Journal of Studies in Social Sciences
ISSN 2201-4624
Volume 7, Number 1, 2014, 1-18

Psychometric Item Analysis and Validation of Indonesian Version of Intragroup Conflict and Group Atmosphere Scale

Dwi Tyastuti[1,2], Hirotaka Onishi[1], Fika Ekayanti[2] and Kiyoshi Kitamura[1]

[1]International Research Center for Medical Education (IRCME), Graduate School of Medicine, the University of Tokyo, Japan

[2]Faculty of Medicine and Health Sciences, Syarif Hidayatullah State Islamic University, Jakarta, Indonesia

Corresponding author: Dwi Tyastuti, Faculty of Medicine, Central Building, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033

Abstract. Interprofessional education (IPE) is a learning model that brings together a variety of professions. This learning has the potential for the emergence of group conflict. It was needed the tools to evaluate the dynamics of conflict that arose within the group in IPE. The objective of this study was to validate the scales of intragroup conflict (IC) and group atmosphere (GA) in Indonesian version. A survey was conducted at the State Islamic University, Jakarta, Indonesia, in 2011 and we collected 302 valid questionnaires from medical, nursing, pharmacy and public health students. The IC and GA scale were adapted in Indonesian language through the cross-cultural adaptation process. The psychometric properties were analyzed by both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA of IC scale on 11-items and GA scale on 8-items accounted for 65.1% and 62.0%, respectively, of the total variance. The final model of IC scale is 10 items three-factors and GA scale is of 8-items three-factors. Both scales demonstrated the satisfactory reliability, adequate convergent validity and acceptable indices of goodness of fit. The study suggested that the Indonesian version of IC and GA scale are valid to assess the attitudes to manage intragroup conflict during IPE.

Keyword: cross-cultural, validity, reliability, factor analysis, intragroup conflict.

Introduction

Interprofessional education (IPE) was defined as a learning concept that "occurs when two or more professions learn with, from and about each other to improve collaboration and the quality of care" (Barr, 2002). This concept is a learning model that brings together a various of professions, and has the potential for the emergence of a conflict within the group. Conflict is the result of tension between team members because of real perceived differences (Wall & Callistar, 1995; De Dreu & Weingart, 2003).

In medical and health teams, the emerging of conflict among professions is triggered by the unclear or overlapping tasks, roles, and responsibility. However, the conflict can be resolved and become the positive resources in the group if each profession understand their roles and responsibility and skilled in conflict management to accomplish the task group (Greer, Sayqi, Aaldering & de Dreu, 2012).

IPE is a new learning concept in Indonesia. In addition, the studies about the dynamics of conflict in the IPE group are limited. There is tremendous potential for research towards the impact of collaborative interprofessional practice on health team performance. Therefore, we developed a model of IPE to learn the dynamics of conflict that arose in the group during interprofessional practice and it was needed a tool to measure it. We conducted a study of psychometric items and validation scale of intragroup conflict in Indonesian version. The scale is developed by Jehn (1993, 1995) and is a valid and reliable scale. In addition, we also adapted the scale of atmosphere in-group (Chatman, 1991; Shah & Jehn, 1993) to assess students' perception on trust in group, respect to others and open discussion in the group. To our knowledge, there is no information about the scales in Indonesia.

Therefore, the aim of the study is to describe the psychometric properties, reliability and validity of the Indonesian version of intragroup conflict and group atmosphere scale, adapted from the original scale through a cross-cultural process (Guillemin, Bombardier & Beaton, 1993), order to be able to assess the dynamic of conflict in-group among undergraduate students who engage in interprofessional learning.

Methods

Instruments

We validated two scales that included (1) scale of atmosphere in-group (GA) (Chatman, 1991; Jehn, 1995) and Scale of conflict in group (IC) (Shah & Jehn, 1993). Both scales was used to measure the conflict among members in-group and group development. The original study showed that the scales had a good reliability and validity (Shah & Jehn, 1993; Jehn, 1995; Bayerl, 2008). The GA scale was made of eight positive statements and CI scale consisted of 11 negative statements. Participants were asked to indicate their strength of agreement on a 7-point Likert-type scale (1=very strongly disagree, 4=neutral, 7= very strongly agree).

Cross-cultural adaptation process

Cross-cultural adaptation process was applied to adapt the original version of all scales that were used in this study in the Indonesian Language according to guideline by Guillemin (Guillemin et al., 1993). The process was started by forward and backward translations (van de Vijver & Hambleton, 1996). The translation was proficient in both the Indonesian and English languages.

The first phase was the translation of the questionnaire (forward translation) by two language experts. The second phase was aimed to assess the correspondence between the original version and the translated version. The scale in the Indonesian version was re-translated into English (backward translation) by different language experts from the first phase to check for inconsistencies. Researchers evaluated both the original questionnaire and the questionnaire after being translated backward. This was carried out to check the inconsistencies.

Statistic analysis

Evaluation of a scale or questionnaires involved reliability and validity tests of the scale (Bagozzi & al., 1988; Malhotra & Birks, 2007). The reliability

and validity testsof the Indonesian version questionnaires were carried out by applying exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) method.

Criteria for assessment of overall of fit model included (1) loading factor for each items >.7; (2) approximate fit indexes >.90; (3) standard of root mean square residual <.08; (4) all items in regression weights are significant; (5) free from Heywood Case (a negative value of item loading or variance for each items) (Bagozzi & al., 1988; Boomsma, 2000; Hair, 2005; Kline, 2005). For approximate fit indexes, the Normed Fit Index (NFI), Tucker-Lewis fit index (TLI) and the Comparative Fit Index (CFI) exceeds .90 would indicate model fit(Byrne, 2001; Schumacker & Lomax, 2004). The value of RMS (Root Mean Square) represented the square root of the average or mean of the covariance residuals was less than .08 (Browne & Cudeck, 1993)and ideally less than .05 (Steiger, 1990) but the upper confidence interval of the RMS should not exceed .08 (Hu & Bentler, 1998). The relative chi-square is the chi square index divided by degree of freedom which values should be less than 2 or 3 (Kline, 2005; Ullman, 2006).

We tested reliabilities of scales by assessing internal consistency reliability, item reliability and constructs reliability (Fornell & Larcker, 1981; Hair, 2005). Cronbach`s α was performed to measure internal consistency reliability calculated by using SPSS Windows 16.0. Cronbach`s alpha more than 0.07 described a good factor (Hair, 2005), however Cronbach`s or alpha value of 0.06 and more was considered acceptable (Nunnally, 1967; Johnson & Stevens, 2001). Assessment of internal structure model included test of convergent validity and discriminant validity. Convergent validity was a test to assess construct validity being measured that included composite (constructs) reliability (CR) and average variance extracted (AVE) (Fornell & Larcker, 1981; Hair, 2005). If the AVE is less than .50, then the variance due to measurement error is greater than the variance due to the construct. In this case, the convergent validity of the construct is questionable" (Fornell & Larcker, 1981). For composite reliability, a threshold value of 0.50 was set for testing the convergent validity (Bagozzi & al., 1988; Chau, 1997; Hair, 2005).

Discriminant validity was an indicator to assess unidimentional of constructs (Fornell & Larcker, 1981; Chau, 1997; Hair, 2005). Fornell and Larcker (1981) presented a formula for assessing discriminant validity of two or more factors by comparing the AVE of each construct with the shared variance between constructs. If the AVE for each construct is greater than its shared variance with any other construct, discriminant validity is supported (Fornell & Larcker, 1981; Hair, 2005; Farrell, 2010).

To compare the difference in measurement model and to indicate the fit model, there were two indicators for selecting a model: (1) if the measurement model was nested model, then the study employed X^2 difference test (Kahn, 2006) and (2) otherwise, by applying the AIC (Akaike Information Criterion) index which AIC has no such restriction (Anderson, 2008). The final model of Indonesian version was the model that had the lowest AIC score among all measurement models.

All analyses were conducted using SPSS versions 16.0 and AMOS 18.

Participants

To obtain a reliable and valid scale using EFA and CFA, larger samples were needed. There were no consensuses about number of samples among researcher. Some guidelines about number of samples were proposed such as argument from Kline (2005) and Hair (2005) suggested minimum number of samples was 10-20 times the number of parameters (Hair, 2005; Kline, 2005).Thus, using the guideline from Kline (2005) and Hair (2005), this study needed at least 110 participants to test the Indonesian version of IC and GA scale that consisted of 11 and 8 items, respectively. The survey was done from the 20th to 28th February 2011 among first and second-year students from medicine, nursing, pharmacy and public health. We distributed 350 questionnaires and 320 questionnaires were returned.

Data collection

Data was collected by distributing the questionnaire in classroom over a two-week period. A faculty member made a schedule to set the day, time and

length of time required to fill out the questionnaire. Faculty members who were not involved in teaching activities assisted in the distribution and collection of questionnaires.

Ethical consideration

All students gave written and verbal informed consent after receiving both written and verbal information about this study. The Ethics Committees of both the University of Tokyo and the State Islamic University, Syarif Hidayatullah Jakarta, Indonesia, approved the research.

Result

Total of 302 data (94% from 320) were valid and have met the minimum criteria of total sample for EFA and CFA analysis. Excisions of several questionnaires were done in the analysis of this study that was aimed to meet the requirements of analysis either by using EFA or CFA. A total of 18 questionnaires were excluded for the following reasons: (1) 7 questionnaires had missing items and result of the missing value analysis (MVA) showed that MCAR values were >0.05, which means that the missing value is not random, (2) for analyzing outliers univariate, 11 questionnaires had Z-score values of less than -3 and more than +3 (Kline, 2005).

For analysis of data normality, after excluding non-valid questionnaires, total of 302 samples was normal indicated by the value of skewness<3 (between -0.795 to +0.950) and kurtosis <10 (-1.800 to +0.934) (Kline, 2005). The presence of multicollinearity of medium level in the data was indicated by a value of 32 from sample moment calculation (Montgomery, Peck & Vining, 2001), and there was no singularity (r>.95) on those items from the analysis result by AMOS.

Distribution of respondents by gender showed that the number of female students (n=196, 65.3%) was higher than male students (n=104, 34.7%). The number of medical students (61.3%) who participated in this study was greater than those of public health students (13.3%), nursing students (13.3%), and pharmacy students (12.0%). Second year (72.0%) who participated in this study was higher than first year students (28.0%).

The total samples (n=302) were randomly divided into 2 split-half samples by SPSS 16.0, number of samples for each analysis (EFA or CFA) was 151 samples. Data from the sub-sample (n=151) were analyzed by EFA with maximum likelihood methods and promax rotation. The second sub-samples (n=151) were used for the analysis of factor structure using CFA.

Result of factor analysis of IC scale

EFA result of IC scale

All items of the Indonesian version IC scale was included in analysis factor by EFA based on result of the Kaiser-Meyer-Olkin (KMO) Measure Sampling Adequacy for 0.906, Bartlett's Test of Spefericity (α) was significant (<0.001), and the value of the correlation (r) of the anti-image correlation for all items were between 0.766 – 0.930 (standard r>.55). By promax rotation method, the IC scale formed two constructs based on eigenvalue >1 and scree plot graphs.

Unlike the original version (Jehn, 1995), constructs of the Indonesian version consisted of first construct (formed by items 1-5) and second construct (built by items 6 to 11). Only item 8 (*there are differences of opinions in my team*) yielded factor loading < 0.05 (see Table 1). In addition, factor that was accounted for the greatest variance was first factor (58.5%) of the total variance (67.0%).

Factor loading of the first factor ranged from .689 – .891 and the item with the highest factor loading was item 3 (*There are tensions among members of my team*). On the second factor, factor loading of each item on this subscale was between .335 - .839 and item 9 (*There are disagreements about who should do what in my team*) possessed the largest factor loading (.839). Internal consistency test showed that Cronbach's alpha of IC scale was 0.90 (first factor) and 0.84 (second factor).

Table 1.Mean, SD and loading factor of EFA result of IC scale(n=151)

Items	*Mean (SD)*	Subscale 1	Subscale 2
1There is a lot of friction among team members.	4.54 (1.15)	.689	.315
2. There are personality conflicts evident in my team.	4.92 (1.20)	.875	.542
3. There are tensions among members of my team.	5.07 (1.13)	.891	.509
4. There are emotional conflicts among members of my team.	5.18 (1.08)	.840	.500
5. Members in my team disagree with opinions regarding the work being done.	4.95 (0.60)	.731	.568
6. There are conflicts about ideas in my team.	4.64 (1.05)	.689	.612
7. There is a conflict about work in my team.	4.78 (0.99)	.728	.704
8. There are differences of opinions in my team.	3.54 (0.99)	.149	.335
9. There are disagreements about who should do what in my team.	4.34 (1.02)	.593	.839
10. There is a conflict about task responsibilities in my team.	4.61 (1.03)	.720	.800
11. There is a disagreement about resource allocation in my team.	4.49 (0.99)	.620	.831
Total variance explained (67.0%)		58.5	10.5
Reliability of each sub-scale (α total .91)		.90	.84

CFA result of IC scale

Initial model: The initial model was a measurement model derived from the results of EFA. The two-factor model with 11 items (IC-A1: items 1-5; IC-A2: items 6-11) evidenced that t-value of all items was significant ($p < .001$) and all loading factor of items were >0.5 (except item 8), suggesting that item 8 was not represented in the construct being measured. Similarly, measurement model of 11 items three-factor model (referred to the original study; IC-B1, IC-B2, IC-B3) demonstrated unfit model due to the item 8 has low loading factor (0.298). Therefore, the initial model and the 11 items three-factor model were not a fit model for Indonesian context since there was a value of factor loading <0.5 (Hair, 2005). Hence, the modification of the measurement model at the later stage was done by eliminating the item 8.

Table 2. Multilevel CF: reliability, validity and fit indices of initial and modified models (n=151)

Indicator	Reliability-Validity			Index Fit						AIC
	α	CR	AVE	CMIN/df	NFI	TLI	CFI	RMSEA	SRMR	
IC Scale										
1. Initial Model (11 items)										
IC-A1 (item 1 – 5)*	.89	.89	.63	2.29	.91	.93	.95	.09	.05	144.36
IC-A2 (item 6 – 11)	.84	.85	.51							
2. Second Model (11 items)	.91									
IC-B1 (item 1 – 4)	.89	.89	.66							
IC-B2 (item 5 – 8)†	.69	.75	.47	1.979	.93	.95	.96	.08	.05	131.58
IC-B3 (item 9 – 11)	.89	.85	.65							
3. Third Model (10 items)	.92									
IC-C1 (item 1 – 4)	.89	.89	.69							
IC-C2 (item 5 – 7)	.86	.83	.69	1.864	.94	.96	.97	.08	.04	106.76
IC-C3 (item 9 – 11)	.89	.85	.72							
GA Scale										
1. Initial Model (8 items)	.91									
GA-A1 (Item 1 – 8)	.92	.94	.67	3.988	.91	.90	.93	.15	.05	111.76
2. Second Model (8 items)										
GA-B1 (item 1 – 3)	.87	.88	.71	2.855	.97	.97	.98	.08	.02	59.66
GA-B2 (item 4 – 5)	.84	.84	.73							
GA-B3 (item 6 – 8)	.91	.91	.77							

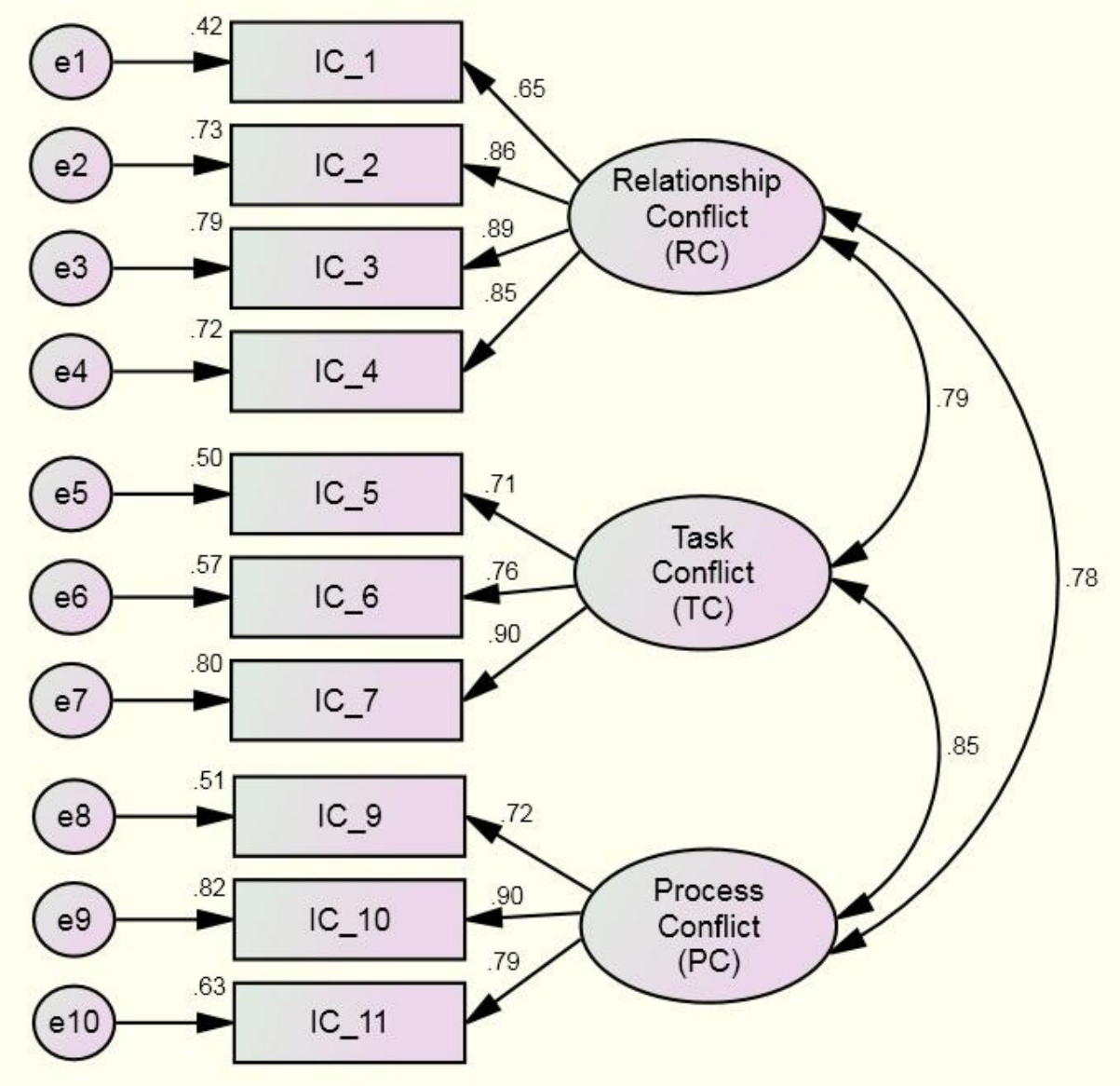
(*). (†)= item 5 of IC scale has loading <0.5

Initial model=based on EFA result; Second model=based on original study; Third model=based on CFA result

Three-factor version, 10 items (Figure 1): The 10 items three-factor model (RC was represented by items 1 to 4; TC consisted of items 5 to 7; PC formed by items 9 to 11) was analyzed and all values of factor loading ranged between .60 and .90. Notwithstanding Hugh (2001) suggested that the best loading factor was >.7 however the loading factor >.5 was also quite satisfactory, especially in the explorative research(Johnson & Stevens, 2001). Therefore, we retained all items (10 items) for further analysis and modification of models. Table 2 showed

that value of the Normed chi-square (32, N = 151) = 59.63) was significant (p<.005). Similarly, the relative chi square (1.864) and GOF indices (TLI=.961, NFI=.943, CFI=.973, RMSEA=0.080 and SRMR=0.039) met the criteria as well as produced AIC score (106.76) lower than other model, indicating that the 10 items three-factor model was a fit model.

Figure 1. Measurement model of 10 items three-factor model of IC scale



IC-C1= Relationship Conflict (RC); IC-C2=Task Conflict (TC);
IC-C3=Process Conflict (PC)

Assessment of internal structure of IC fit model

Reliability analysis pointed out that the overall Cronbach`s alpha for the initial model and modified model was 0.91 and 0.92. By deleting item 8, the overall Cronbach's alpha of modified model was higher than the initial model. Cronbach`s alpha values of 10 items three-factor model ranged from 0.86 to 0.89 (see Table 2). Those results denoted that the modified model (three-factor model) had a high level of internal consistency to measure intragroup conflict. Furthermore, convergent validity of fit model showed that CR (construct reliability) values for each factor was >0.7 and RC factor produced highest CR than other factors. This study showed that the reliability values of all constructs by either Cronbach's alpha or CR were almost similar. AVE (average variance extracted) values for the factors were found as an acceptable level that ranged from 0.69 to 0.72.

Focusing in Table 3, this study showed questionable discriminant validity for factor of TC since AVE of TC (0.69) was lower than square correlation between TC and PC (0.70). Only factor of RC and PC indicated acceptable discriminant validity because the AVE was greater than the squared correlation. The results evidenced that RC and PC were unidimentional.

Table 3. The discriminant validity on three-factor model of IC scale (n=151)

Three factor model	RC	TC	PC
RC	**0.69***		
TC	0.67	**0.69***	
PC	0.59	0.70	**0.72***

Notes: *Diagonal elements report of the AVE, and other matrix entries report the squared correlation estimation between them
RC=relationship conflict; TC=task conflict; PC=professional conflict.

Result of factor analysis of GA scale

EFA result of GA scale

EFA on 8 items was accepted because three measures met the criteria. Firstly, Kaiser-Meyer-Olkin (KMO) measure sampling adequacy was 0.907, above the recommended benchmark of 0.6. Secondly, Bartlett's test of sphericity

(α) was significant(χ^2 (28) = 1643, p < 0.0001). Thirdly, the diagonals of the anti-image correlation matrix for all items were between 0.866-0.938, above the standard of 0.5. By using promax rotation with maximum likelihood, EFA of GA scale produced only one factor (Table 4) and the total variance was 62.0%. The factor loadings for items ranged from 0.628 to 0.837 and the item with the highest factor loading was item 6 (*in my team we discuss issues openly*). As to the internal consistency test, Cronbach's α had a value of .92.

Table 4. Mean, SD and loading factor of EFA result of GA Scale (n=151)

Items		*Mean (SD)*	Subscale 1
1.	Members in my team trust each other.	5.25 (0.84)	.628
2.	Members in my team feel comfortable delegating to others in the team.	5.14 (0.79)	.648
3.	Team members are truthful and honest.	5.25 (0.78)	.787
4.	Team members respect each other.	5.40 (0.82)	.825
5.	Team members respect each others' ideas.	5.40 (0.87)	.768
6.	In my team we discuss issues openly.	5.36 (0.78)	.837
7.	In my team we communicate openly.	5.40 (0.78)	.794
8.	In my team conflicts are dealt with openly.	5.20 (0.87)	.714
	Total variance explained		62.0
	Reliability of each sub-scale		.92

CFA Result of GA scale

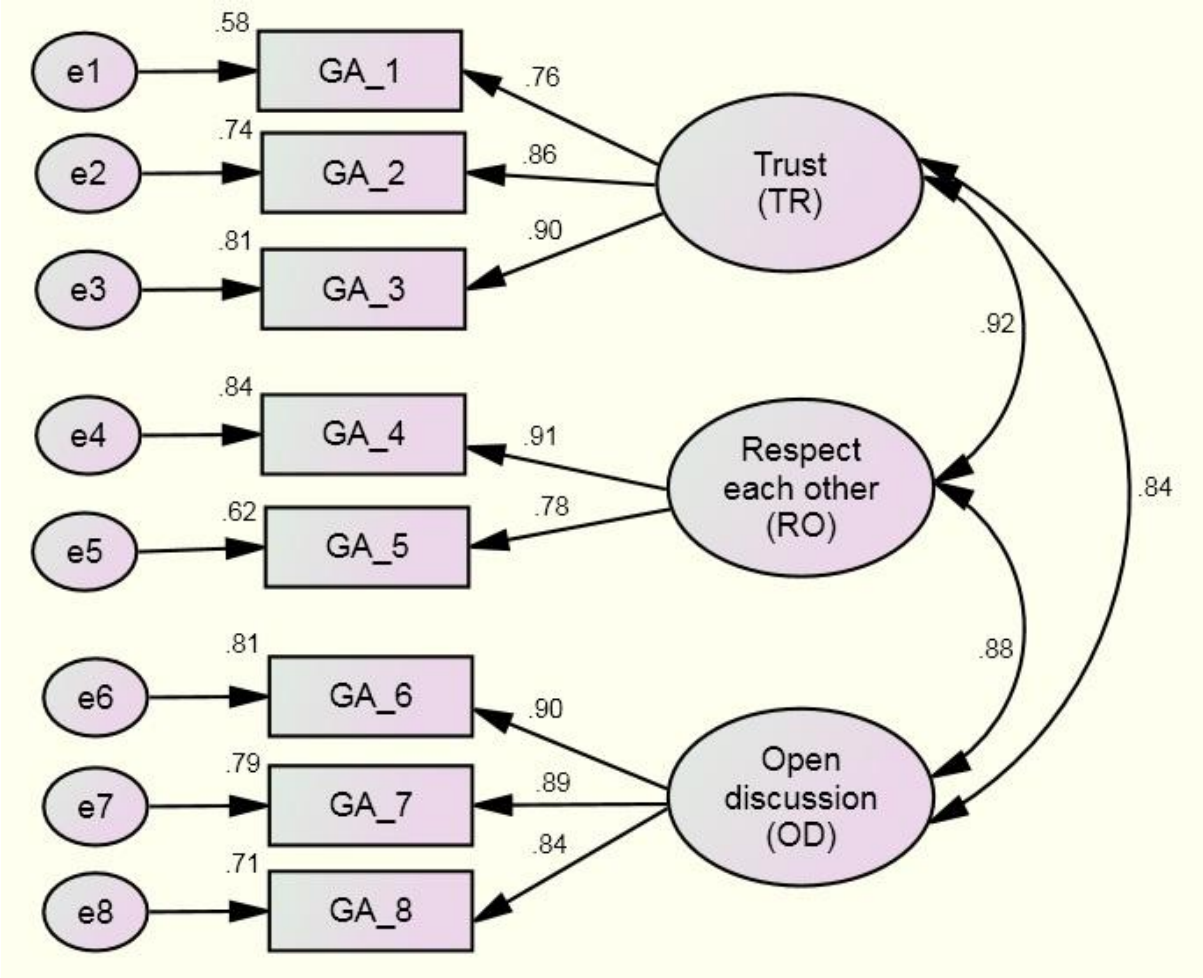
The EFA's analysis was followed by CFA for the Indonesian version to examine reliability and validity of the scale and generate fit model of the Indonesian version.

Initial model: In this study, an initial model (which was consisted of eight items) was a measurement model derived from EFA result. The initial model demonstrated loading factors that met criteria and the indices of goodness of fit index (Normed chi-square, NFI, TLI, RMSEA, SRMR) met the benchmark.

Three-factor version, 8 items (Figure 2): The 8 items three-factor model is similar to the original model. In this study, the 8 items three-factor model (TR

was represented by items 1 to 3; RO consisted of items 4 and 5; OD formed by items 6 to 8) was analyzed and all values of factor loading ranged between .76 and .91. Fit indices showed that all criteria of fit index met the criteria and AIC score of 8 items three-factor model (AIC=59.66) was lower than one-factor model (AIC=111.76), see Table 1.

Figure 2. Measurement model of 8 items three-factor of GA scale



GA-B1= Trust (TR); GA-B2= Respect each other (RO); GA-B3= Open Discussion (OD)

Assessment of internal structure of GA fit model

Reliability analysis pointed out that the overall alpha coefficient of the GA scale model was 0.92, and alpha coefficient for the factors were more than standard. Furthermore, the convergent validity that indicated by AVE and CR

were satisfied. The CR of each construct was satisfied (>0.7) and AVE for all factors met criterion.

Table 5 showed the discriminant validity of three-factor model of GA scale. The study indicated the questionable discriminant validity for all constructs because the AVE of those constructs (TR, RO, and OD) were lower than square correlation between each constructs. The results evidenced that TR, RO and OD were two-dimensional.

Table 5. The discriminant validity on three-factor model of GA scale (n=151)

Three factor model	TR	RO	OD
TR	**0.71***		
RO	0.85	**0.77***	
OD	0.71	0.77	**0.73***

Notes: *Diagonal elements report of the AVE, and other matrix entries report the squared correlation estimation between them.
TR= Trust; RO=respect each other; OD=open discussion

Discussions

This was the first Indonesia study to conduct cross-cultural adaptation and psychometric testing of scales for assessment of attitude to deal with conflict and improve group atmosphere by medical, nursing, pharmacy and public health students, namely the intragroup conflict (IC) scale and group atmosphere (GA) scale. In this study we found that the Indonesian version of the IC and GA scale demonstrated to be a reliable and valid scale.

Through CFA processes, the EFA result of both scales were tested and did not offer the most desirable fit to our data (see Table 2). The item "there are differences of opinions in my team" of IC scale was deleted from this study since factor loading of this item was 0.335, which means that only 34% of variation in this item is explained by IC scale. The reconstructing of IC scale measurement model generated 10 items three-factor model which can be labeled as relationship conflict (RC), task conflict (TC) and process conflict (PC) similar to the original study (Jehn, 1995). Likewise, on GA scale, reconstructing the measurement model by using the original measurement model yielded 8 items

three-factor model which can be labeled as factor of trust (TR), respect to other (RO) and open discussion (OD) refer to the original study (Shah & Jehn, 1993).

Referring to criteria of the fitted CFA model (Brown, 2006), we can conclude that the final CFA with this three-factor model of both scales showed satisfactory data. Absolute fit indices in this study consisted of Normed chi-square, NFI, TLI, CFI, RMSEA, and SRMR, and all generated acceptable results (Bagozzi & Yi, 1988; Hair, 2005). Similarly, the fit model hadthe lowest AIC score among all measurement models (Anderson, 2008). These results indicated that the fit model had a very good fitness to our data.

Internal consistency of fit model of both scales, as determined by the alpha coefficient were adequate in this study (alpha coefficient = 0.92 (IC scale) and 0.91 (GA scale) for the entire sample group. The results presented in Table 2 attested to the high internal consistency of the instrument in which all values were above the suggested 0.70 level for scale robustness (Nunnally & Bernstein, 1994). However, the internal consistency of IC subscales (ranged from .83 to .89) was lower than the original scale (ranged from .93 to .94) (Shah & Jehn, 1993; Jehn, 1995). Convergent validity of fit model of both scales showed that the AVE and CR of all factors yielded acceptable values (Chau, 1997; Farrell, 2010). It means that all factors are well explained by its observed variables.

Nevertheless, we have the discriminant validity issues. Test of discriminant validity of IC scale yielded unidimensional constructs for RC and PC whereas the construct of TC was two-dimensional. While the discriminant validity test of GA scale demonstrated that all factors were two-dimensional. The results indicated that some observed variables correlated more highly with variables outside their factor than with the variables within their factor. This fact showed that a high alpha value does not necessarily indicate that a factor is unidimensional. The alpha coefficient is not sufficient for measuring the dimensionality of a construct or factor (Tavako & Dennick, 2011). This is evidenced by factor of TC, TR, RO and OD in which the alpha coefficients of those factors were high but those factors were multidimesional.

Several factors contributed to the emergence of discriminant validity problems and low factor loading. In this study, the existence of cultural diversity

among respondents such as gender, ethnic, belief and local cultures (Pashaei, Razaghi, Foroushani & Tabatabaei, 2013) as well as students' experiences in working in group (De Dreu & Weingart, 2003) may have influenced the responses to the items on Indonesian version. In addition, the effect of translation into a particular language played a significant role in this study and the role of cultural differences must be considered (Pashaei et al., 2013).

Although the fit model of IC and GA scale of Indonesian versions are favorable, there are several limitations. First, our sample was confined to a school of health professionals in Indonesia that might not represent all Indonesian students. The findings may be difficult to generalize because the sample was only derived from one institution. Second, the sampling method was non-probability, which may produce sampling bias. Nonetheless, the sample represented students from various multicultural and social backgrounds, which is reassuring. Third, this psychometric study in Indonesian context resulted the questionable discriminant validity of adapted scale.

Conclusion

The present study contributes to the literature in a couple of ways. First, the translation and adaptation of intragroup scale and group atmosphere scale into the Indonesian language by a cross-cultural adaptation process was successful and the Indonesian version produced the CFA fit model with the highest reliability, the satisfied convergent validity and the acceptable goodness of fit index. Second, an item of Indonesian version intragroup conflict scale exhibited misfit measurement within the model and was excluded. Consequently, further revision and assessment of the intragroup conflict psychometric properties in the Indonesian version is recommended.

Acknowledgement

We would like to thank all students at Faculty of Medicine and Health Sciences (FMHS), Syarif Hidayatullah State Islamic University, Jakarta, Indonesia who participated in this study.

References

[1] Anderson, D. R. (2008). Model Based Inference in the Life Sciences. Colorado State University Fort Collins, CO, Springer Science.

[2] Bagozzi, R. & al., e. (1988). On the evaluation of structural equation models. Journal of the Academy of Marketing Science 16: 74-94.

[3] Bagozzi, R. P. & Yi, Y. (1988). On the evaluation of structural equation models. Journal of the Academy of Marketing Science 16: 74-94.

[4] Barr, H. (2002). Interprofessional education : Today, yesterday and tomorrow. Learning and Teaching support Network: Centre for Health Sciences and Practice.

[5] Bayerl, P. S. (2008). Co-learning in health care: Evaluation of the interprofessional training program at the Bedlam Longitudinal Clinic.

[6] Boomsma, A. (2000). Reporting analyses of covariance structures. Structural Equation Modelling 7(3): 461-483.

[7] Brown, T. A. (2006). Confirmatory factor analysis for applied research. New York, The Guilford Press.

[8] Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. Testing Structural Equation Models. Beverly Hills, CA, Sage**:** 136-162.

[9] Byrne, B. M. (2001). Strutural equation modeling with AMOS, EQS, and LISREL: comparative approaches to testing for the factorial validity of a measuring instrument. International Journal of Testing 1(1): 55-86.

[10] Chatman, J. (1991). Matching people and organizations: Selection and socialization in public accounting firms. Adminstrative Science Quarterly 36: 459-484.

[11] Chau, P. Y. K. (1997). Reexaming a model for evaluating information center succes using a structural equation modelling approach. Decision Sciences 28(2): 309-334.

[12] De Dreu, C. K. W. & Weingart, L. R. (2003). Task Versus Relationship Conflict, Team Performance, and Team Member Satisfaction: A Meta-Analysis. Journal of Applied Psychology 88(4): 741-749.

[13] Farrell, A. M. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). Journal of Business Research 63: 324-327.

[14] Fornell, C. & Larcker, D. F. (1981). Evaluation of Structural Equation Model with Unobservables Variable and Measurement Error. Journal of Marekting Research 18(1): 39-50.

[15] Greer, L. L., Sayqi, O., Aaldering, H. & de Dreu, C. (2012). Conflict in medical teams: opportunity or danger. Medical Education 46(10): 935-942.

[16] Guillemin, F., Bombardier, C. & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 46(12): 1417-1432.

[17] Hair, J. F. (2005). Multivariate data analysis. Upper Saddle River, NJ, Prentice Hall.

[18] Hu, L.-t. & Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification. Psychological Methods 3(4): 424-453.

[19] Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. Administrative Science Quarterly 40(2): 256-282.

[20] Johnson, B. & Stevens, J. J. (2001). Exploratory and confirmatory factor analysis of the school level environment questionnaire (SLEQ). Learning Environment Research 4: 325-344.

[21] Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice principles, advances, and applications. The Counseling Psychologist 34(5): 684-718.

[22] Kline, R. B. (2005). Principles and practice of structural equation modeling. New York, Guilford.

[23] Malhotra, N. K. & Birks, D. F. (2007). Marketing Research: an applied orientation. Essex, UK, Person Education International/Prentice Hall.

[24] Montgomery, D. C., Peck, E. A. & Vining, G. G. (2001). Introduction to linear regression analysis. New York, Wiley.

[25] Nunnally, J. C. (1967). Psychometric theory. New York, McGraw-Hill.

[26] Nunnally, J. C. & Bernstein, I. H. (1994). Psychometric theory. New York, McGraw-Hill.

[27] Pashaei, T., Razaghi, O. M., Foroushani, A. R. & Tabatabaei, M. G. (2013). Assessing the validity and reliability of the Farsi version of inventory drug-taking situations. Iran J Psychiatry 8(2): 80-85.

[28] Schumacker, R. E. & Lomax, R. G. (2004). A Beginner's Guide to Structural Equation Modeling. Mahwah, NJ, Lawrence Erlbaum Associates, Inc.

[29] Shah, P. P. & Jehn, K. A. (1993). The interaction of friendship, conflict, and task. Group Decision and Negotiation 2: 149-165.

[30] Steiger, J. H. (1990). Structural Model Evaluation and Modification : An Internal Estimation Approach. Multivariate Behavioral Research 25(2): 173-180.

[31] Tavako, M. & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education 2: 53-33.

[32] Ullman, J. B. (2006). Structural equation modeling: reviewing the basics and moving forward. Journal of Personality Assessment 87(1): 35-50.

[33] van de Vijver, F. J. R. & Hambleton, R. K. (1996). Translating tests: Some practical guidelines European Psychologist 1(2): 89-99.

[34] Wall, J. A. & Callistar, R. R. (1995). Conflict and its Management. Journal of Management 21: 515-558.