

Extending self-leadership research to the East: Measurement equivalence of the Chinese and English versions of the MSLQ

Jessie Ho,¹ Paul L. Nesbit,¹ Denise Jepsen² and Soseh Demirian³

¹Macquarie Graduate School of Management, ²Department of Business and ³Department of Psychology, Macquarie University, North Ryde, Sydney, New South Wales, Australia

In response to calls regarding the applicability of self-leadership measurement in Eastern (collectivistic) and Western (individualistic) cultures, this study examined the psychometric properties and measurement invariance of the modified self-leadership questionnaire (MSLQ). The sample consisted of 395 Chinese students and 241 Australian students. Results revealed that the modified self-leadership questionnaire exhibited a satisfactory condition of psychometric properties across cultures. A series of multi-sample confirmatory factor analyses demonstrated the cross-cultural similarity of an 11-factor model across the Chinese and Australian samples. The modified self-leadership questionnaire was also found to possess measurement invariance, suggesting that it is appropriate for cross-cultural research assessing differences of self-leadership behaviour across Chinese and Western cultures. Implications for future research are also discussed.

Key words: collectivistic culture, individualistic culture, measurement equivalence, modified self-leadership questionnaire (MSLQ), psychometric properties, self-leadership.

In order to succeed in increasingly complex and turbulent markets, organizations often adopt decentralized, organic-type organizational structures (Conger & Kanungo, 1988). Within these organizations, members at all levels are being encouraged to take greater responsibility for their own job tasks and work behaviours. This trend toward more flexible and decentralized organizations and greater employee autonomy has drawn attention to the concept of self-leadership (Manz, 1986; Manz & Neck, 2004; Manz & Sims, 1980, 2001). Self-leadership is defined as a systematic set of strategies through which individuals influence themselves toward higher levels of performance and effectiveness. (Manz, 1986; Manz & Neck, 2004). Self-leadership incorporates the principles of self-regulation (e.g. Carver & Scheier, 1998) and self-management (e.g. Luthans & Davis, 1979; Manz & Sims, 1980), but goes beyond these to include self-determined governing standards and a unique set of self-motivation techniques (Manz, 1986; Neck & Houghton, 2006).

Self-leadership strategies are typically classified into three categories, namely behaviour-focused strategies, natural reward strategies, and cognitive or thought pattern strategies (Manz & Neck, 2004; Prussia, Anderson, & Manz, 1998; Sims & Manz, 1996). Behaviour-focused self-leadership involves using action-oriented strategies to

accomplish tasks that are difficult or are neither enjoyable nor motivating. Sims and Manz (1996) identified various behaviour-focused self-leadership strategies, including self-observation, self-goal setting, self-reward, self-punishment, and self-cueing strategies. Natural reward strategies are designed to enhance intrinsic motivation for better performances which include focusing one's attention on the pleasant aspects of a given job and engaging in job- or task-redesign (Manz & Neck, 2004). Constructive thought strategies involve visualizing successful performance, engaging in positive self-talk, and examining individual beliefs and assumptions to align cognitions with desired behaviour (Neck & Manz, 1996; Neck, Stewart, & Manz, 1995). Research on the use of self-leadership strategies has found it to be effective at enhancing performance in clinical, athletic, and educational settings (Neck & Manz, 1992), and in employment contexts (Neck & Manz, 1996; Stewart, Carson, & Cardy, 1996).

Establishing measurement invariance in self-leadership assessment

As is often the case with organizational theories, the majority of self-leadership research has been conducted in the USA and as such its development and supporting research may reflect Western cultural values (Alves *et al.*, 2006; Neck & Houghton, 2006; Neubert & Wu, 2006). Consequently, there have been calls for more empirical research to examine the intercultural aspects of self-leadership (Alves *et al.*, 2006; Neck & Houghton, 2006). Developing an understanding of the applicability of self-leadership

Correspondence: Jessie Ho, Macquarie Graduate School of Management, Macquarie University, Balaclava Road, North Ryde, Sydney, NSW 2109, Australia. Email: ccjessie@hkccepolyu.edu.hk

Received 15 September 2010; accepted 15 September 2011.

across different cultures may help to improve organizational and managerial effectiveness in multinational corporations. However, research on cross-cultural comparison of self-leadership is rare due to the lack of a measurement instrument applicable in both Eastern and Western cultures. Thus, developing a reliable scale with measurement invariance is a critical step which ensures that any observed differences truly reflect differences between cultures on the underlying construct, and are not due to systematic biases in the way people from different cultures interpret a given measure in a conceptually different manner (Vandenberg & Lance, 2000). As such, the purpose of this study is to examine the psychometric properties and the extent of measurement invariance of one available self-leadership instrument, the modified self-leadership questionnaire (MSLQ; Ho & Nesbit, 2009) using samples drawn from Western and Eastern cultures.

Refinement and extension of the self-leadership scale

The first published self-leadership assessment instrument, Anderson and Prussia's (1997) 50-item Self-Leadership Questionnaire (SLQ), was based to a large extent on the self-leadership research of Manz and Sims (1991) and Manz (1992). Nevertheless, Anderson and Prussia's SLQ suffered from a number of psychometric problems and required further refinement. Subsequently, a revised Self-Leadership Questionnaire (RSLQ) was developed by Houghton and Neck (2002) based on Anderson and Prussia's (1997) scale. They eliminated items that loaded on the unanticipated factors and reworded a few items. Houghton and Neck (2002) also reported that the RSLQ demonstrated a greater degree of reliability and construct validity than the earlier SLQ.

As noted earlier, as much of the existing research has been done in the USA, Alves *et al.* (2006) called for more empirical research to address the applicability of self-leadership theory across cultures. The authors argued that while self-leadership behaviour is a generally universal concept, different cultures value different attributes and practices. Thus, self-leadership behaviours may be applied differently across cultures. However, only two empirical studies have directly addressed this issue by exploring the application of self-leadership measurement in a non-Western context. Neubert and Wu (2006) tested Houghton and Neck's (2002) RSLQ in a Chinese context. They found that four out of nine self-leadership factors had unsatisfactory reliability and failed to emerge in the same manner as factors found in Houghton and Neck (2002)'s original validation study conducted in the U.S. context. This suggested that although RSLQ was found to be a valid scale with promising reliability and construct validity in U.S. samples

(Houghton & Neck, 2002), its generalizability to the Chinese context was problematic. Furthermore, using the five-factor model of self-leadership (excluding those four factors with unsatisfactory reliability) that included goal setting, visualizing successful performance, self-talk, self-reward, and self-punishment, Neubert and Wu (2006) found that self-leadership was positively related to self-reports of in-role performance and creativity.

To further enhance the generalization of the self-leadership concept to the Chinese context, Ho and Nesbit (2009) modified the RSLQ. Following the work by Neubert and Wu (2006), they refined or deleted some ambiguous items of four existing dimensions (self-observation, evaluations of beliefs and assumptions, natural rewards, and self-punishment) found to have low reliabilities in Neubert and Wu's (2006) validation study. They also suggested that three component dimensions, namely natural rewards, self-observation, and evaluating beliefs and assumptions, may be more valid for individualist cultures but less so for collectivistic cultures. According to Markus and Kitayama (1991), the cultures of North America and Northern and Western Europe have been identified as generally individualistic that possess independent self-construal. Individualists value independence, autonomy and the expression of one's unique configuration of needs, rights, and capacities. In contrast, the cultures of Easterners, such as Japanese, Chinese, and Koreans, have been identified as collectivistic, possessing interdependent self-construal (Markus & Kitayama, 1991). Collectivists strive to find a way to fit in with significant others, and to fulfil their obligations as part of social networks. Thus, in considering the nature of self-leadership in non-Western cultures, such as Chinese cultures, Ho and Nesbit (2009) suggested that more attention needs to be given to the issue of relationships and collective efforts in the measurement of relevant behaviours.

Ho and Nesbit (2009) argued that the original items of these three component dimensions were strongly shaped by individualistic cultural values with a strong focus centred either on the 'task' or the 'person'. Natural rewards emphasizes the personal intrinsic value of a task, self-observation highlights the role of observing one's actions relative to goals, and evaluating beliefs and assumptions concerns examining one's thoughts, especially self-defeating thoughts that detract from successful task performance. Thus, these three subscales were extended to the development of three new factors, 'relation-based natural rewards', 'social-oriented evaluation of beliefs and assumptions' and 'relation-based self-observation', which incorporated social/relation-based features typically associated with collectivism (Markus & Kitayama, 1991). Relation-based natural rewards measures a person's tendency to find enjoyment through maintaining social relations, group harmony and belongingness. Social-oriented evaluation of beliefs and assumptions measures a person's tendency to identify

their own beliefs and assumptions in conflict with the opinions of their group members and authority figures. Relation-based self-observation measures a person's tendency to keep track of their performance in meeting the expectation of authority figures or group members.

Even though there are 12 subscales in the MSLQ, the exploratory and confirmatory factor analysis done by Ho and Nesbit (2009) consistently identified an 11-factor structure rather than the 12-factor structure with good internal consistency. The items of the 'relation-based self-observation' consistently merged with those of the 'task-based self-observation' (from the original subscale) to form a single factor. This suggests that when Chinese participants observe progress on the task or in reaching goals they also assess whether they are working smoothly and harmoniously with team members. Overall, the study by Ho and Nesbit (2009) provided evidence that the 38-item MSLQ is a reliable measure with a stable 11-factor structure that captures different aspects of self-leadership theory. Furthermore, Ho and Nesbit (2009) also found a strong relationship between the self-leadership scale and general self-efficacy in the Chinese context, which provided further evidence for the construct validity of the MSLQ.

Despite the promising evidence supporting the construct validity and reliability of the MSLQ, whether it is equivalent across Chinese and Western cultures is unknown. Thus this study investigates the measurement invariance of the MSLQ to determine: (1) whether the three relation-based dimensions (relation-based natural reward, social-oriented evaluation of beliefs and assumptions, and relation-based self-observation) can be generalized to the Western culture; (2) whether the empirical-supported 11-factor model can be replicated in another Chinese sample; and (3) whether the original theorized 12-factor rather than the 11-factor model is more appropriate for the Western, individualistic population.

Social/relation-based self-leadership dimensions and western culture

Despite the individualistic culture typically associated with Western societies (Hofstede, 1980; Markus & Kitayama, 1991), it can be argued that these three social/relation-based dimensions of self-leadership may be applicable to Western cultures on the basis that the need for 'relatedness' is regarded as a universal human value (Alderfer, 1972). For example, self-determination theory (Deci & Ryan, 1985, 2000) views people as 'active, growth-oriented organisms who are naturally inclined toward integration of their psychic elements into a unified sense of self and integration of themselves into larger social structures' (Deci & Ryan, 2000, p. 229). Self-determination theory further stresses that the satisfaction of psychological needs of relat-

edness is essential for the healthy development and well-being of all individuals regardless of culture (Deci & Ryan, 2000).

The relevance of social relations values among Western cultures is also reflected in the increasing use of teams within Western organizations (Devine, Clayton, Phillips, Dunford, & Melner, 1999). Effective teamwork has become a crucial component of an organization's success (Kozlowski & Ilgen, 2006; Smolek, Hoffman, & Moran, 1999) and this effectiveness is to a certain extent dependent on the social and interpersonal abilities of team members (Stevens & Campion, 1999).

In sum, the above research points to the social foundation of team effectiveness which suggests the applicability of relatedness/social constructs within Western culture. While the social/relation-based refinement of self-leadership dimensions was originally conceived with collectivist cultures in mind (Ho & Nesbit, 2009) it is expected that the social/relation-based concept is applicable to Western populations as well. Given these arguments, we next outline a study that seeks to assess the cross-cultural comparability of the MSLQ. As is required, ethics approval for this study was given by the Ethics Review Committee of the researcher's University.

Method

Participants

Participants were from Hong Kong and Australia and therefore represented two diverse cultures. Hofstede (1980) and Oyserman, Coon, and Kimmelmeyer (2002) found that Australia is one of the most individualistic societies in the world and that collectivist countries could be found in Asia, for example, China, Japan and Hong Kong. In particular, Hofstede (1980) found that Hong Kong samples were lower than their Australian counterparts on individualism and uncertainty avoidance but higher on power distance. These characteristics were typically identified as collectivist. The collectivist cultural value is assumed to be present among the Chinese participants in this study, especially given the strong Chinese social influences, such as exposure to Mainland Chinese influences, and the fact that the Hong Kong students' education was mostly conducted in Cantonese.

Chinese sample. Responses were collected from 395 full-time Chinese students (64% female; 36% male) undertaking business classes at a community college in Hong Kong. This student sample ranged in ages from 18 to 29 (Mean = 19.9, S.D. = 1.16). Ninety-nine percent of the participants were 18–24 years old. The questionnaires were completed anonymously and participation was voluntary without compensation.

Australian sample. The sample was made up of 241 Australian full-time students who were born and had lived all their lives in Australia (69% female; 31% male). All respondents were recruited from two public universities located in Sydney. Ages ranged from 17 to 50 (Mean = 20.4; S.D. = 4.55). Ninety percent of the respondents were between the ages of 17–24 years. All students were undertaking psychology classes and received course credit for their participation. In order to maximize the possible differentiation between the two cultural groups, only data from respondents with European heredity in the Australian sample were used in the analyses.

Measure

Self-leadership was measured using the MSLQ developed by Ho and Nesbit (2009). The MSLQ consists of 38 items describing various behaviours associated with self-leadership and participants using a five-point Likert-type scale (1 = not all accurate; 2 = a little accurate; 3 = somewhat accurate; 4 = mostly accurate; 5 = completely accurate) to indicate how accurate each behaviour describes them. The 12 subscales include Visualizing Successful Performance (two items); Self-goal Setting (four items); Self-talk (three items); Self-reward (three items); Self-punishment (four items); Task-based Natural Reward (four items); Relation-based Natural Reward (four items); Task-based Self-observation (two items); Relation-based Self-observation (two items); Individual-oriented Evaluation of Beliefs and Assumptions (five items); Social-oriented Evaluation of Beliefs and Assumptions (three items); Self-cueing (two items). Sample items include 'I use my imagination to picture myself performing well on important tasks', 'I consciously have goals in mind for my work efforts', 'I think that the enjoyment gained from work is more important than external rewards', and 'When I differ from others' opinions, I try to modify my thinking to avoid conflicts so as to maintain harmony'. See Ho and Nesbit's (2009) validation research for the complete questionnaire.

Adapting self-leadership measure for ensuring cross-cultural equivalence

When a psychological instrument developed in one cultural group is applied in a different cultural context, one cannot assume that psychometric properties are invariant and that the meanings of the scores will be identical. Comparisons between cultural groups are appropriate when empirical evidence demonstrates that 'meaning and dimensional structure of the construct as well as the items comprising the measuring instrument are group-equivalent' (Byrne *et al.*, 2009, p. 95). Measurement invariance is achieved when both the item content and psychometric properties (i.e. validity and reliability) of the instrument are equivalent

across groups (Byrne *et al.*, 2009; Van de Vijver & Poortinga, 2005). According to Leung (2008), when adapting an instrument from one language to another language, cross-cultural equivalence should include: (1) linguistic equivalence; (2) conceptual equivalence; and (3) scalar equivalence. To achieve linguistic equivalence, a key requirement is the accurate translation of the materials for different linguistic groups.

In this study, before the MSLQ could be administered to the Hong Kong respondents, translation into the Chinese language was required. We followed Brislin's (1980) translation/back-translation procedure to create a Chinese version of the questionnaire. The first author, who is bilingual (Chinese and English) and had lived in the United States for four years, translated the MSLQ from English into Chinese. After the Chinese version of the MSLQ was generated, a translation professional then back translated it to English. Three Hong Kong University lecturers examined the original version in English and the back-translated English version and found no back-translation discrepancies.

'Conceptual equivalence' refers to the similarity of the meaning of construct across different cultural groups. For evaluating conceptual equivalence across cultures, one simple way is to determine the factor structure and factor loading of a scale. Conceptual equivalence may exist if a scale shows a similar factor structure and factor loading across cultural groups (see Vandenberg & Lance, 2000; Van de Vijver & Leung, 1997). For making meaningful comparison of means of target variables across cultures, conceptual equivalence is a necessary but insufficient requirement. Scalar equivalence needs to be established (Byrne, 1998; Meredith, 1993). Scalar equivalence can be obtained when the measurement instrument is on the same ratio scale with the same origin in each cultural group (Leung, 2008; Van de Vijver & Tanzer, 1997). Multi-group confirmatory factor analysis (MGCFA) is a common statistical technique used for establishing both the conceptual equivalence and scalar equivalence. Technical procedures for assessing these two levels of equivalence through MGCFA will be outlined in the next section.

Procedure to assess measurement invariance

We used MGCFA and followed a procedure of testing measurement invariance outlined in previous theoretical work (Cheung & Rensvold, 1999; Jöreskog, 1971; Meredith, 1993; Vandenberg & Lance, 2000). In brief, when testing for measurement equivalence across cultures, sets of parameters are constrained in a logically ordered, increasingly restrictive fashion. The first step is to compare the factor structure of the MSLQ across the two cultural groups. Given that in earlier research, Ho and Nesbit (2009)

found an 11-factor model instead of the hypothesized 12-factor model it is necessary to determine which factor structure (11- vs. 12-factor) represents the best model fit for each cultural group. Confirmatory factor analyses of the proposed measurement model are performed separately for the Chinese and Australian groups. If the same factor structure is found in the Chinese and Australian samples, testing of configural invariance would be continued.

The next step would be a test of equivalent factor structures across groups to determine if the patterns of fixed and free factor loadings are comparable (Vandenberg, 2002). The aim is to examine if respondents belonging to different cultural groups have the same number of factors with a similar pattern of factor loadings (Millsap & Everson, 1991; Riordan & Vandenberg, 1994). Configural invariance is also specified as the baseline model against which the subsequent more restrictive models are compared (Salzberger, Sinkovics, & Schlegelmilch, 1999).

Once configural invariance is established, metric invariance, which indicates whether the Chinese and Australian participants are interpreting the items of the survey in the same way, is assessed. This test is conducted by imposing equality constraints on the factor loadings. We followed the recommendation of Cheung and Rensvold (1999) and performed a series of CFA tests using every other item on the subscale as a temporary referent item. If metric invariance exists, the strength of the relationship between specific scale items and the underlying constructs would be the same across groups (Cheung & Rensvold, 2000).

Once metric invariance is established, the fourth step is to explore scalar invariance with a test of equivalence of indicator intercepts (Cheung & Rensvold, 2000). This test is conducted by imposing equality constraints on item intercepts and is conducted to determine whether there are cross-cultural differences in agreement bias (also known as 'acquiescence response style bias') (see Cheung & Rensvold, 2000).

For each of the four steps listed above, evaluation of local areas of misfit would be performed through scrutinizing standardized residuals and modification index (Byrne, 1998). Some researchers advocate that a final step of measurement invariance testing is the need for equivalence of the error variance across groups (Byrne, Shavelson, & Muthen, 1989; Drasgow, 1984, 1987). However, Byrne (1998), Steenkamp and Baumgartner (1998) and Meredith (1993) have suggested that testing the invariance of error variance is not necessary as it represents an overly restrictive test of the data. They argue that only scalar invariance is a necessary condition for the purpose of conducting mean comparisons across cultures. Some researchers even suggest that latent-mean comparisons can still be made under partial scalar equivalence (Byrne *et al.*, 1989; Millsap & Kwok, 2004; Steenkamp & Baumgartner, 1998).

Based on these researchers' guidelines, we exclude the invariant testing of error variance.

Estimation and evaluation of model fit

Overall model fit in MGCFA is commonly evaluated through a chi-square (χ^2) test. However, this test is extremely sensitive to sample size, which often signals statistically significant misfit even for trivial departures from perfect fit. Thus, this measure was not used in present study. While there are no golden rules for assessment of model fit, reporting a variety of indices is necessary for capturing different aspects of model fit (Brown, 2006). Following the recommendations of Hu and Bentler (1999) and Cheung and Rensvold (2000), absolute fit was assessed by the standardized root mean square residual (SRMR) and parsimonious fit was assessed by the root mean square error of approximation (RMSEA). Incremental fit was assessed by the comparative fit index (CFI). Typically a good model fit requires values below 0.08 for SRMR, values around 0.06 for RMSEA (Hu & Bentler, 1999) and values higher than 0.90 for CFI (Bentler & Bonett, 1980). Akaike Information Criterion (AIC) was also used for comparing models with smaller values being better than larger values.

In addition to the measurement equivalence procedure mentioned above, the process also involves testing the fit of a series of increasingly restrictive models against a baseline model with the equality of estimated parameters (e.g. equal factor loading or intercepts). To determine the degree of equivalence, differences in CFI (Δ CFI) and between nested equivalence models is used as a more reliable measure of model fit than the Chi-squared difference test (Brannick, 1995; Cheung & Rensvold, 2002; Kelloway, 1995). Chen (2007) further suggests that when the sample size is small (total $N \leq 300$) and sample sizes are unequal as shown in this study, for the testing of metric (factor loadings) and scalar invariance (intercept), a change of ≤ -0.005 in CFI supplemented by a change of ≤ 0.010 in RMSEA would indicate invariance model.

Results

Data screening

Preliminary examination of the data showed no serious problems with non-normality. All item variables from the Chinese and Australian samples had skewness values < 0.75 and kurtosis values < 1.03 . These values were well below the guidelines of univariate normality (skewness values > 2 and kurtosis values > 7) suggested by Curran, West, and Finch (1996). Hence, maximum likelihood is the appropriate estimation method in the confirmatory factor analyses.

Factor structure: testing the 11-factor and 12-factor measurement models

Multi-sample analyses were performed using maximum likelihood estimation method based on a covariance matrix. Results of the confirmatory factor analyses for each independent sample showed that both the 12-factor and 11-factor models yielded an acceptable model fit to the data. Most fit indices indicated that the 12-factor 38-item model, labelled as Model A, was an adequate model across the Chinese and Australian sample (Chinese sample: $\chi^2 = 1095.97$; SRMR = 0.049; RMSEA = 0.046; CFI = 0.904; AIC = 1388. Australian sample: $\chi^2 = 1054.39$; SRMR = 0.067; RMSEA = 0.056; CFI = 0.893; AIC = 1355). On the other hand, the fit indices of the 11-factor, 38-item measure, labelled as Model B, also yielded very similar results across both samples (Chinese sample: $\chi^2 = 1113.40$; SRMR = 0.048; RMSEA = 0.046; CFI = 0.906; AIC = 1386. Australian sample: $\chi^2 = 1070.41$; SRMR = 0.066; RMSEA = 0.056; CFI = 0.893; AIC = 1356). This result revealed that the items of relation-based self-observation and task-based self-observation could be merged to form a single factor in both Chinese and Australian student samples. Using the principle of parsimony, we recognize the 11-factor model as providing the best model fit to the data. Furthermore, the two relation-based dimensions, namely relation-based natural reward and social-oriented evaluation of beliefs and assumptions, presented as two independent, distinct factors within the 11-factor model across the cultural groups.

As suggested by Byrne (1998), we further examined the modification index and standardized residuals to improve the model fit of the 11-factor model. The modification index demonstrated that two items cross-loaded with other unexpected factors (one item from the self-goal setting construct – ‘I think about goals that I intend to achieve in the future’ and another item from the task-based natural reward construct – ‘I seek out activities in my work that I enjoy doing’). Following the suggestion to delete ‘doublet’ items by Anderson and Gerbing (1988) in conjunction with content considerations, we dropped these two items and tested an alternative Model C that contained 36 items for both samples. Model C showed a somewhat better fit compared to Model B as the fit indices for Model C were slightly better than those of Model B for both samples ($\chi^2 = 978.5$; SRMR = 0.046; RMSEA = 0.045; CFI = 0.910; AIC = 1232, and $\chi^2 = 957.5$; SRMR = 0.064; RMSEA = 0.055; CFI = 0.901; AIC = 1210 for the Chinese and Australian group respectively). In particular, the AIC value of Model C was obviously smaller than that of Model B in both samples.

Consequently, these results suggested that the 11-factor of Model C with 36 items provided the best model fit. We also examined the standardized factor loading estimates for the MSLQ’s items on their latent factors for both groups

separately. All of the factor loading coefficients were positive, statistically significant and moderately high for both Chinese and Australian samples.

Internal consistency reliability

Means, standard deviations, factor reliability and correlation estimates across the Chinese and Australian student samples are provided in Table 1. In general, the factor correlations of both samples were not high, demonstrating that self-leadership’s 11 dimensions represented by these factors are theoretically distinct constructs. In addition, most factors had moderate to high reliability which meet adequate reliability standards. However, the alpha of task-based natural rewards, and social-oriented evaluation of beliefs and assumptions in the Australian student samples and visualizing successful performance in the Chinese student sample were 0.69, 0.67 and 0.67 respectively, which were slightly below but still very close to the threshold of 0.70 suggested by Nunnally (1978).

Configural invariance. Given the good fit of the 36-item 11-factor model to both samples, a multi-group CFA was used for the testing of measurement invariance. First, the equality of the same factor structure was tested across the Chinese and Australian groups (configural invariance). As shown in Table 2, all fit indices indicate a good model fit for both groups (SRMR = 0.047, RMSEA = 0.035, CFI = 0.903).

Metric invariance. Given that configural invariance was supported, we continued to test for metric invariance where factor loadings were constrained to be equal across groups. As described earlier, we used Chen’s (2007) suggested ΔCFI of ≤ -0.005 and $\Delta RMSEA$ of ≤ 0.010 as indicators that the constrained model of metric invariance adequately fit the data. As shown in Table 2, the difference in CFI and RMSEA value between the baseline model (configural invariance) and metric invariance model fell below the critical difference of 0.005 and 0.01. We concluded that metric invariance was supported.

Scalar invariance. Because the metric invariance model was supported, the next step was to test for scalar invariance. Scalar invariance was tested by constraining the intercepts of the 36 indicators to be the same across the two groups. The difference in CFI between the scalar invariance model (Model 3) and the metric invariance model (Model 2) was larger than 0.005. Thus the model of full scalar invariance was not supported. However, as argued by Byrne *et al.* (1989) and Meredith (1993), full metric and scalar invariance is not a necessary condition for making meaningful comparison of means across cultures provided that at least two items per factor exhibit metric and scalar

Table 1 Means, standard deviations, factor reliability estimates and factors correlations for Chinese and Australia students samples (Model C – 36 items)

Factor	Australia Students ($N = 241$)											α		
	1	2	3	4	5	6	7	8	9	10	11		M	SD
1. Visualizing successful performance	1.0	0.35**	0.33**	0.26**	0.13**	0.14**	0.17**	0.40**	0.33**	0.13*	0.19**	3.14	1.03	0.74
2. Self-goal setting	0.51**	1.0	0.33**	0.27**	0.18**	0.22**	0.20**	0.46**	0.47**	0.28**	0.52**	3.45	0.92	0.78
3. Self-talk	0.42**	0.25**	1.0	0.32**	0.25**	0.15**	0.14**	0.34**	0.36**	0.21**	0.21**	3.40	1.13	0.92
4. Self-reward	0.19**	0.27**	0.21**	1.0	0.17**	0.15**	0.20**	0.26**	0.15**	0.20**	0.24**	3.45	1.06	0.92
5. Self-punishment	0.17**	0.13**	0.18**	0.03	1.0	-0.04	0.07	0.27**	0.24**	0.25**	0.12**	3.48	0.86	0.84
6. Task-based natural rewards	0.19**	0.32**	0.21**	0.16*	0.09	1.0	0.42**	0.26**	0.28**	0.27**	0.04	3.01	0.79	0.69
7. Relation-based natural rewards	0.28**	0.34**	0.18**	0.15*	0.07	0.47**	1.0	0.45**	0.40**	0.57**	0.07	3.07	0.77	0.73
8. Self-observation†	0.44**	0.60**	0.17**	0.25**	0.21**	0.33**	0.42**	1.0	0.60**	0.43**	0.20**	3.31	0.72	0.70
9. Individual-oriented evaluation of beliefs and assumptions	0.50**	0.39**	0.39**	0.19**	0.14*	0.33**	0.36**	0.43**	1.0	0.46**	0.19**	2.90	0.87	0.85
10. Social-oriented evaluation of beliefs and assumptions	0.15**	0.20**	0.20**	0.06	0.15*	0.22**	0.52**	0.25**	0.41**	1.0	0.11*	2.87	0.78	0.67
11. Self-cueing	0.23**	0.48**	0.18**	0.17*	0.14*	0.21**	0.24**	0.21**	0.20**	0.17**	1.0	3.75	1.14	0.83
Chinese Students ($N = 395$)														
M	3.20	3.32	3.33	3.78	3.54	3.18	3.45	3.32	3.12	3.41	3.37			
SD	0.81	0.74	0.88	0.80	0.74	0.75	0.62	0.63	0.63	0.64	0.99			
α	0.67	0.77	0.82	0.83	0.78	0.70	0.72	0.73	0.73	0.71	0.81			

* $p < 0.05$; ** $p < 0.01$. Values below the diagonal represent the factor correlations for Australian students and values above the diagonal represent the factor correlations for Chinese students. †This factor includes the items of two subscales (task-based self-observation and relation-based self-observation) which are merged into one single factor in CFA.

Table 2 Results for measurement invariance tests

Model description	χ^2	df	SRMR	RMSEA	CFI	Model comparison	Δ CFI	Δ RMSEA
Model 1: Configural invariance (baseline model)	1935.86	1078	0.047	0.035	0.903	–		
Model 2: Full metric invariance	1988.12	1103	0.047	0.036	0.900	2 vs. 1	–0.003	0.001
Model 3: Full metric and full scalar invariance	2315.33	1139	0.048	0.038	0.868	3 vs. 2	–0.032	0.002
Model 4: Full metric and partial scalar invariance	2162.38	1130	0.047	0.038	0.898	4 vs. 2	–0.002	0.002

CFI, comparative fit index; RMSEA, the root mean square error of approximation; SRMR, Standardized root mean squared residual.

invariance. In order to assess whether a reasonable degree of partial scalar invariance could be achieved, a modification index (MI) was used to locate intercepts that are not invariant across the Chinese and Australian sample (Steenkamp & Baumgartner, 1998).¹ This examination revealed that the significant increase in χ^2 value was due to a lack of scalar invariance of eight indicators from seven factors: self-reward (one item), self-punishment (one item), relationship-based natural reward (one item), task-based natural reward (one item), task and relationship-based self observation (one item), individual-oriented evaluation of beliefs and assumption (two items), social-oriented evaluation of beliefs and assumption (one item). Relaxing the constraints of equal intercepts for these eight indicators yielded substantial improvement in fit as compared to the full scalar invariance model (see Table 2). This partial scalar invariance model (Model 4) was evaluated against the metric invariance model (Model 2). Because the difference of CFI and RMSEA fell below the critical difference of 0.005 and 0.010, we concluded that partial scalar invariance was supported (see Table 2).

Discussion

Self-leadership theory has developed largely within the context of the culture of the USA. Despite the growing academic interest in exploring the influence of culture on individuals' use of self-leadership strategy, the intercultural aspects of self-leadership have received limited attention to date. One possible reason is the lack of a measurement tool that is applicable across different cultures. Without a reliable scale that is demonstrated to be equivalent across cultural groups, any conclusion drawn from comparing the mean scores of self-leadership constructs is questionable. The purpose of this study was to examine the measurement equivalence of a survey instrument (MSLQ) using Chinese and Australian samples representing Eastern and Western cultures respectively.

Our findings provide additional support for the stability of the 11-factor model that was found in previous validation research (Ho & Nesbit, 2009). This study provided solid

support for configural invariance, suggesting that the overall factor structure was demonstrated to be equivalent across the Chinese and Australian groups. This suggested that respondents from Chinese and Western cultures did not differ from one another in terms of the conceptual meaning attached to all self-leadership factors. The metric invariance model was also solidly supported, which further suggests that Chinese and Australian groups used the rating scale in similar ways. However, our findings found that there was a lack of scalar invariance of eight indicators from seven factors. This suggested that the intercepts of these eight indicators are non-equivalent across cultures. Non-equivalent intercepts may be caused by cross-cultural differences in acquiescence response style, which occurs when participants in one cultural group systematically give higher or lower responses than do participants in another group (Cheung, 2008). Furthermore, respondents' higher tendency to respond positively to items in one culture leads to a higher scale origin in that particular culture. This tendency could be explained in terms of social desirability or a belief that a higher score is a better score (Cheung & Rensvold, 2000). Nevertheless, although the MSLQ did not meet full scalar invariance, according to Byrne *et al.* (1989) and Steenkamp and Baumgartner (1998), full scalar invariance is not a necessary condition for comparison of the means of latent factors across cultural groups, provided that at least two items per factor exhibit metric and scalar invariance which was achieved in this study. Thus the MSLQ has met the condition of partial scalar invariance across Western and Eastern cultures.

Internal reliability estimates provide additional evidence for the psychometric properties of both Chinese and English versions of the MSLQ. In most cases, the alpha coefficients were found to be acceptable and comparable across the Chinese and Australian samples. However, the subscale of task-based natural rewards and social-oriented evaluation of beliefs and assumptions in the Australian sample and visualizing successful performance in the Chinese sample did not reach acceptable reliability, although they were both close to the accepted level. Thus it is suggested that caution should be taken when interpreting scores calculated from these factors.

Given that configural, metric and partial scalar invariance were established, our findings of measurement invariance between the Chinese and Australian samples provide evidence that researchers could use this instrument to make valid cross-cultural comparisons (see Little, 1997; Vandenberg & Lance, 2000).

Limitations and future research

As with most research, this study is subject to limitations. The use of a student sample within a narrow age range may limit the generalizability of results, despite the cultural appropriateness of the sample used. Future researchers may consider using a less homogeneous, non-student sample such as employees in a work setting with a wider age range. It would also be of interest to examine whether the MSLQ self-leadership measure could be generalized to other Western and non-Western countries such as Japan, Korea and Malaysia. The marginal reliability estimates of two factors in the Australian sample ('task-based natural rewards', 'social-oriented evaluation of beliefs and assumptions') and one factor in the Chinese sample ('visualizing successful performance') suggest further item modification which may improve the reliability of this scale.

As noted earlier, based upon the research findings of some cross-cultural studies (Hofstede, 1980; Oyserman *et al.*, 2002), our study assumed that the Hong Kong subjects were collectivists, whereas the Australian subjects were classified as individualists. Without measuring the cultural differences of these two samples in terms of the individualism–collectivism orientation (Singelis, Triandis, Bhawuk, & Gelfand, 1995), it is possible that these young Hong Kong students may be open to greater influence from Western culture and be more individualistic in their values, like the Australian sample. However, we are not aware of any empirical research arguing for a convergence of cultural values among young Hong Kong Chinese to Western individualism. Indeed, a number of surveys of Hong Kong residents have found a growing identification with China (Wong, 2010). Furthermore, the Chinese students in our sample speak Cantonese in their education and in their daily lives. While it is unlikely that these two

samples are similar in terms of their cultural value orientation, we nevertheless suggest that future studies utilize specific measures of individualism–collectivism in their research.

Although the overall factor structure was found to be similar across Eastern and Western cultural groups in our study, suggesting that the MSLQ possesses conceptual equivalence, additional research is needed to further confirm the conceptual equivalence of self-leadership measure through examining the nomological network across cultures. According to Leung (2008, p. 61), 'nomological network provides the definitive answer to the question of conceptual equivalence. If a construct is related to similar antecedents and consequences across different cultures, it must be conceptually similar.' Indeed, the self-leadership literature has suggested a number of antecedents and outcomes thought to be associated with the application of self-leadership strategies. These include self-efficacy, personality traits such as conscientiousness and internal locus of control, positive affect, job satisfaction, and psychological empowerment. Future research should continue to investigate the correlations between the MSLQ's factors and other constructs in order to identify similarities and differences in these correlations across cultures.

In conclusion, this study provides evidence that the MSLQ is an appropriate measure to make cross-group comparisons between English and Chinese-speaking participants. This study serves as a springboard for future research to identify potential differences in self-leadership behaviours between Eastern and Western populations.

End note

1. Steenkamp and Baumgartner (1998) recommended that invariance constraints could be relaxed for those intercepts where the MIs are highly significant (both in absolute magnitude and in comparison with the majority of other MIs). In our study, MIs corresponding to eight intercepts were all larger than 40. These eight MIs were much higher than the majority of MIs of other intercepts. Thus, these eight items were identified as non-invariant across the Chinese and Australian samples.

References

- Alderfer, C. P. (1972). *Existence, Relatedness, and Growth: Human Needs in Organizational Settings*. New York: Free Press.
- Alves, J. C., Lovelace, K. J., Manz, C. C., Matsupura, D., Toyasaki, F. & Ke, K. (2006). A cross-cultural perspective of self-leadership. *Journal of Managerial Psychology*, 21, 338–359.
- Anderson, J. & Gerbing, D. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Journal of Applied Psychology*, 103, 411–423.
- Anderson, J. S. & Prussia, G. E. (1997). The self-leadership questionnaire: Preliminary assessment of construct validity. *Journal of Leadership Studies*, 4, 119–143.
- Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588–606.
- Brannick, M. T. (1995). Critical comments on applying covariance structure modeling. *Journal of Organizational Behavior*, 16, 201–213.
- Brislin, R. W. (1980). Translation and content analysis of oral and written materials. In: H. C. Triandis & W. W. Lambert, eds. *Handbook of Cross-Cultural Psychology*,

- Vol. 2. pp. 349–444. Boston, MA: Allyn & Bacon.
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford.
- Byrne, B. M. (1998). *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum.
- Byrne, B. M., Oakland, T., Leong, F. T. L., Van de Vijver, F. J. R., Hambleton, R. K. & Cheung, F. M. (2009). A critical analysis of cross-cultural research and testing practices: Implications for improved education and training in psychology. *Training and Education in Professional Psychology*, 3, 94–105.
- Byrne, B. M., Shavelson, R. J. & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456–466.
- Carver, C. S. & Scheier, M. F. (1998). *Attention and Self-Regulation: A Control Theory Approach to Human Behavior*. New York: Springer-Verlag.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504.
- Cheung, G. W. (2008). Testing equivalence in the structure, means, and variances of higher-order constructs with structural equation modeling. *Organizational Research Methods*, 11, 593–613.
- Cheung, G. W. & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*, 25, 1–27.
- Cheung, G. W. & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross Cultural Psychology*, 31, 187–212.
- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness of fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255.
- Conger, J. & Kanungo, R. (1988). The empowerment process: Integrating theory and practice. *The Academy of Management Review*, 13, 639–652.
- Curran, P. J., West, S. G. & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16–29.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum Press.
- Deci, E. L. & Ryan, R. M. (2000). The ‘what’ and ‘why’ of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, 227–268.
- Devine, D. J., Clayton, L. D., Phillips, J. L., Dunford, B. B. & Melner, S. B. (1999). Teams in organizations: Prevalence, characteristics, and effectiveness. *Small Group Research*, 30, 678–711.
- Drasgow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin*, 95, 134–135.
- Drasgow, F. (1987). Study of the measurement bias of two standardized psychological tests. *Journal of Applied Psychology*, 72, 19–29.
- Ho, J. & Nesbit, P. L. (2009). A refinement and extension of the self-leadership scale for the Chinese context. *Journal of Managerial Psychology*, 24, 450–476.
- Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills, CA: Sage Publications.
- Houghton, J. D. & Neck, C. P. (2002). The revised self-leadership questionnaire: Testing a hierarchical factor structure for self-leadership. *Journal of Managerial Psychology*, 17, 672–691.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Jöreskog, K. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36, 409–426.
- Kelloway, F. K. (1995). Structural equation modeling in perspective. *Journal of Organizational Behavior*, 16, 215–224.
- Kozlowski, S. W. J. & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, 7, 77–124.
- Leung, K. (2008). Methods and measurement in cross-cultural management. In: P. B. Smith, M. F. Peterson & D. C. Thomas, eds. *The Handbook of Cross-Cultural Management Research*, pp. 59–66. Thousand Oaks, CA: Sage Publications.
- Little, T. D. (1997). Mean and covariance structures (MACS) analysis of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, 32, 53–76.
- Luthans, F. & Davis, T. (1979). Behavioral self-management (BSM): The missing link in managerial effectiveness. *Organizational Dynamics*, 8, 42–60.
- Manz, C. C. (1986). Self-leadership: Toward an expanded theory of self-influence processes in organizations. *Academy of Management Review*, 11, 585–600.
- Manz, C. C. (1992). *Mastering Self-Leadership: Empowering Yourself for Personal Excellence*. Prentice-Hall, NJ: Englewood Cliffs.
- Manz, C. C. & Neck, C. P. (2004). *Mastering Self-Leadership: Empowering Yourself for Personal Excellence (3rd Ed)*. Pearson Prentice Hall, NJ: Upper Saddle River.
- Manz, C. C. & Sims, H. P. (1980). Self-management as a substitute for leadership: A social learning theory perspective. *Academy of Management Review*, 5, 361–367.
- Manz, C. C. & Sims, H. P. Jr (1991). Super-leadership: Beyond the myth of heroic leadership. *Organizational Dynamics*, 19, 18–35.
- Manz, C. C. & Sims, H. P. Jr (2001). *New Superleadership: Leading Others to Lead Themselves*. San Francisco, CA: Berrett-Koehler.
- Markus, H. & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525–543.
- Millsap, R. E. & Everson, H. (1991). Confirmatory measurement model comparisons using latent means. *Multivariate Behavioral Research*, 26, 479–497.
- Millsap, R. E. & Kwok, O. (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological Methods*, 9, 93–115.
- Neck, C. P. & Houghton, J. D. (2006). Two decades of self-leadership theory and research: Past developments, present trends, and future possibilities. *Journal of Managerial Psychology*, 21, 270–295.
- Neck, C. P. & Manz, C. C. (1992). Thought self-leadership: The impact of self-talk and mental imagery on performance. *Journal of Organizational Behavior*, 12, 681–699.
- Neck, C. P. & Manz, C. C. (1996). Thought self-leadership: The impact of mental strategies training on employee behavior, cognition, and emotion. *Journal of Organizational Behavior*, 17, 445–467.
- Neck, C. P., Stewart, G. & Manz, C. C. (1995). Thought self-leadership as a framework for enhancing the performance of performance

- appraisers. *Journal of Applied Behavioral Science*, 31, 278–302.
- Neubert, M. J. & Wu, J. C. (2006). An investigation of the generalizability of the Houghton and Neck revised self-leadership questionnaire to a Chinese context. *Journal of Managerial Psychology*, 21, 360–373.
- Nunnally, J. C. (1978). *Psychometric Theory* (2nd Ed.). New York: McGraw-Hill.
- Oyserman, D., Coon, H. M. & Kimmelmeir, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analysis. *Psychological Bulletin*, 128, 3–72.
- Prussia, G. E., Anderson, J. S. & Manz, C. C. (1998). Self-leadership and performance outcomes: The mediating influence of self-efficacy. *Journal of Organizational Behavior*, 19, 523–538.
- Riordan, C. M. & Vandenberg, R. J. (1994). A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management*, 20, 643–671.
- Salzberger, T., Sinkovics, R. R. & Schlegelmilch, B. B. (1999). Data equivalence in cross-cultural research: A comparison of classical test theory and latent trait theory based approaches. *Australasian Marketing Journal*, 7, 27–38.
- Sims, H. P. Jr & Manz, C. C. (1996). *Company of Heroes: Unleashing the Power of Self-Leadership*. New York: Wiley.
- Singelis, T. M., Triandis, H. C., Bhawuk, D. P. S. & Gelfand, M. J. (1995). Horizontal and vertical aspects of individualism and collectivism: A theoretical and measurement refinement. *Cross-Cultural Research*, 29, 240–275.
- Smolek, J., Hoffman, D. & Moran, L. (1999). Organizing teams for success. In: E. Sundstrom, ed. *Supporting Work Team Effectiveness*, pp. 24–62. San Francisco, CA: Jossey-Bass.
- Steenkamp, J. B. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national research. *Journal of Consumer Research*, 25, 78–90.
- Stevens, M. J. & Campion, M. A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management*, 25, 207–228.
- Stewart, G. L., Carson, K. P. & Cardy, R. L. (1996). The joint effects of conscientiousness and self-leadership training on employees self-directed behavior in a service setting. *Personnel Psychology*, 49, 143–155.
- Van de Vijver, F. J. R. & Leung, K. (1997). *Methods and Data Analysis for Cross-Cultural Research*. Newbury Park, CA: Sage.
- Van de Vijver, F. J. R. & Poortinga, Y. H. (2005). Conceptual and methodological issues in adapting tests. In: R. K. Hambleton, P. F. Merenda & C. D. Spielberger, eds. *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*, pp. 39–64. Mahwah, NJ: Erlbaum.
- Van de Vijver, F. J. R. & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 47, 263–279.
- Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, 5, 139–158.
- 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, 4–69.
- Wong, N. (2010, June 23). Hongkongers becoming more Chinese as handover recedes. *The Standard*, p. 9.

Copyright of Asian Journal of Social Psychology is the property of Wiley-Blackwell and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.