Muslim Unity, IIUM, Kuala Lumpur.
- Al-Amar, Fawzia Kassim (2008). Intellectual characteristics of Muslim personality and job characteristics as determinants for organizational citizenship behavior among Yemeni secondary school in Sana'a. Unpublished doctoral dissertation, International Islamic University Malaysia, Kuala Lumpur.
- Byrne, M. B., Oakland, T., Fredrick, T. L., ve de Vijver, J.F., Hambleton, R. K., Cheung, M. F., Barrtram, D. (2009). A critical analysis of cross- culturalresearch and testing practices: implications for improved education and training in Psychology. *Training and Education* in Professional Psychology, 2 (2), 94-105.
- Brtne, B. M. (2004). Structural Equation Modeling With AMOS, EQS, and LISREL: Comparative Approaches to Testing for the Factorial Validity of a Measuring Instrument. International Journal of Testing, 1(1), 55–86.
- Byrne, B. M. & Watkins, D. (2003). The issue of measurement equivalence revisited. *Journal of Cross-cultural Psychology*, 34, 155-175.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement equivalence. *Psychological Bulletin*, 105, 456-466.
- Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). Cross-cultural research methods, New York: Wiley.
- Bielet, J. (2003). Cross-cultural equivalence with structural equation modelling. In Harkness, J.
 A., Van de Vijver, R. J. F., & Hohler, P. PH (Eds.). Cross- cultural survey methods. New Jersey: John Wiley & Sons, Inc.
- Braun, M. (2003). Errors in comparative survey research: an overview. In Harkness, J. A., Van de Vijver, R. J. F., & Hohler, P. PH (Eds.). Cross- cultural survey methods. New Jersey: John Wiley & Sons, Inc
- Bentler, P.M. (1990). Comparative Fit Index in Structural Models. *Psychology Bulletin*, 107 (2), 238-246.
- Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.



- Cheung, F. M., Kwong, J. Y., & Zhang, J. (2003). Clinical validation of the Chinese personality assessment inventory. *Psychological assessment*, 15(1), 89-100.
- Cheung, M. F. (2004). Use of western and indigenously developed personality Tests in Asia. *Applied Psychology: An International Review*, 53(2), 173-191.
- Cheung, G. W, & Rensvold, R. B. (2001). Evaluating Goodness -of -fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 2, 233-255.
- Comsa, M.(2010). How to compare means of latent variables across countries and waves: testing for invariance measurement: an application using Eastern European societies. *Sociology*, 24, 6, 636- 669.
- Lonner, J. W.(1990). An over view of cross-cultural testing and assessment. In R.W. Brislin (Ed.), *Applied Cross- Cultural Psychology*(pp. 56-76). London Sage Publications Ltd.
- Lievens, F. & Anseel, F. (2004). Confirmatory factor analysis and invariance of an organizational citizenship behavior measure across samples in a Dutch-speaking context. Journal of Occupational and Organizational Psychology, 27, 77, 299-306.
- Liu, C. Borg, I., Spector, E. P. (2004). Measurement equivalence of the German job satisfaction survey used in a multinational organization: implications of Schwartz's cultural model. Journal of Applied Psychology, 89, 6, 1070-1082.
- Lucas, N. Mongkuo, M. Y., Taylor, A., & Ngwainmbi, J. (2012). HIV prevention behavior scale: testing for factorial invariance across lower class and upper class historically black college students. *IOSR Journal of Pharmacy (IOSRPHR), 2, 4, 58-65.*
- Lumpkin. A. (2008). Teachers as roles models teaching character and moral virtues. JOPERD, 27, 2, 45-49.
- Ghorpade, J., Hattrup,K. Lackriz, J. R. (1999). The of personality measures in cross-cultural research: A test of three personality scales in two countries. *Journal of Academic Psychology*, 84, 5, 670-679.
- He, J., & van de Vijver, F. (2012). Bias and Equivalence in Cross-Cultural Research. Online Reading in Psychology and Culture, 2(2).
- Hambleton, R. K., Merenda, P.F. & Spielberger, C. D. (Eds). Adapting educational and psychological tests for cross-cultural assessment. London: Lawrence Erlbaum Associates, Publishers.



- Hu, Li-Tze, & Benter, M. P. (1995). Evaluating Model fit. In R. H. Hoyle, (Ed.), *Structural Equation Modeling Concepts, Issues, and Applications*. Thousand Oaks: Sage Publications.
- Hair, J. F. Jr., Andrson, R. E., Tatham, R. E., & Black, W. F., (1998). *Multivariaat data analysis* (5 th ed.). New Jersey: Prentice- Hall International, Inc.
- Johnson, E. C., & Meade, A. W. (2007, April). The Role of Referent Indicators in Tests of Measurement Invariance. Paper presented at the 22nd Annual Meeting of the Society for Industrial and Organizational Psychology, New York.
- Olver, E. M., Neumam, S. C., Wong, C. S., Hare, D. R. (2012). The structural and predictive properties checklist-revised in Canadian Aboriginal and ono- Aboriginal offenders. Psychological Assessment, doi: 10.1037/a0029840.
- Milfont, L. T., Fischer, R. (2010). Testing measurement invariance across groups: applications in cross-cultureal research. International Journal of Psychological Research, 3, 1, 11-121.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2006, August). *The Utility of Alternative Fit Indices in Tests of Measurement Invariance*. Paper presented at the annual Academy of Management conference, Atlanta, GA.
- Meade, A. W. (2005, April). Sample Size and Tests of Measurement Invariance. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles, CA.
- Mead, A. W. & Bauer, D. J. (2007). Power and precision of confirmatory of factor analysis of measurement invariance. *Structural Equation Modeling*, 14(4), 611–635.
- Nguyen, T. H. Kitner-Triolo, M., Evans, K. M., Zonderman, B. A. (2004). Factorial invariance of the CES-D in low socioeconomic status African Americans compared with a nationally representative sample. Psychiatry Research 126, 177-187.
- Naglieri, A. J., Taddei, S., Williams, M. K. (2004). Multigroup confirmatory factor analysis of U. S. and Italian children's performance on the PASS theory of Intelligence as measured by the cognitive assessment system. Psychology Assessment, Doi:
- Raju, S. N., Laffitte, J. L. & Byrne, M. B.(2012). Measurement equivalence: a comparison of Methods baed on confirmatory factor analysis and item response theory. *Journal of Applied Psychology*, 87, 3, 517-529.
- Siedlecki, L. K., Manly, J. J., Brickman, A. M., Schupf, N., Tang, Ming-Xin, Stem, Y., (2010). Do neuropsychological tests have the same meaning in Spanish speakers as they do in



English speakers?. Neuropsychology, 24, 3, 204-211.

- Schmitt, N. &Kuljanin, G. (2008). Measurement invariance: a review of practice and implications. *Human Resource Management*, 18, 210-22.
- Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. New Jersey: Lawrence Erlbaum Association, Inc.
- Schermelleh-Engel, K., Moosbrugger, H., & Müler, H. (2003). Evaluating the fit ofstructural equation models: tests of significance and descriptive goodness- of –fitmeasures. *Methods* of Psychology Research Online 2003, 18(2), 23-74. Retrieved September 5, 2007. <u>http://</u> www.mprr-online.de.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods, 3 (1): 4-69.
- Van de Vijver, F. J., R., (2003). Bias and substantive analyses. In Harkness, J. A., Van de Vijver, R. J. F., & Hohler, P. PH (Eds). Cross- cultural survey methods. New Jersey: John Wiley & Sons, Inc
- Yen, W. K. S. & Lan, O. S. (2013). Measurement Invariance of students' parent action questionnaire (SPAQ) among three ethics groups in Malaysia. *International Journal of Educational and psychological Assessment*, 2, 12, 123-139.
- Pellegrine, E. K. &Scandure, A. T. (2005). Construct equivalence across groups: unexplored issue in mentoring research. *Educational and Psychological Measurement*, 65, 2, 323-335.
- Wang, S. (2006). Validation and invariance of factor structure of the ECPE and MELAB across gender. Spaan Fellow Working Papers in Second or Foreign Language Assessment, 4, 41-56. Retrieval from http//: English Language Institute, University of Michigan 41