RESEARCH ON TRANSLATIONS OF TESTS

Validation of the Dyadic Coping Inventory With Chinese Couples: Factorial Structure, Measurement Invariance, and Construct Validity

Feng Xu

University of Zurich and Guangdong Construction Polytechnic,
Guangdong, China

Peter Hilpert

University of Washington

Ashley K. Randall
Arizona State University

Qiuping Li
Jiangnan University

Guy Bodenmann
University of Zurich

The Dyadic Coping Inventory (DCI, Bodenmann, 2008) assesses how couples support each other when facing individual (e.g., workload) and common (e.g., parenting) stressors. Specifically, the DCI measures partners' perceptions of their own (Self) and their partners' behaviors (Partner) when facing individual stressors, and partners' common coping behaviors when facing common stressors (Common). To date, the DCI has been validated in 6 different languages from individualistic Western cultures; however, because culture can affect interpersonal interactions, it is unknown whether the DCI is a reliable measure of coping behaviors for couples living in collectivistic Eastern cultures. Based on data from 474 Chinese couples ($N = 948$ individuals), the current study examined the Chinese version of the DCI's factorial structure, measurement invariance (MI), and construct validity of test scores. Using 3 cultural groups (China, Switzerland, and the United States [U.S.]), confirmatory factor analysis revealed a 5-factor structure regarding Self and Partner and a 2-factor structure regarding Common dyadic coping (DC). Results from analyses of MI indicated that the DCI subscales met the criteria for configural, metric, and full/partial scalar invariance across cultures (Chinese-Swiss and Chinese-U.S.) and genders (Chinese men and women). Results further revealed good construct validity of the DCI test scores. In all, the Chinese version of the DCI can be used for measuring Chinese couples' coping behaviors, and is available for cross-cultural studies examining DC behaviors between Western and Eastern cultures.

Keywords: Chinese couples, construct validity, dyadic coping, factorial structure, measurement invariance

Supplemental materials: <http://dx.doi.org/10.1037/pas0000329.supp>

Research consistently shows that individuals' stressful experiences are closely associated with mental and physical health problems, such as mood disorders (e.g., depression) and poor immune

and cardiovascular system functioning (e.g., Larzelere & Jones, 2008; Thoits, 2010). Stressful experiences can spill over into one's romantic relationship, negatively affecting partners' relationship behaviors and outcomes (Buck & Neff, 2012; Randall & Bodenmann, 2009). However, such negative spillover effects can be reduced if couples are capable of coping with the stress together, which is referred to as DC (Bodenmann, 2005). Findings over the last decade show that DC is positively associated with relationship satisfaction as well as physical and psychological well-being (e.g., Bodenmann, Meuwly, & Kayser, 2011; Bodenmann, Pihet, & Kayser, 2006; Papp & Witt, 2010).

The 37-item DCI (Bodenmann, 2008) was developed to measure DC behaviors within couples, and is the most widely used measure to assess DC behaviors (see the meta-analysis by Falconier, Jackson, Hilpert, & Bodenmann, 2015). The original German version of the DCI has been translated into 16 languages (Simmons & Lehmann, 2012) and has been validated in six languages (Table 1). To date, all validation studies have been carried out in Western cultural contexts. Because culture can affect interpersonal interactions (Kim, Sherman & Taylor, 2008), it is unknown whether the

This article was published Online First May 16, 2016.

Feng Xu, Department of Psychology, University of Zurich and Guangdong Construction Polytechnic, Guangdong, China; Peter Hilpert, Department of Psychiatry and Behavioral Sciences, University of Washington; Ashley K. Randall, Counseling and Counseling Psychology, Arizona State University; Qiuping Li, Medical College, Jiangnan University; Guy Bodenmann, Department of Psychology, University of Zurich.

Peter Hilpert contributes to the manuscript equally as the first author and shares the first authorship. This research was supported by the China Scholarship Council (Grant 201308440256).

Correspondence concerning this article should be addressed to Feng Xu, Department of Psychology, University of Zurich, Binzmuehlestrasse, 14/23, Zurich, Switzerland, or to Peter Hilpert, Department of Psychiatry and Behavioral Sciences, University of Washington, 1959 Northeast Pacific Street, Office BB1545, Seattle, WA 98195-4944. E-mail: xufenggz@gmail.com or hilpert@uw.edu

Table 1
Summary of Prior DCI Validation Studies

Citation	Language	Sample	Factorial structure	Construct validity	Measurement invariance	Analysis	Items
Bodenman (2008)	German	Individuals	4/1	RAS; COPE		EFA	All items
Donato et al. (2009)	Italian	Couples	5/2 ^a			CFA	All items
Ledermann et al. (2010)	German, Italian, French	Individuals	4/1	PQ; CPQ		CFA	2 items deleted in the Italian and French sample (2, 17)
Falconier et al. (2013)	Spanish	Couples	5/2	CSI; DAS; CTS-R		CFA	7 items deleted (2, 3, 15, 17, 18, 23, 26)
Vedes et al. (2013)	Portuguese	Individuals	5/2	RAS; B-SRQ	Across genders	CFA	All items
Levesque et al. (2014)	English	Couples	4/1	DAS		EFA, CFA	7 items deleted (2, 3, 8, 17, 18, 23, 35)
Randall et al. (2015)	English	Individuals	5/2	RAS; COPE	Across genders and cultures (U.S.– Swiss)	CFA	6 items deleted (2, 3, 9, 17, 18, 24)

Note. 4/1 = 4-factor and 1-factor structure; 5/2 = 5-factor and 2-factor structure; RAS = Relationship Assessment Scale; COPE = Coping Orientations to Problems Experienced Scale; PQ = Partnership Questionnaire; CPQ = Communication Pattern Questionnaire; CSI = Coping Strategies Inventory; DAS = Dyadic Adjustment Scale; CTS-R = Conflict Tactics Scale-Revised; B-SRQ = Brief Sound Relationship Questionnaires; EFA = exploratory factor analysis; CFA = confirmatory factor analysis.

^a Indicates that the Italian version was validated on the basis of the old version of DCI with 39 items across Self, Partner and Common which contains an additional factor (2 items) for Common DC.

DCI is an adequate measure of stress and coping processes in Eastern cultures, such as China. Therefore, the aim of the present study was to validate the Chinese version of DCI.

A Cultural Perspective of Chinese Couples' Stress and Coping Behaviors

Chinese marital culture has been fundamentally shaped by Confucianism and collectivism (Ebrey, 2006; Marshall, 2008). Confucianism emphasizes the importance of marriage and family, which are seen as the foundation of social harmony (Huang & Grove, 2012). Maintaining functional relationships and family harmony are thought to be achieved by fulfilling prescribed gender roles. Typically, husbands take on the responsibility of holding a career, whereas wives take care of the children and manage daily household tasks (Huang & Grove, 2012; Wang, Schoebi, & Perez, 2010). Because of the focus on prescribed gender roles, women in Chinese society may provide more support to their male partners during times of stress (Chen & Li, 2007).

Collectivism is associated with the aim of achieving interpersonal harmony by sacrificing one's own goals for the benefit of the group's goals (Oyserman, Coon, & Kimmelmeier, 2002). Furthermore, in accordance with Chinese familial harmony, partners are expected not to share personal problems and stressful experiences (Kim, Sherman, & Taylor, 2008; Taylor et al., 2004). Given this, stress and coping processes in Chinese couples may be different when compared to couples from more Westernized and individualistic cultures, which place more of an emphasis on individual goals and personal benefits (Oyserman et al., 2002).

In the last three decades, China has undergone remarkable social and economic developments which, in turn, has influenced modern life for Chinese families (Hesketh, Lu, & Xing, 2005; Lu, 2006; Shek, 2006; Zhang, Foley & Yang, 2013). Moreover, although economic reforms have improved living standards, Chinese couples face issues related to economic stress (Quek & Knudson-Martin, 2006) and work-family conflicts (Choi, 2008; Zhang et al., 2013). For example, financial hardship (i.e., financial stress) is

forcing most couples to become dual-career couples (Quek & Knudson-Martin, 2006). As a result of the new dual-career model, the traditional model of the support seeker (husbands as breadwinners) and the support provider (wives as housewives) is changing, which in turn could affect couples' stress and coping behaviors. Thus, there is a need for a well-validated coping measure with sound psychometric properties to assess and interpret Chinese couples' coping behaviors.

The DCI

The original German version of the DCI (37 items) was developed to assess how partners help one another cope with individual stressors (e.g., workload) as well as how partners cope together with common stressors (e.g., parenting; Bodenmann, 2008). The DCI is comprised of five factors related to individual stressors and two factors related to common stressors. The five factors related to individual stressors include: (a) Stress Communication, (b) Emotion-Focused Supportive DC (i.e., expressing understanding), (c) Problem-Focused Supportive DC (i.e., giving advice), (d) Delegated DC (i.e., helping the partner with their responsibilities), and (e) Negative DC (i.e., mocking or ignoring the partner's problems). For each of the five factors, participants rate their self-reported behaviors (Self; 15 items in total) and their perceptions of partner's behaviors (Partner; 15 items in total). As such, item contents and number of items per factor are equivalent across Self and Partner. The two factors related to common stressors include: (a) Emotion-Focused Common DC (i.e., sharing emotions when suffering common stress), and (b) Problem-Focused Common DC (i.e., working out solutions for common stress mutually). Thus, partners rate how they and their partners engage in DC when faced with common stressors (Common DC; 5 items). The DCI includes two additional items that specifically evaluate self-rated satisfaction with the employment of DC strategies; however, these two items are generally excluded from validation studies because they do not assess any theoretically conceptualized DC behaviors (Falconier, Nussbeck, & Bodenmann, 2013; Ledermann et al., 2010).

Prior studies have validated the DCI in different languages (see Table 1). Overall, three psychometric properties should be taken into account when validating the Chinese version of the DCI. First, two different solutions of the 37-item DCI's factorial structure have been reported. Using confirmatory factor analysis (CFA) or exploratory factor analysis (EFA), three studies proposed the theoretically suggested five-factor solution for Self and Partner and two-factor solution for Common DC (35 items, e.g., Falconier et al., 2013), whereas another three studies reported a four-factor solution for Self and Partner and one-factor solution for Common DC (35 items; e.g., Ledermann et al., 2010).¹ In the latter solution, items from the two subscales of Problem- and Emotion-Focused DC by Self and Partner have been combined into a single factor for the four-factor structure (labeled as Supportive DC; 5 items) and items from the two subscales of Problem- and Emotion-Focused DC have been combined for the one-factor structure (labeled as Common DC; 5 items). In addition, not all previous validation studies validated all items in their CFA or EFA models, indicating that some items did not load on conceptually expected latent factors. Two studies validated the complete 37-item DCI (Bodenmann, 2008; Vedes, Nussbeck, Bodenmann, Lind, & Ferreira, 2013), whereas four studies excluding different numbers of items (range: 2–7; Falconier et al., 2013; Ledermann et al., 2010; Levesque, Lafontaine, Caron, & Fitzpatrick, 2014; Randall, Hilpert, Jimenez-Arista, Walsh, & Bodenmann, 2015). Excluding items resulted in an unequal amount of items for the final solution across Self and Partner. Given these inconsistencies, it is necessary to find a common factorial structure which applies to both Western and Eastern people.

Second, confirming MI is a critical step for validating an instrument, as diverse cultural groups may interpret psychological phenomena differently due to cultural characteristics (Schwartz et al., 2014). Thus far, only the English (Randall et al., 2015) and Portuguese (Vedes et al., 2013) DCI validation studies have tested for MI across genders, providing evidence for full configural, metric, and scalar invariance in the five-factor and two-factor solution. Moreover, Randall and colleagues (Randall et al., 2015) further tested for MI across cultures (U.S.–Swiss) and results supported full configural, metric invariance and partial scalar invariance.

Third, it is important to examine construct validity of the DCI test scores. Prior DCI validation studies have tested construct validity of DCI test scores by correlating DCI subscales to other measures, such as relationship satisfaction (convergent validity) and individual coping (discriminant validity). As the DCI assesses husbands' and wives' behaviors from self-rated as well as from partner-rated perspectives, construct validity of the DCI test scores (i.e., convergent and discriminant validity) is best tested using a multitrait-multimethod matrix (MTMM, Campbell & Fiske, 1959).

Present Study

The present study sought to validate the Chinese version of the DCI by examining three corresponding aims: (a) factorial structure, (b) MI of the factorial structure, and (c) construct validity of DCI test scores. First, to determine the best-fitting model, the factorial structure was examined by comparing the five-factor and two-factor solution and the four-factor and one-factor solution within samples from three countries: China, Switzerland (where

the DCI originated), and the U.S. (where our measure was translated from). In order to find the best-fitting model, we compared models with five-factor and two-factor solution and the four-factor and one-factor solution using all items as well as dropping items that have been found problematic in prior studies (e.g., Randall et al., 2015). It is crucial to note that the factorial structure of the Chinese version of the DCI should uphold the same structure between cultures in order to further test for MI across cultures and genders.

Second, in order to test for MI, we conducted several CFA analyses to determine whether the DCI constructs were equivalent to configural, metric, and scalar invariance levels across cultures and genders. *Configural invariance* implies that the same factorial structure can be established across groups (i.e., culture and gender), indicating that samples from different groups interpret latent constructs in the same way. *Metric invariance* implies that factor loadings are equivalent for corresponding items across groups, indicating that a construct has similar meaning across groups and predictive associations with other constructs can be compared across groups (Schwartz et al., 2014). Additionally, *scalar invariance* examines whether each item has the same item intercept across groups, indicating that latent factor means across groups can be directly compared (Groth-Marnat, 2009; Schwartz et al., 2014). Assuming all three types of MI can be confirmed, this would provide evidence that items and constructs of the DCI are not unique to Western couples; rather, this would indicate that the DCI is also a valid measure for Chinese couples.

Finally, we used the MTMM approach (Campbell & Fiske, 1959) to test for construct validity of the DCI test scores. This approach allowed us to examine for (a) *convergent validity*: whether a trait (e.g., husbands' Stress Communication) can be captured across different methods (Self- vs. Partner-ratings); (b) *discriminant validity*: whether each trait comprises a unique variability and, therefore, each trait is distinct from each other; and (c) *method variance*: whether the variability between traits is attributed to either the individual trait or to different methods of assessment (Davis, 1971). In order to test for convergent validity, discriminant validity, and method variance, three MTMM matrices were computed for husbands' behaviors (DC by Self and Partner), wives' behaviors (DC by Self and Partner), and their Common DC behaviors.

Method

Participants and Procedure

Data were collected from a convenience sample of heterosexual couples from three metropolitan cities (Beijing, Shanghai, and Guangzhou) and several provinces (e.g., Guangdong, Jiangsu, Shanxi, and Jiangxi) in China. Participants were recruited by distributing flyers at parent meetings in schools and to administrative managers of different enterprises (civil institutions, companies, schools, and hospitals). Participants had to meet the following

¹ The Italian version was validated on the basis of the old 41-item DCI by testing two additional items for Negative DC and 2 additional items for Common DC (Donato et al., 2009) and these four items were dropped in the widely used 37-item DCI by the original author (Bodenmann, 2008).

criteria in order to be able to participate: (a) married (legal age to marry in China is 22 years old for males and 18 years old for females), and (b) currently living with their spouses. Interested participants who met the inclusion criteria contacted research assistants via e-mail or phone and were then sent a letter including information about the study, two participation agreements, two sets of questionnaires, and two preaddressed and stamped envelopes. Couples were instructed to sign the participation agreements, fill out their questionnaires independently, and send the scales back to the research assistant in separate envelopes. Participants were notified that their data would be kept anonymous and confidential for scientific purposes.

Approximately 600 couples showed interest in the study, and 501 couples (84%) sent back the questionnaires. From 501 couples, data from 27 couples (5.4%) were excluded because they either did not fill out the questionnaires or husbands' and wives' answers were identical. The final sample included data from 474 couples ($N = 948$ individuals). The husbands' mean age was 36.5 years ($SD = 7.7$) and the wives' mean age was 34.4 years ($SD = 7.3$). On average, couples had been married for 9.4 years ($SD = 7.9$). Eighty-four percent of all couples had at least one child ($M = 1.0$, $SD = .66$, range 0–3). Approximately 17% of husbands reported to hold a senior high school degree or other diplomas, 24% reported that they received a 3-year vocational college diploma, 40% finished an undergraduate degree, and 19% held a graduate degree. Approximately 17% of wives reported a senior high school degree or other diplomas, 27% reported that they received a 3-year vocational college degree, 38% finished an undergraduate degree, and 18% held a graduate degree. Husbands reported their employment as follows: 35% work in civil and public institutions, 14% work in state-owned enterprises, and 33% work in private companies, while 17% are self-employed. Wives reported their employment as follows: 38% work in civil and public institutions, 14% work in state-owned enterprises, and 36% work in private companies, while 10% are self-employed. Only 2% of wives were unemployed and reported being housewives. On average, husbands earned a monthly salary of 6,000 to 10,000 CNY (approximately \$1,000 to \$1,700 USD during data collection in 2014), whereas wives earned around 2,000 to 6,000 CNY (approximately \$350 to \$1,000 USD).²

Swiss and U.S. Samples

To evaluate factorial structure and MI we utilized a sample of heterosexual couples from Switzerland and a sample of heterosexual individuals currently in a romantic relationship from the U.S. The Swiss sample included 330 couples ($N = 660$ individuals). Couples ranged in age from 18 to 65 years old ($M_{\text{men}} = 41.4$, $SD = 9.2$; $M_{\text{women}} = 40.0$, $SD = 9.0$) and had been in a relationship for an average of 12.9 years ($SD = 8.6$). This data was collected in Switzerland for a randomized controlled trial (Bodenmann, Hilpert, Nussbeck, & Bradbury, 2014). The U.S. sample included 969 heterosexual individuals currently in a romantic relationship (29.4% men, 70.6% women). These individuals ranged in age from 18 to 75 years old ($M_{\text{men}} = 32.1$, $SD = 12.3$; $M_{\text{women}} = 33.2$, $SD = 12.0$). Participants reported being in a relationship with their partner for an average of 10.3 years ($SD = 10.1$). This data was used in the English validation study of the DCI (Randall et al., 2015).

Translation Procedures of the DCI

We translated the Chinese version of the DCI from the English version (Randall et al., 2015), as researchers would like to accept a validated instrument in English (Ziegler & Bensch, 2013). Furthermore, the English version has been validated with a Canadian sample with college students (Levesque et al., 2014) and a U.S. sample with individuals in a romantic relationship (Randall et al., 2015). To translate the English version of the DCI into Chinese we followed recommendations by Sireci and colleagues (Sireci, Wang, Harter, & Ehrlich, 2006). One Chinese-English translator and one Chinese psychology professor translated the English version of the DCI into Chinese, and then two bilingual translators (Chinese and English) back-translated the DCI separately. Translators only detected minor discrepancies (three sentences in Chinese) which were discussed and solved by mutual agreements. To guarantee full consistency of all items, the back-translated English version was checked by the author of the original instrument.

Measures

DCI. The DCI (Bodenmann, 2008) is a self-report inventory of 37 items measured on a 5-point Likert scale, ranging from 1 (*not at all/very rarely*) to 5 (*very often*). To assess couples' stress and coping behaviors, the DCI includes five subscales for the Self and corresponding subscales for the Partner: (a) Stress Communication by Self (e.g., "I tell partner openly how I feel and appreciate his/her support"; 4 items) and by Partner (e.g., "My partner tells me openly how she/he feels and appreciates my support"; 4 items), (b) Emotion-Focused Supportive DC by Self (e.g., "I show empathy and understanding to my partner"; 3 items) and by Partner (e.g., "My partner shows empathy and understanding to me"; 3 items), (c) Problem-Focused Supportive DC by Self (e.g., "I tell my partner that his/her stress is not that bad and help him/her to see the situations in a different light"; 2 items) and by Partner (e.g., "My partner helps me to see stressful situations in a different light"; 2 items), (d) Delegated DC by Self (e.g., "I take on things that my partner would normally do in order to help him/her out"; 2 items) and by Partner (e.g., "My partner takes on things that I normally do in order to help me out"; 2 items), and (e) Negative DC by Self (e.g., "I blame my partner for not coping well enough with stress"; 4 items) and by Partner (e.g., "My partner blames me for not coping well enough with stress"; 4 items). The DCI also assesses two Common DC behaviors: (a) Emotion-Focused Common DC (e.g., "We help each other relax with such things like massage, taking a bath together, or listening to music together"; 2 items), and (b) Problem-Focused Common DC (e.g., "We try to cope with the problem together and search for ascertained solutions"; 3 items). Additionally, the DCI includes two additional items for assessing couples' satisfaction with the effectiveness of DC behaviors.

² According to reports released by the Ministry of Human Resources and Social Security of China in 2014, the average salary varies from city to city on the basis of the economy level. The average salary in the developed areas, like Beijing and Shanghai, is around 6000 CNY per month (around \$1,000 USD) while the average salary in less developed areas, like Western parts of China, is around 2000 CNY (around \$350 USD). Thus, our participants' average salary falls into the category of the middle class.

Data Analytic Strategies

Aim 1: Testing competing CFA models for factorial structure. In order to find the best-fitting model, we used CFA modeling to test competing DCI measurement models (5-factor and 2-factor vs. 4-factor and 1-factor) within three cultural groups (Chinese, Swiss, and U.S.) for both genders. Given that some items were problematic in the U.S. sample (Randall et al., 2015), we relied on modification indices (MIs) and parameter estimates (PE) to provide information about potential model misfit (see Brown, 2015). Based on this information, we allowed items' error estimates to covary. Moreover, if statistically justifiable, we removed problematic items to achieve acceptable model fit in all samples (Brown, 2015). In order to test CFA models, we used the robust maximum likelihood estimator adjusting for nonnormally distributed data. The full information maximum likelihood was used to handle missing data. We relied on common fit indices to evaluate model to data fit: χ^2 , comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean squared residual (SRMR). Using a chi-square to test for differences between models is frequently used, however, this test is too sensitive for a large sample size ($N = 300$; Kline, 2011; MacCallum, Browne, & Cai, 2006). Given this, we relied on the principle that the χ^2/df ratio should be smaller than 3 (Schermelele-Engel, Moosbrugger, & Muller, 2003) and made use of the alternative criteria (Akaike information criterion [AIC]; see Brown, 2015) to compare models. A good or acceptable model to data fit is indicated by a CFI above .95, a RMSEA smaller than .08, and a SRMR smaller than .08 (Hu & Bentler, 1999; McDonald & Ho, 2002). Lower values of AIC indicate a better model fit than compared models (Brown, 2015).

Aim 2: Testing MI. We used the best-fitting factorial structure derived from Aim 1 to examine MI across Chinese-Swiss and Chinese-U.S. cultures for both genders. Specifically, we used a series of nested multigroup CFA models with increasing parameter constraints to test for configural, metric, and scalar invariance. Configural invariance was tested in a model where the factorial structure was invariant across groups with no equality constraints imposed, relying on common model fit indices (χ^2 , CFI, RMSEA, and SRMR; Chen, 2007; Schwartz et al., 2014). Metric invariance was tested by fitting models where factor loadings on respective items were constrained to be equal across groups. Scalar invariance was tested by additionally constraining intercepts to be equal across groups. In order to evaluate metric and scalar invariance models, we followed Chen's (2007) simulation study, which suggests that ΔCFI , $\Delta RMSEA$, and $\Delta SRMR$ could be used to inspect changes in model fit between nested models. The difference of fit indices between these models should be smaller than .010 for ΔCFI , .015 for $\Delta RMSEA$, and .030 for $\Delta SRMR$ in a large sample size ($N > 300$; Chen, 2007). Given that full scalar MI models could not be found, we tested for partial scalar MI. We relied on the highest MIs to identify which item intercepts needed to be freely estimated (Brown, 2015), which would indicate that some but not all intercepts were found to be equivalent across groups. We followed Schwartz and colleagues' (2014) suggestion that partial scalar invariance is still tenable if less than half of item intercepts are freely estimated.

In addition to MI analyses across cultures, we further examined MI across genders within the Chinese sample after restructuring

the dyadic data. Following recommendations by Verhofstadt and colleagues (Verhofstadt, Buysse, Rosseel, & Peene, 2006), we randomly allocated couples to two groups and then selected women from one group and men from the second group. To avoid selection bias, we repeated this procedure 1,000 times and calculated average fit indices across models within the 1,000 samples.

Finally, given that full or partial scalar MI can be established, we tested whether latent means were significantly different between groups. One group was chosen as a reference group with its latent means fixed to zero (e.g., Chinese men) whereas factor means of the other group was freely estimated (e.g., U.S. and Swiss men; Schwartz et al., 2014).

Aim 3: Testing construct validity. The MTMM approach is based on a correlation matrix. We computed three correlation matrices (husbands' behaviors, wives' behaviors and Common DC behaviors) to test for convergent validity, discriminant validity, and method variance. Convergent validity can be confirmed if the validity correlation coefficients (monotrait-heteromethod) are different from zero (Campbell & Fiske, 1959) and exceed a minimum correlation of $r \geq .25$ (Zou, Schimmak, & Gere, 2013). Discriminant validity can be confirmed if convergent validity coefficients (monotrait-heteromethod) are significantly larger than coefficients based on different traits and different methods (heterotrait-heteromethod) involving dependent correlations with a common index (e.g., comparing r_{12} with r_{13} ; 84 comparisons). We used one-tailed Steiger's z test (Steiger, 1980) to compare if traits are independent of each other. Finally, an influence of the method can be detected if the heterotrait-monomethod coefficients are significantly larger than the heterotrait-heteromethod coefficients. This was also tested with one-tailed Steiger's z test (Steiger, 1980; 42 comparisons).

The statistical software package R (R version 3.1; R, Development Core Team, 2008) was used to compute descriptive statistics and correlation matrices. The lavaan package (Rosseel, 2012) in R was used to test all structural equation models (CFA and multigroup CFA) and the cocor package (Diedenhofen & Musch, 2015) to compare correlation coefficients.

Results

Aim 1: Testing Competing CFA Models for Factorial Structure

CFA models (35-item version) for the four-factor and one-factor solution as well as for the five-factor and two-factor solution were computed separately for each sample across both genders. Results revealed that the four-factor and one-factor solution did not fit the data well for both genders in three samples (Tables 2 and 3). For the Chinese sample, the five-factor solution (where we separated the Supportive DC subscale into Problem- and Emotion-Focused Supportive DC subscales) and the two-factor solution (where we separated the Common DC subscale into Problem- and Emotion-Focused Common DC subscales) yielded a good model to data fit. However, the five-factor and two-factor solution for Swiss and U.S. samples did not show adequate fit. After a closer examination of items, we found four problematical items from two subscales of Stress Communication by Self (Items 2 and 3) and by Partner (Items 17 and 18). Based on MIs and PE, we found that covariances between items in each subscale (Self/Partner) were extremely high for three samples: Items

Table 2
Fit Indices for CFA Models of DCI Across Three Samples for Men

Models	Items	χ^2	df	CFI	SRMR	RMSEA (90% CI)	AIC	$\Delta\chi^2$	p
Chinese (N = 474)									
4-factor/Self	15	179.27	84	.930	.044	.049 (.040, .058)	17,376		
5-factor/Self	15	156.38	80	.945	.041	.045 (.035, .054)	17,354	21.34	.00
5-factor/Self ^a	13	104.61	55	.958	.036	.044 (.032, .055)	15,033	51.92	.00
4-factor/Partner	15	155.56	84	.955	.048	.042 (.033, .052)	17,467		
5-factor/Partner	15	124.87	80	.972	.043	.034 (.023, .045)	17,438	32.25	.00
5-factor/Partner ^a	13	84.29	55	.979	.039	.034 (.020, .046)	15,236	40.59	.03
1-factor/Common	5	47.70	5	.924	.054	.134 (.104, .167)	5,604.1		
2-factor/Common ^a	5	12.50	4	.985	.023	.067 (.030, .107)	5,564.4	33.24	.00
Swiss (N = 330)									
4-factor/Self	15	181.66	84	.918	.057	.057 (.046, .067)	12,870		
5-factor/Self	15	161.23	80	.932	.052	.053 (.041, .064)	12,654	17.22	.00
5-factor/Self ^a	13	84.59	55	.968	.037	.038 (.021, .054)	10,895	75.11	.00
4-factor/Partner	15	219.27	84	.900	.065	.066 (.056, .077)	13,259		
5-factor/Partner	15	180.63	80	.926	.061	.059 (.048, .069)	13,025	39.13	.00
5-factor/Partner ^a	13	71.18	55	.985	.032	.028 (.000, .045)	11,213	110.81	.00
1-factor/Common	5	86.43	5	.862	.094	.211 (.174, .250)	4,417		
2-factor/Common ^a	5	2.56	4	1.00	.013	.000 (.000, .065)	4,333	66.15	.00
U.S. (N = 285)									
4-factor/Self	15	221.13	84	.898	.069	.079 (.067, .091)	9,573.4		
5-factor/Self	15	153.85	80	.945	.057	.059 (.046, .073)	9,509.7	84.87	.00
5-factor/Self ^a	13	88.47	55	.973	.039	.048 (.029, .065)	8,177.2	65.97	.00
4-factor/Partner	15	272.79	84	.873	.074	.092 (.081, .104)	9,831.1		
5-factor/Partner	15	228.77	80	.900	.068	.084 (.072, .096)	9,790.0	43.65	.00
5-factor/Partner ^a	13	140.83	55	.945	.054	.077 (.062, .092)	8,366.6	88.86	.00
1-factor/Common	5	66.13	5	.857	.074	.212 (.175, .251)	3,369.3		
2-factor/Common ^a	5	9.9	4	.986	.023	.073 (.013, .132)	3,291.0	39.38	.00

Note. df = degrees of freedom; CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; 90% CI = 90% confidential interval; AIC = Akaike information criterion; $\Delta\chi^2$ = chi-square difference test. Two evaluation items (Items 36 and 37) were excluded for the analysis and (Items 2, 3, 17, and 18) were deleted for testing final models.

^a Indicates results for the final model.

2 and 3 (e.g., U.S. men/women: MIs = 26.20/28.57, PE = .26/.17); Items 17 and 18 (e.g., U.S. men/women: MIs = 29.35/53.09, PE = .28/.29). Subsequently, we allowed error estimates of two pairs of items to covary. However, the models still did not show satisfactory fit to the U.S., sample. (e.g., U.S. men: CFI = .93; U.S. women CFI = .92). In order to attain a common factorial structure for three samples and further test for MI across samples, we excluded four problematic items of the Stress Communication subscale (two for Self, two for Partner). Excluding these items is further justified as they theoretically overlap with the Delegated DC subscale (e.g., "I ask my partner to do things for me when I have too much to do"). Therefore, we recomputed CFA models without these four items, resulting in satisfactory model indices across three samples for both genders (see Table 2 and 3). AIC values for the final five-factor and two-factor model across three samples for both genders were the lowest, further strengthening the robustness of the final models. Thus, the five-factor and two-factor solution based on a 31-item version was used to test all subsequent steps. Figure 1 shows standardized factor loadings for the final models (factor loadings for all studied models can be viewed in the supplemental material).

Aim 2: Testing MI

MI across cultures. Analyses demonstrated good data fit for configural invariance across the Chinese-Swiss sample (Table 4) as well as for the Chinese-U.S. sample (Table 5), indi-

cating that the best-fitting factorial structure was confirmed within three cultural groups. Furthermore, we found full metric invariance with satisfactory model indices for the Chinese-Swiss sample as well as for the Chinese-U.S. sample. Finally, we tested for scalar invariance by constraining item intercepts to be equal across groups. Because most models differed significantly from metric invariance models, we allowed four intercepts with the highest MIs to vary freely for Self and Partner subscales (Self: Items 23, 26, 22, and 28; Partner: Items 6, 7, 14, and 15) and two intercepts to vary freely for the Common DC subscale (Items 32 and 33) for the Chinese-Swiss sample. Similarly, we allowed four intercepts with the highest MIs to vary freely for Self and Partner subscales (Self: Items 3, 26, 28, and 30; Partner: Items 7, 8, 12, and 14) and two items for the Common DC subscale (Items 33 and 35) for the Chinese-U.S. sample. Overall, further analyses showed that partial scalar invariance could be established. It is worth noting that scalar invariance can still be maintained if less than half of the intercepts are freely estimated (Schwartz et al., 2014).

MI across genders. Table 6 shows results of MI analyses across genders (Chinese men and women) for the best-fitting model (5-factor and 2-factor solutions, 31 items). Findings confirmed full configural, metric, and scalar invariance for Self, Partner, and Common DC subscales. Results fully support the assumption that both Chinese husbands and wives respond to each item within each subscale in a similar way.

Table 3
Fit Indices for CFA Models of DCI Across Three Samples for Women

Models	Items	χ^2	df	CFI	SRMR	RMSEA (90% CI)	AIC	$\Delta\chi^2$	p
Chinese (N = 474)									
4-factor/Self	15	157.42	84	.952	.049	.043 (.033, .052)	17,686		
5-factor/Self	15	128.90	80	.968	.045	.036 (.025, .046)	17,658	27.89	.00
5-factor/Self ^a	13	92.66	55	.969	.044	.038 (.026, .050)	15,502	35.73	.08
4-factor/Partner	15	153.81	84	.952	.049	.043 (.033, .052)	18,184		
5-factor/Partner	15	141.20	80	.968	.045	.036 (.025, .046)	18,178	13.09	.01
5-factor/Partner ^a	13	86.69	55	.969	.044	.038 (.026, .050)	15,900	55.17	.00
1-factor/Common	5	68.36	5	.863	.061	.164 (.132, .197)	5,713.7		
2-factor/Common ^a	5	8.89	4	.989	.025	.051 (.000, .092)	5,648.9	84.71	.00
Swiss (N = 330)									
4-factor/Self	15	203.91	84	.908	.057	.062 (.052, .073)	13,105		
5-factor/Self	15	187.10	80	.918	.053	.060 (.050, .071)	13,093	15.60	.00
5-factor/Self ^a	13	114.61	55	.945	.042	.054 (.041, .068)	11,405	73.04	.00
4-factor/Partner	15	223.39	84	.928	.059	.068 (.058, .078)	13,708		
5-factor/Partner	15	181.90	80	.947	.054	.059 (.049, .070)	13,667	34.86	.00
5-factor/Partner ^a	13	92.65	55	.977	.037	.043 (.028, .058)	11,942	89.90	.00
1-factor/Common	5	117.39	5	.829	.105	.247 (.210, .287)	4,473.8		
2-factor/Common ^a	5	8.799	4	.993	.021	.057 (.000, .109)	4,427.7	99.67	.00
U.S. (N = 285)									
4-factor/Self	15	319.04	84	.890	.056	.066 (.059, .073)	22,904		
5-factor/Self	15	233.09	80	.929	.050	.054 (.047, .062)	22,812	93.14	.00
5-factor/Self ^a	13	97.49	55	.978	.031	.035 (.024, .045)	19,453	139.05	.00
4-factor/Partner	15	358.94	84	.921	.061	.072 (.065, .079)	23,895		
5-factor/Partner	15	275.30	80	.944	.058	.062 (.055, .069)	23,798	77.12	.00
5-factor/Partner ^a	13	129.08	55	.977	.031	.046 (.037, .055)	20,343	153.23	.00
1-factor/Common	5	162.89	5	.863	.083	.216 (.192, .242)	8,245.5		
2-factor/Common ^a	5	8.49	4	.997	.010	.035 (.000, .057)	8,044.8	123.57	.00

Note. df = degrees of freedom; CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of Approximation; 90% CI = 90% confidential interval; AIC = Akaike information criterion; $\Delta\chi^2$ = chi-square difference test. Two evaluation items (Items 36 and 37) were excluded for the analysis and (Items 2, 3, 17, and 18) were deleted for testing final models.

^a Indicates results for the final model.

Descriptive and latent mean comparisons. After establishing scalar invariance, we conducted descriptive statistics and latent mean comparisons across cultures and genders (Table 7). Comparing latent means across cultures, we found a number of significant differences across Chinese-Swiss and Chinese-U.S. samples. In general, Westerners scored higher in Supportive, Delegated, and Common DC behaviors whereas Chinese scored higher in Negative DC. Comparing latent means between Chinese husbands and wives, we found that Chinese couples only differed in Stress Communication behaviors: husbands scored higher in Stress Communication by Self whereas wives rated higher in Stress Communication by Partner.

Reliability. The reliability of test scores is shown in Table 7 and results revealed that the reliability of most test scores were good or acceptable; however, two of them were poor ($\alpha = .51$ for husbands' Problem-Focused Supportive DC by Self and $\alpha = .52$ for wives' Delegated DC by Partner). Notably, lower reliability of test scores can be accounted for subscales consisting of two items (de Vaus, 2002). In line with current suggestions (e.g., George & Mallery, 2003; Nunnally & Bernstein, 1994), the reliability of test scores between .50 and .60 is poor but still acceptable.

Aim 3: Testing Construct Validity

Table 8 displays three MTMM correlation matrices. In all three matrices, convergent validity coefficients (monotrait-

heteromethod) were significant and higher than the minimum value of $r = .25$. Overall, we found moderate convergent correlation coefficients for husbands' behaviors ($r_{\text{average}} = .32$, range: .27 to .36), for wives' behaviors ($r_{\text{average}} = .32$, range: .26 to .37), and for the two dimensions of Common DC behaviors ($r_{\text{emotion-focused}} = .38$, $r_{\text{problem-focused}} = .38$). These findings provide evidence for convergent validity of the DCI test scores. Examination of discriminant validity of DCI test scores showed that convergent validity coefficients (monotrait-heteromethod) were significantly larger than discriminant validity coefficients (heterotrait-heteromethod) in 75 out of 84 comparisons, indicating that each trait comprises a unique variance. Of note, nonsignificant results were found in four comparisons for husband' behaviors (e.g., coefficient comparisons between Emotion-Focused Supportive DC by Self and Stress Communication by Partner) and five comparisons for wives' behaviors (e.g., coefficient comparisons between Delegated DC by Self and Problem-Focused Supportive DC by Partner). Overall, these results supported discriminant validity of the DCI test scores. Finally, we tested for method variance and found that heterotrait-monomethod coefficients were significantly larger than heterotrait-heteromethod coefficients in 24 of 42 comparisons, which indicates a rater's evaluation bias. Results of all the correlation comparisons are not presented here but are available upon request from the authors.

Figure 1. The standardized factor loadings for five-factor structure for DC by Self (husbands and wives) and for DC by Partner (husbands and wives) as well as two-factor structure for Common DC (husbands and wives). The two models are based on the final CFA models. 1/16 and likewise numbers refer to the items by Self and Partner, while 31 and likewise numbers refer to Common DC items. DC = dyadic coping.

Discussion

The DCI is a widely used self-report measure to assess couples' stress and coping behaviors (see Falconier et al., 2015); however, to date, the DCI had not been validated for use in non-Western cultures. The goal of the present study was to validate a Chinese version of the DCI by establishing the best-fitting factorial structure and MI as well as examining construct validity of test scores.

The Best-Fitting Factorial Structure

Our first aim was to test competing CFA models in order to find the best-fitting factorial structure. Prior DCI validation studies tested either a five-factor and two-factor solution (e.g., Randall et al., 2015) or a four-factor and one-factor solution (e.g., Ledermann et al., 2010), but no study had compared both solutions. Comparing the two solutions allows us to decide which one is more appropriate for use in empirical studies. Therefore, we tested and compared both solutions within three cultural groups using all 35 items. Our results of comparing these two solutions supported the modified five-factor and two-factor solution by excluding four items from the subscale of Stress Communication (Self: Items 2 and 3; Partner: Items 17 and 18), which is in line with prior validation studies (e.g., Randall et

al., 2015). Confirmation of the five-factor and two-factor solution indicates that participants from all three cultures are able to distinguish between problem- and emotion-focused coping behaviors (not found in the 4-factor and 1-factor solution). Additionally, confirming the five-factor and two-factor solution provides additional support for all theoretically formulated factors of the DCI.

MI and Latent Mean Comparisons

The second aim was to use the best-fitting model to test whether MI could be established across three cultural groups (Chinese, Swiss, and the U.S.) and across genders. Findings of the multigroup CFA analyses demonstrate full configural, metric invariance, and full/partial scalar invariance. First, the establishment of full configural invariance indicates that same items are associated with the same latent factors across cultures. Second, results suggest that factor loadings for all items are consistent across Western and Eastern populations. In relation to scalar invariance, we only found partial MI. Cultural factors, such as collectivistic values, can affect the intercept of items (Chen, 2007; Schwartz et al., 2014), which could suggest that couples who live in collectivistic cultures may conceptualize their behaviors differently compared with couples who live in individualistic cultures. In addition, findings of MI across genders supported full configural, metric, and scalar invariance, indicating that factorial structure, factor

Table 4
Model Indices for Measurement Invariance Across the Chinese-Swiss Sample

Invariance types	Goodness-of-fit indices					Comparison of nested model						
	χ^2	df	CFI	SRMR	RMSEA	Contrast	$\Delta\chi^2$	Δdf	ΔCFI	$\Delta SRMR$	$\Delta RMSEA$	p
Men												
DC by Self												
Configural	190.75	110	.962	.037	.042							
Metric	202.10	118	.960	.040	.041	2 vs. 1	11.35	8	.002	.003	.001	.16
Full scalar	384.97	126	.894	.056	.065	3 vs. 2	182.87	8	.066	.016	.024	.00
Partial scalar	216.74	123	.956	.042	.043	4 vs. 2	14.64	5	.004	.002	.002	.01
DC by Partner												
Configural	156.06	110	.981	.036	.032							
Metric	171.98	118	.978	.041	.033	6 vs. 5	15.92	8	.003	.005	.001	.05
Full scalar	200.54	126	.970	.044	.038	7 vs. 6	28.56	8	.008	.003	.005	.00
Common												
Configural	16.97	8	.993	.019	.052							
Metric	24.04	11	.990	.033	.053	9 vs. 8	7.07	3	.003	.014	.001	.08
Full scalar	99.46	14	.936	.062	.120	10 vs. 9	75.42	3	.066	.029	.067	.00
Partial scalar	24.36	12	.991	.033	.049	11 vs. 9	.32	1	.001	.000	.004	.57
Women												
DC by Self												
Configural	205.57	110	.957	.043	.045							
Metric	233.34	118	.948	.050	.048	13 vs.12	27.77	2	.009	.007	.003	.00
Full scalar	375.58	126	.886	.064	.069	14 vs.13	142.24	8	.062	.014	.021	.00
Partial scalar	238.05	121	.947	.050	.048	15 vs.13	4.71	3	.001	.000	.000	.20
DC by Partner												
Configural	179.01	110	.976	.038	.039							
Metric	202.09	118	.970	.046	.041	17 vs.16	23.08	8	.006	.008	.002	.00
Full scalar	306.32	126	.936	.060	.059	18 vs.17	104.23	8	.034	.014	.018	.00
Partial scalar	236.46	122	.959	.048	.047	19 vs.17	34.37	4	.011	.002	.006	.00
Common												
Configural	19.63	8	.992	.023	.059							
Metric	30.96	11	.986	.039	.066	21 vs.20	11.33	3	.006	.016	.007	.02
Full scalar	89.32	14	.947	.058	.113	22 vs.21	58.36	3	.039	.019	.047	.00
Partial scalar	31.28	12	.987	.039	.062	23 vs.21	.32	1	.001	.000	.004	.57

Note. df = degrees of freedom; CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; DC = dyadic coping; $\Delta\chi^2$ = chi-square difference test. Items (2, 3, 17, and 18) and two evaluation items (Items 36 and 37) were excluded for the analysis.

loadings, and item intercepts are invariant between Chinese husbands and wives. These findings confirm other DCI validation studies that have examined MI across genders (Randall et al., 2015; Vedes et al., 2013). This indicates that the Chinese version of DCI can be used to compare mean differences between Chinese husbands and wives.

In the subsequent steps, we compared latent group means between samples and found several mean differences between cultural groups. Chinese couples scored lower in DC behaviors such as Emotion- and Problem-Focused Supportive DC, Delegated DC, and Common DC, and higher in Negative DC than Swiss and U.S. couples. Surprisingly, Chinese couples reported using less positive but more negative coping behaviors. These differences may be due to cultural perceptions of how one should behave within the romantic relationship (e.g., different perceptions of intimacy; Marshall, 2008) or the cultural heterogeneity within one cultural group (Quek, Knudson-Martin, Rue, & Alabiso, 2010). For example, Chinese couples and Western couples are supposed to perceive negative or positive DC behaviors asymmetrically due to culturally different expectations (Ebrey, 2006; Oyserman et al., 2002). However, lacking other relationship variables in the current study may remind us to interpret these results with caution.

According to gender differences within Chinese couples, we found that couples exhibited differences in Stress Communication behav-

iors; otherwise, they reported the same amount of DC behaviors. This indicates that both husbands and wives can support each other on a regular basis, which in general might reflect the modern egalitarian relationship mode (Shek, 2006; Quek et al., 2010). Therefore, it is possible that Chinese couples may have shifted to share equal roles in tackling familial affairs (e.g., household chores, jobs, parenting; see Wang et al., 2010), which, in turn, seems to equalize their levels of stress expression and support provision.

Furthermore, we found acceptable to good reliability of the DCI test scores for most subscales with the exception of two subscales where the reliability of test scores was poor (men's Problem-Focused Supportive DC ($\alpha = .51$) by Self and women's Delegated DC by Partner ($\alpha = .52$). Because these two subscales are composed of just two items, results are still plausible (de Vaus, 2002). Future research might focus on exploring the reliability of test scores of these two subscales with different samples.

Construct Validity

This is the first DCI validation study to use an MTMM approach to test for convergent validity, discriminant validity, and method variance. Testing for convergent validity, we found that all convergent validity coefficients (monotrait-heteromethod) were signifi-

Table 5
Model Indices for Measurement Invariance Across the Chinese-U.S. Sample

Invariance types	Goodness-of-fit indices					Contrast	Comparison of nested models					
	χ^2	df	CFI	SRMR	RMSEA		$\Delta\chi^2$	Δdf	ΔCFI	$\Delta SRMR$	$\Delta RMSEA$	p
Men												
DC by Self												
Configural	193.99	110	.965	.037	.045							
Metric	218.26	118	.958	.046	.048	2 vs. 1	24.27	8	.007	.009	.003	.02
Full scalar	328.08	126	.916	.055	.066	3 vs. 2	109.82	8	.042	.009	.018	.00
Partial scalar	240.50	123	.951	.048	.051	4 vs. 2	22.24	5	.007	.002	.003	.00
DC by Partner												
Configural	222.58	110	.959	.044	.053							
Metric	234.48	118	.957	.046	.052	6 vs. 5	11.9	8	.002	.002	.001	.17
Full scalar	287.94	126	.941	.050	.059	7 vs. 6	53.46	8	.016	.004	.007	.00
Partial scalar	253.22	122	.952	.048	.054	8 vs. 6	18.74	4	.005	.002	.002	.01
Common												
Configural	27.94	8	.986	.023	.082							
Metric	47.36	11	.974	.050	.094	10 vs. 9	19.42	3	.012	.027	.012	.00
Full scalar	87.09	14	.947	.061	.118	11 vs.10	39.73	3	.027	.011	.024	.00
Partial scalar	48.24	12	.974	.050	.090	12 vs.10	.88	1	.000	.000	.004	.35
Women												
DC by Self												
Configural	190.01	110	.974	.036	.036							
Metric	233.59	118	.963	.048	.042	14 vs.13	43.58	8	.011	.012	.006	.00
Full scalar	386.24	126	.916	.060	.061	15 vs.14	152.65	8	.047	.012	.019	.00
Partial scalar	267.84	122	.953	.052	.046	16 vs.14	34.25	4	.010	.004	.004	.00
DC by Partner												
Configural	216.44	110	.977	.034	.042							
Metric	231.76	118	.975	.039	.042	18 vs.17	15.32	8	.002	.005	.000	.05
Full scalar	320.30	126	.957	.049	.053	19 vs.18	88.54	8	.018	.010	.011	.00
Partial scalar	269.96	122	.967	.044	.047	20 vs.18	38.20	4	.008	.005	.005	.00
Common												
Configural	21.30	8	.995	.016	.054							
Metric	47.12	11	.986	.040	.076	22 vs.21	25.82	3	.009	.024	.022	.00
Full scalar	72.73	14	.977	.045	.085	23 vs.22	25.61	3	.009	.005	.009	.00

Note. *df* = degrees of freedom; CFI = comparative fit index; SRMR = Standardized Root Mean Square residual; RMSEA = Root Mean Square Error of Approximation; DC = dyadic coping; $\Delta\chi^2$ = chi-square difference test. Items (2, 3, 17, and 18) and two evaluation items (items 36 and 37) were excluded for the analysis.

cant and above the minimal criteria ($r = .25$), confirming convergent validity of the DCI test scores. Nevertheless, convergent validity coefficients were all in a lower to medium range. Such lower correlations can be explained by the way items were constructed (Bodenmann, 2008). The DCI items assess the frequency of coping behaviors over an unspecific time span, which means that participants have to estimate their own and their partners' coping behaviors across situations (and time). Therefore, it is not reasonable to expect high correlations across raters.

According to discriminant validity of test scores, we found that convergent validity coefficients (monotrait-heteromethod) were significantly larger than discriminant validity coefficients (heterotrait-heteromethod) in 75 out of 84 comparisons. This establishes discriminant validity of the DCI test scores as it suggests that each trait assesses a unique and discriminant aspect of DC. Finally, we tested for method variance (i.e., if variability between traits can only be explained by the traits or also based on differences). We found that same-rater coefficients (heterotrait-monomethod) were significant larger than cross-rater coefficients (heterotrait-heteromethod) in 24 out of 42 comparisons, indicating a rater' evaluation bias. However, such evaluation biases are common for many scales (e.g., Anusic, Schimmack, Pinkus, & Lockwood, 2009; Zou et al.,

2013). Altogether, the construct validity of DCI test scores are supported.

Limitations

Limitations of the current study must be addressed, despite its strengths. First, the use of a convenience sample may limit the generalizability of findings to all Chinese couples (and non-Western couples). Utilizing a random stratified sampling technique based on socioeconomic factors (i.e., age, income, and education) would be more beneficial for improving the applicability of the DCI. Second, although couples were mailed the questionnaires packets and were instructed to complete the measures independently, we cannot fully rule out the possibility that one partner completed both sets of questionnaires. Although data from couples with identical answers was excluded in our analyses, future research should utilize methodologies that allow us to control for independent reports (e.g., make use of online assessment tools through giving couples separate accounts). Third, three samples used to test for MI differed in age and education. Although it could have been advantageous to match couples based on age, income, education, one must also consider that income and education may simply be distinct

Table 6

Model Indices for Measurement Invariance Across Genders for the Chinese Sample

Invariance types	Goodness-of-fit indices					Comparison of nested model						
	χ^2	<i>df</i>	CFI	SRMR	RMSEA	Contrast	$\Delta\chi^2$	Δdf	ΔCFI	$\Delta SRMR$	$\Delta RMSEA$	<i>p</i>
DC by Self												
Configural	201.34	80	.956	.047	.053							
Metric	293.23	170	.950	.057	.055	2 vs. 1	91.89	90	.006	.010	.002	.55
Full scalar	299.36	180	.950	.058	.055	3 vs. 2	6.13	10	.000	.001	.000	.00
DC by Partner												
Configural	181.83	80	.951	.047	.050							
Metric	298.20	170	.946	.056	.056	5 vs. 4	116.37	90	.005	.009	.006	.90
Full scalar	304.95	180	.946	.057	.049	6 vs. 5	6.75	10	.000	.001	.007	.61
Common												
Configural	8.52	4	.986	.028	.065							
Metric	16.91	11	.986	.032	.068	8 vs. 7	8.39	7	.000	.004	.003	.09
Full scalar	24.15	14	.985	.036	.058	9 vs. 8	7.24	3	.001	.004	.010	.55

Note. *df* = degrees of freedom; CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; DC = dyadic coping; $\Delta\chi^2$ = chi-square difference test. Items (2, 3, 17, and 18) and two evaluation items (Items 36 and 37) were excluded for the analysis.

among the three countries. Fourth, because the study was based on cross-sectional data and the test–retest reliability could not be computed. Fifth, two subscales with two items each showed lower reliability of test scores. Future research should further examine reliability of test scores for these two subscales. Finally, we found a rater's evaluation bias, which cannot be overcome using a self-report measure. Alternatively, an observational behavioral technique (Raush, Barry, Hertl, & Swain, 1974) could be used to replicate findings of the DCI scale.

Theoretical and Practical Implications

This study has a number of important theoretical and practical implications for future research. First, the established factorial structure supports the theoretical conceptualization of DC behaviors beyond Western cultures (Bodenmann, 2005). Second, the validation of the Chinese version of the DCI may open the door for clinicians and researchers to use the measure to examine stress-coping processes in a collectivistic culture.

Table 7

Descriptive Statistics and Reliability for the Chinese Sample and Group Mean Differences Across all Three Samples

DCI subscales	Comparing latent mean difference (PE)											
	Manifest variables Chinese sample				Across genders							
					Across cultures							
	H		W		H	W	H			W		
	<i>M (SD)</i>	α	<i>M (SD)</i>	α			Chinese	Chinese	Swiss	U.S.	Chinese	Swiss
Self												
Stress Communication	3.43 (.71)	.70	3.21 (.63)	.78	.000	.163	.000	.181	.629	.000	.503	.531
Emotion-Focused Supportive DC	3.46 (.72)	.69	3.53 (.69)	.67	.000	−.017	.000	.336	.514	.000	.341	.571
Problem-Focused Supportive DC	3.40 (.75)	.51	3.55 (.68)	.60	.000	−.102	.000	.190	.211	.000	.093	.272
Delegated DC	3.40 (.71)	.60	3.50 (.70)	.60	.000	−.028	.000	.068	.101	.000	.169	.156
Negative DC	2.40 (.87)	.79	2.40 (.85)	.80	.000	−.137	.000	−.785	−.700	.000	−.186	−.639
Partner												
Stress Communication	3.21 (.70)	.73	3.38 (.65)	.75	.000	−.176	.000	.286	−.230	.000	.364	.198
Emotion-Focused Supportive DC	3.34 (.78)	.70	3.40 (.73)	.72	.000	−.106	.000	.560	.433	.000	.394	.454
Problem-Focused Supportive DC	3.42 (.78)	.63	3.34 (.77)	.62	.000	.029	.000	.136	.073	.000	.226	.130
Delegated DC	3.32 (.80)	.66	3.31 (.83)	.52	.000	−.072	.000	−.011	.216	.000	.187	.176
Negative DC	2.46 (.68)	.79	2.42 (.82)	.75	.000	−.085	.000	−.772	−.341	.000	−.086	−.197
Common												
Emotion-Focused Common DC	3.62 (.68)	.80	3.61 (.72)	.76	.000	.030	.000	.622	.172	.000	.109	.237
Problem-Focused Common DC	3.20 (.90)	.60	3.30 (.82)	.66	.000	−.126	.000	−.584	−.568	.000	.263	.052

Note. All results are based on the final scalar invariance model. Items (2, 3, 17, and 18) and two evaluation items (Items 36 and 37) were excluded for the analysis. In comparison with gender differences, we chose Chinese men as the reference group and a significant difference is indicated in bold ($p < .05$; two-tailed) for women. For comparisons across cultures, we chose the Chinese sample (husbands and wives respectively) as the reference group and if we found a difference within samples from Switzerland or the U.S., the significant estimate is indicated in bold ($p < .05$; two-tailed). DC = dyadic coping; α = Cronbach's alpha; PE = parameter estimates of factor means; *M* = mean; *SD* = standard deviation; H = husbands; W = wives.

Table 8

Three Multitrait-Multimethod Matrices for Husbands' Behaviors, Wives' Behaviors, and Common DC Behaviors

Methods and traits	Husbands' behaviors								
	Husbands' self-report					Wives' partner-report			
	SC	EDC	PDC	DDC	NDC	SC	EDC	PDC	DDC
Husbands' self-report									
SC									
EDC	.31								
PDC	.22	.52							
DDC	.25	.45	.52						
NDC	−.02	−.22	−.13	−.18					
Wives' partner-report									
SC	.36	.17	.12	.18	−.02				
EDC	.21	.29	.16	.15	−.08	.43			
PDC	.22	.20	.35	.20	−.04	.42	.51		
DDC	.20	.20	.20	.27	−.12	.27	.35	.47	
NDC	−.03	−.13	−.07	−.08	.35	−.14	−.21	−.09	−.10
	Wives' behaviors								
Methods and traits	Wives' self-report					Husbands' partner-report			
	SC	EDC	PDC	DDC	NDC	SC	EDC	PDC	DDC
Wives' self-report									
SC									
EDC	.43								
PDC	.39	.60							
DDC	.39	.43	.47						
NDC	−.02	−.25	−.13	−.06					
Husbands' partner-report									
SC	.37	.29	.20	.15	−.14				
EDC	.21	.34	.20	.15	−.08	.49			
PDC	.21	.17	.26	.12	−.04	.37	.59		
DDC	.16	.12	.14	.28	.10	.37	.46	.52	
NDC	−.07	−.08	−.01	.00	.36	−.08	−.21	−.13	−.18
	Common DC behaviors								
Methods and traits	Husbands		Wives						
	(1)	(2)	(3)	(4)					
Husbands									
Emotion-focused CDC (1)									
Problem-focused CDC (2)	.45								
Wives									
Emotion-focused CDC (3)	.37	.17							
Problem-focused CDC (4)	.24	.38	.46						

Note. The convergent validity diagonals (monotrait-heteromethod coefficients) are indicated in bold; all correlation coefficients are significant ($p < .05$; two-tailed). SC = stress communication; EDC = emotion-focused supportive dyadic coping; PDC = problem-focused supportive dyadic coping; DDC = delegated dyadic coping; NDC = negative dyadic coping; DC = dyadic coping; CDC = common dyadic coping.

Third, the Chinese version of the DCI can be used to examine such stress-coping processes in Chinese couples, which could provide a new insight for clinicians who provide therapy and counseling for Chinese couples.

Conclusion

The DCI (Bodenmann, 2008) is a useful measure for relationship researchers and mental health professionals who work with couples trying to manage stress in their relationships. The current study demonstrates that the Chinese version of the DCI is a valid and reliable instrument, which can be used to examine stress and coping behaviors in Chinese couples. Moreover, this measure is a

valuable tool for conducting cross-cultural research between Western and Eastern cultures, which can further inspire additional research on how stress may affect couples' coping behaviors in cross-cultural contexts.

References

- Anusic, I., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2009). The nature and structure of correlations among Big Five ratings: The halo-alpha-beta model. *Journal of Personality and Social Psychology*, 97, 1142-1156. <http://dx.doi.org/10.1037/a0017159>
- Bodenmann, G. (2005). Dyadic coping and its significance for marital

- functioning. In T. A. Revenson, K. Kayser, & G. Bodenmann (Eds.), *Couples coping with stress: Emerging perspectives on dyadic coping* (pp. 33–50). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/11031-002>
- Bodenmann, G. (2008). *Dyadisches Coping Inventar: Testmanual* [Dyadic Coping Inventory: Test manual]. Bern, Switzerland: Huber.
- Bodenmann, G., Hilpert, P., Nussbeck, F. W., & Bradbury, T. N. (2014). Enhancement of couples' communication and dyadic coping by a self-directed approach: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 82, 580–591. <http://dx.doi.org/10.1037/a0036356>
- Bodenmann, G., Meuwly, N., & Kayser, K. (2011). A comparison of two conceptualizations of dyadic coping and their potential for predicting relationship quality and individual well-being. *European Psychologist*, 16, 255–266. <http://dx.doi.org/10.1027/1016-9040/a000068>
- Bodenmann, G., Pihet, S., & Kayser, K. (2006). The relationship between dyadic coping and marital quality: A 2-year longitudinal study. *Journal of Family Psychology*, 20, 485–493. <http://dx.doi.org/10.1037/0893-3200.20.3.485>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Buck, A. A., & Neff, L. A. (2012). Stress spillover in early marriage: The role of self-regulatory depletion. *Journal of Family Psychology*, 26, 698–708. <http://dx.doi.org/10.1037/a0029260>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105. <http://dx.doi.org/10.1037/h0046016>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464–504. http://dx.doi.org/10.1080/1070551070130183410.1207/S15328007SEM0902_5
- Chen, F. M., & Li, T. S. (2007). Marital enqing: An examination of its relationship to spousal contributions, sacrifices, and family stress in Chinese marriages. *The Journal of Social Psychology*, 147, 393–412. <http://dx.doi.org/10.3200/SOCP.147.4.393-412>
- Choi, J. (2008). Work and family demands and life stress among Chinese employees: The mediating effect of work-family conflict. *The International Journal of Human Resource Management*, 19, 878–895. <http://dx.doi.org/10.1080/09585190801993885>
- Davis, H. L. (1971). Measurement of husband-wife influence in consumer purchase decisions. *Journal of Marketing Research*, 8, 305–312. <http://dx.doi.org/10.2307/3149567>
- De Vaus, D. (2002). *Surveys in social research* (5th ed.). London, UK: Routledge.
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, 10, e0121945. <http://dx.doi.org/10.1371/journal.pone.0121945>
- Donato, S., Iafraite, R., Barni, D., Bertoni, A., Bodemann, G., & Gagliardi, S. (2009). Measuring dyadic coping: The factorial structure of Bodenmann's dyadic coping questionnaire in an Italian sample. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, 16, 25–47.
- Ebrey, P. (2006). Confucianism. In D. S. Browning, M. C. Green, & J. W. Jr (Eds.), *Sex, marriage, and family in world religions* (pp. 367–450). New York, NY: Columbia University Press.
- Falconier, M. K., Jackson, J. B., Hilpert, P., & Bodenmann, G. (2015). Dyadic coping and relationship satisfaction: A meta-analysis. *Clinical Psychology Review*, 42, 28–46. <http://dx.doi.org/10.1016/j.cpr.2015.07.002>
- Falconier, M. K., Nussbeck, F., & Bodenmann, G. (2013). Dyadic coping in Latino couples: Validity of the Spanish version of the Dyadic Coping Inventory. *Anxiety, Stress, & Coping*, 26, 447–466. <http://dx.doi.org/10.1080/10615806.2012.699045>
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. 11.0 update (4th ed.). Boston, MA: Allyn & Bacon.
- Groth-Marnat, G. (2009). *Handbook of psychological assessment*. Hoboken, NJ: Wiley.
- Hesketh, T., Lu, L., & Xing, Z. W. (2005). The effect of China's one-child family policy after 25 years. *The New England Journal of Medicine*, 353, 1171–1176. <http://dx.doi.org/10.1056/NEJMp051833>
- 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. <http://dx.doi.org/10.1080/10705519909540118>
- Huang, G. H. C., & Grove, M. (2012). Confucianism and Chinese families: Values and practices in education. *International Journal of Humanities and Social Science*, 2, 10–14.
- Kim, H. S., Sherman, D. K., & Taylor, S. E. (2008). Culture and social support. *American Psychologist*, 63, 518–526. <http://dx.doi.org/10.1037/0003-066X>
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Larzelere, M. M., & Jones, G. N. (2008). Stress and health. *Primary Care: Clinics in Office Practice*, 35, 839–856. <http://dx.doi.org/10.1016/j.pop.2008.07.011>
- Ledermann, T., Bodenmann, G., Gagliardi, S., Charvoz, L., Verardi, S., Rossier, J., . . . Iafraite, R. (2010). Psychometrics of the dyadic coping inventory in three language groups. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie*, 69, 201–212. <http://dx.doi.org/10.1024/1421-0185/a000024>
- Levesque, C., Lafontaine, M.-F., Caron, A., & Fitzpatrick, J. (2014). Validation of the English version of the Dyadic Coping Inventory. *Measurement and Evaluation in Counseling & Development*, 47, 215–225. <http://dx.doi.org/10.1177/0748175614522272>
- Lu, L. (2006). The transition to parenthood: Stress, resources, and gender differences in a Chinese society. *Journal of Community Psychology*, 34, 471–488. <http://dx.doi.org/10.1002/jcop.20110>
- MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychological Methods*, 11, 19–35. <http://dx.doi.org/10.1037/1082-989X.11.1.19>
- Marshall, T. C. (2008). Cultural differences in intimacy: The influence of gender-role ideology and individualism-collectivism. *Journal of Social and Personal Relationships*, 25, 143–168. <http://dx.doi.org/10.1177/0265407507086810>
- McDonald, R. P., & Ho, M. H. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, 7, 64–82. <http://dx.doi.org/10.1037/1082-989X.7.1.64>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, 128, 3–72.
- Papp, L. M., & Witt, N. L. (2010). Romantic partners' individual coping strategies and dyadic coping: Implications for relationship functioning. *Journal of Family Psychology*, 24, 551–559. <http://dx.doi.org/10.1037/a0020836>
- Quek, K., & Knudson-Martin, C. (2006). A push toward equality: Processes among dual career newlywed couples in a collectivist culture. *Journal of Marriage and Family*, 68, 56–69. <http://dx.doi.org/10.1111/j.1741-3737.2006.00233.x>
- Quek, K. M., Knudson-Martin, C., Rue, D., & Alabiso, C. (2010). Relational harmony: A new model of collectivism and gender equality among Chinese American couples. *Journal of Family Issues*, 31, 358–380. <http://dx.doi.org/10.1177/0192513X09351162>
- R, Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>

- Randall, A. K., & Bodenmann, G. (2009). The role of stress on close relationships and marital satisfaction. *Clinical Psychology Review*, 29, 105–115. <http://dx.doi.org/10.1016/j.cpr.2008.10.004>
- Randall, A. K., Hilpert, P., Jimenez-Arista, L., Walsh, K. J., & Bodenmann, G. (2015). Dyadic coping in the U. S.: Psychometrics properties and validity for use of the English version of the Dyadic Coping Inventory. *Current Psychology*. Advance online publication. <http://dx.doi.org/10.1007/s12144-015-9323-0>
- Raush, H. L., Barry, W. A., Hertel, R. K., & Swain, M. A. (1974). *Communication, conflict, and marriage*. San Francisco, CA: Jossey-Bass.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36. <http://dx.doi.org/10.18637/jss.v048.i02>
- Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the fit of the structural equation models: Tests of significance and descriptive goodness-fit measures. *Methods of Psychological Research Online*, 8, 23–74.
- Schwartz, S. J., Benet-Martínez, V., Knight, G. P., Unger, J. B., Zamboanga, B. L., Des Rosiers, S. E., . . . Szapocznik, J. (2014). Effects of language of assessment on the measurement of acculturation: Measurement equivalence and cultural frame switching. *Psychological Assessment*, 26, 100–114. <http://dx.doi.org/10.1037/a0034717>
- Shek, T. D. L. (2006). Chinese family research: Puzzles, paradigms, and policy implications. *Journal of Family Issues*, 27, 275–284. <http://dx.doi.org/10.1177/0192513X05283508>
- Simmons, C. A., & Lehmann, P. (2012). *Tools for strengths-based assessment and evaluation*. New York, NY: Springer.
- Sireci, S. G., Wang, Y., Harter, J., & Ehrlich, E. J. (2006). Evaluating guidelines for test adaptations: A methodological analysis of translation quality. *Journal of Cross-Cultural Psychology*, 37, 557–567. <http://dx.doi.org/10.1177/0022022106290478>
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251. <http://dx.doi.org/10.1037/0033-2909.87.2.245>
- Taylor, S. E., Sherman, D. K., Kim, H. S., Jarcho, J., Takagi, K., & Dunagan, M. S. (2004). Culture and social support: Who seeks it and why? *Journal of Personality and Social Psychology*, 87, 354–362. <http://dx.doi.org/10.1037/0022-3514.87.3.354>
- Thoits, P. A. (2010). Stress and health: Major findings and policy implications. *Journal of Health and Social Behavior*, 51, S41–S53. <http://dx.doi.org/10.1177/0022146510383499>
- Vedes, A., Nussbeck, F. W., Bodenmann, G., Lind, W., & Ferreira, A. (2013). Psychometric Properties and validity of the dyadic coping inventory in Portuguese. *Swiss Journal of Psychology*, 72, 149–157. <http://dx.doi.org/10.1024/1421-0185/a000108>
- Verhofstadt, L. L., Buysse, A., Rosseel, Y., & Peene, O. J. (2006). Confirming the three-factor structure of the quality of relationships inventory within couples. *Psychological Assessment*, 18, 15–21. <http://dx.doi.org/10.1037/1040-3590.18.1.15>
- Wang, Z. Y., Schoebi, D., & Perez, M. (2010). The division of family work in China and Europe: On the role of culture. *Xin Li Ke Xue Jin Zhan* [Advances in Psychological Science], 18, 1668–1678.
- Zhang, M., Foley, S., & Yang, B. Y. (2013). Work-family conflict among Chinese married couples: Testing spillover and crossover effects. *The International Journal of Human Resource Management*, 24, 3213–3231. <http://dx.doi.org/10.1080/09585192.2013.763849>
- Ziegler, M., & Bensch, D. (2013). Lost in translation: Thoughts regarding the translation of existing psychological measures into other languages. *European Journal of Psychological Assessment*, 29, 81–83. <http://dx.doi.org/10.1027/1015-5759/a000167>
- Zou, C., Schimmack, U., & Gere, J. (2013). The validity of well-being measures: A multiple-indicator-multiple-rater model. *Psychological Assessment*, 25, 1247–1254. <http://dx.doi.org/10.1037/a0033902>

Received May 4, 2015

Revision received March 24, 2016

Accepted March 31, 2016 ■