
An Empirical Cross-Cultural Study of Moral Judgment Development in Mainland China

Qian Zhang

(Guangdong University of Foreign Studies, gwzhangqian@126.com)

Stephen J. Thoma

(University of Alabama, sthoma@bamaed.ua.edu)

1. Introduction

In the study of moral judgment development, Kohlberg's Moral Judgment Interview (MJI) has been widely used, whose validity and reliability have been demonstrated to be good (Colby & Kohlberg 1987). However, MJI was criticized by researchers for its complicated scoring system (Rest, Narvaez, Bebeau, & Thoma 1999): The 1200-page Scoring Criteria makes it hard for researchers to master, and the one-by-one interview makes the test difficult to be widely applied. What's more, the productive evaluation in MJI only focuses on what is explicitly expressed by the interviewee, neglecting the invisible part that the interviewee fails to express by words (Rest, Narvaez, Bebeau & Thoma 1999).

Based on Kohlberg's theory of moral judgment and aiming at solving the problems of MJI, Rest constructed an objective measure of moral judgment, Defining Issues Test (DIT) in 1970s which develops into DIT2 in 1990s (Rest & Narvaez, 1998; Rest, Narvaez, & Thoma 1999).

Different from the production task of MJI in which one's moral judgment competency is evaluated by his oral expression of his attitude towards a moral dilemma, DIT2 is an objective recognition task in which one's moral judgment competency is evaluated by his rating and ranking of certain items about a moral dilemma (Rest, Narvaez, & Thoma 1999). DIT2 presents five moral dilemma stories for which the participants are asked to make an action choice, rate 12 items in terms of importance in helping them to make that action decision, and then rank the four most important items from above and put them in order.

From the ratings and rankings, three important developmental indices are calculated: Personal Interest schema focuses one's personal welfare or benefits of family and close friends, which is the lowest level of moral reasoning measured by DIT2; Maintaining Norms schema takes a further step from Personal Interest by focusing on people's conformation to the laws and principles in the society, which is the medium level of moral reasoning measured by DIT2; Postconventional Thinking schema represents the ability to consider an action decision from the perspective of intuitively appealing ideals, which is the highest level of moral reasoning.

Through the calculation of rating and ranking, the researcher works out the participant's score in each schema. Higher scores in each schema indicate that the participant takes the items which represent the designated schema as particularly important. The researcher uses N2 score in DIT2 to show the participant's emphasis of a more advanced thinking (Postconventional) and de-emphasis of a lower thinking (Personal Interest) (Thoma 2006). Basically speaking, the higher one scores in Postconventional Thinking, the lower he will score in Personal Interest. Therefore higher N2 scores indicate a participant's more advanced moral reasoning. The score range for each schema is from 0 to 95. (Rest, Narvaez, Bebeau, & Thoma 1999).

As a good measurement tool in America (Davison, Robbins, & Swanson 1978; Martin, Shafto, & VanDeise 1977; Rest 1979), DIT was introduced into China in 1980s (Xinyin 1988; Haigen & Boshu 1997). The research in mainland China is almost in theoretical studies (Xinyin 1988; Haigen & Boshu 1997; Shaogang & Huihong 2004; Shaogang 2006). Some scholars from Hong Kong and Taiwan did some empirical research, but all their research used DIT1 (Hing-Keung & Wing-Shing 1987; Hing-Keung 1988; Hau-siu & Daniel 2002; Yi-Hui & Chieh-Yu 2008). The current research is the first comparative empirical study with DIT2 in mainland China. In this study, 113 Chinese participants take part in the DIT2 test and their scores are calculated by the Center of Ethical Studies in University of Alabama. Research based on the result of Chinese participants is analyzed and compared with American equivalents in this paper.

There are two goals to the research presented here: (a) To test the pattern of DIT2 with

Chinese participants, and (b) To investigate the differences between Chinese participants and American counterparts.

2. Method

2.1. Participants

Participants include 113 individuals from various regions in China, with 57 males and 56 females. The average age is 34.7. Within the sample of the highest education level, 6 participants were from middle school (5.4%), 7 from high school (6.3%), 81 getting bachelor degree (72.3%), and 18 getting graduate degree (16.1%). All the participants have been working between 5 years to 25 years after getting their highest diploma.

2.2. Materials and Procedure

The current research uses the Chinese version of DIT2. Translating DIT2 into Chinese was done by the writer, with the guidance of Dr. Stephen J. Thoma, the major developer of DIT1 and DIT2. After the translation, two college English teachers in China were invited to review the Chinese version and the original

English version to see whether there is any discrepancy in understanding on the language level.

Moreover, another two Chinese who are sufficient in both American culture and Chinese culture were invited to take the test with both Chinese and original English versions. These two share similar background, both being born and receiving education in China until getting their bachelor degree, both getting master and PhD degree in America and staying in America for more than 8 years (one is 8 years and another is 14 years), both getting a teaching job in American university after graduation. They were invited to review the two versions to guarantee there is no discrepancy in the understanding on the cultural level.

Participants take the Chinese version of DIT2 in a natural condition and the scores were compared with the online mega sample of American participants from 2011 to 2014. Differences in Postconvention (P), Maintaining Norms (MN), Personal Interests (PI), and N2 scores were investigated with regard to the following variables: Sex, Education Level, Religious Orthodoxy and Humanitarian/Liberalism.

3. Results

3.1. Differences in P, MN, PI, and N2 Scores by Religious Orthodoxy

One-way ANOVA is conducted to investigate differences in moral judgment developmental indices based on Religious Orthodoxy. Different from the result of mega sample of American participants which shows there is a significantly positive relationship between Religious Orthodoxy and moral reasoning (see Table 1) , the result in the current research reveals that with Chinese participants there is no significant relationship within except P score (see Table 2).

	Sum of Square	df	Mean Square	F	Sig.
Post Conventional (P score)					

Between group	205593.965	12	17132.830	69.726	.000
Within group	3309329.268	13468	245.718		
Sum	3514923.233	13468			
Personal Interest (Stage 2/3)					
Between group	20907.697	12	1742.308	11.157	.000
Within group	2103255.421	13468	156.167		
Sum	2124163.118	13468			
Maintain Norms (Stage 4) Between					
group	246662.603	12	20555.217	113.967	.000
Within group	2429095.177	13468	180.360		
Sum	2675757.780	13468			
N2SCORE					
Between group	147712.004	12	12309.334	53.402	.000
Within group	3104430.931	13468	230.504		
Sum	3252142.934	13468			

Table 1. Correlation of Religious Orthodoxy with the DIT2 Indices among American participants

	Sum of Square	df	Mean Square	F	Sig.
Post Conventional (P score)					
Between group	2373.789	7	339.113	2.662	.014
Within group	13377.216	105	127.402		
Sum	15751.005	112			
Personal Interest (Stage 2/3) Between					
group	647.476	7	92.497	.788	.599
Within group	12321.752	105	117.350		
Sum	12969.228	112			
Maintain Norms (Stage 4) Between					
group	1062.159	7	151.737	1.406	.211
Within group	11335.538	105	107.958		
Sum	12397.698	112			
N2SCORE					
Between group	1484.687	7	212.098	2.014	.060
Within group	11059.658	105	105.330		
Sum	12544.345	112			

Table 2. Correlation of Religious Orthodoxy with the DIT2 Indices among Chinese participants

3.2. Differences in P, MN, PI, and N2 Scores by Humanitarian/Liberalism

One-way ANOVA was conducted to investigate differences in moral judgment developmental indices based on Humanitarian/Liberalism. The result of the mega sample of American participants reveals that Humanitarian/Liberalism has significant relationship with all the moral judgment developmental indices. The higher Personal Interest score/Postconventional score/N2 score is, the

more pronounced Humanitarian/Liberalism is, while at the same time the lower Maintaining Norms score is.

However, the result in the present research doesn't show exactly the same trend with the Chinese participants. As shown in Table 3, there is significant relationship between Humanitarian/Liberalism and Personal Interest score. The higher Personal Interest score is, and the lower Maintaining Norms score is, the more pronounced Humanitarian/Liberalism is (see Figure 1 & 2). But it shows no significant relationship between Humanitarian/Liberalism with N2 score and P score. (see Figure 3 & 4)

	Sum of Square	df	Mean Square	F	Sig.
Post Conventional (P score)					
Between group	654.572	5	130.914	.928	.466
Within group	15096.433	107	141.088		
Sum	15751.005	112			
Personal Interest (Stage 2/3)					
Between group	1671.674	5	334.335	3.167	.010
Within group	11297.554	107	105.585		
Sum	12969.228	112			
Maintain Norms (Stage 4)					
Between group	703.737	5	140.747	1.288	.275
Within group	11693.960	107	109.289		
Sum	12397.698	112			
N2SCORE					
Between group	794.240	5	158.848	1.447	.214
Within group	11750.105	107	109.818		
Sum	12544.345	112			

Table 3. Correlation of Humanitarian with the DIT2 Indices among Chinese Participants

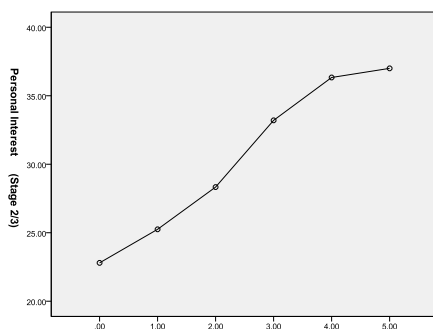


Figure 1. PI Scores by Humanitarian Liberalism

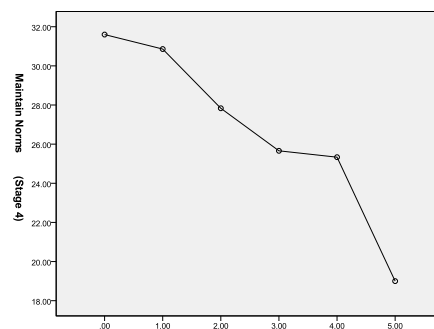


Figure 2. MN Scores by Humanitarian Liberalism

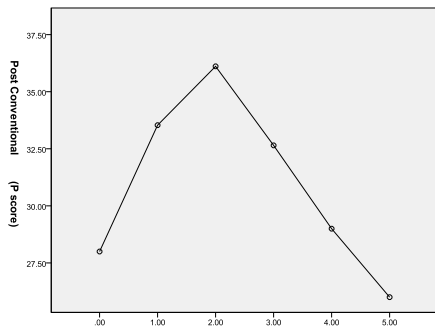


Figure 3. P Scores by Humanitarian Liberalism

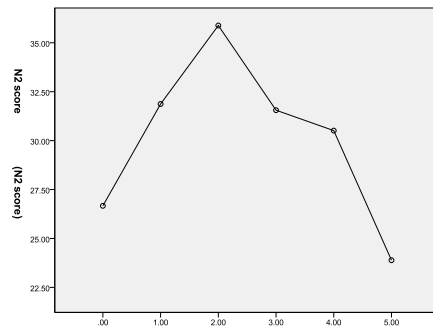


Figure 4. N2 Scores by Humanitarian Liberalism

3.3. Differences in P, MN, PI, and N2 Scores by Sex

Descriptive statistics analysis was conducted to investigate the differences in moral judgment developmental indices between males and females. The result shows the same pattern with that of American participants, that is, females in this research also outscore males, but the magnitude of the difference is quite small (see Table 4).

Sex	N	N2		P		MN		PI	
		M	SD	M	SD	M	SD	M	SD
Male	57	32.60	9.5	33.3	10.7	28.4	10.4	29.1	9.9
Female	56	32.62	11.6	33.6	13.0	27.4	10.7	29.4	11.7

Table 4. Descriptive Statistics by Sex among Chinese participants

3.4. Differences in P, MN, PI, and N2 Scores by Education

One-way ANOVA was conducted to investigate differences in moral judgment developmental indices based on education. Result of mega sample in America shows that the higher education level is, the more advanced moral reasoning will be (see Table 5). Descriptive statistical analysis in the current research shows the same trend as with the growth of educational level, there is a declining tendency in Personal Interest score and a rising tendency in Postconventional score and N2 score (see Table 6).

Education Level	N	N2		P		MN		PI	
		M	SD	M	SD	M	SD	M	SD
Grade 7-9	576	36.3	14.6	36.3	15.7	32.7	13.8	25.3	12.3
Grade 10-12	546	41.9	15.1	43.3	16.4	28.0	14.0	23.6	11.8
Seniors	2458	37.4	15.6	37.4	16.3	33.3	14.2	23.9	12.6
MS Degree	788	33.0	14.7	34.6	14.6	31.8	13.6	27.5	12.3

Table 5. DIT Means and Standard Deviations for Schema Scores and N2 Scores by Education Level for American participants from 2010-2014

Education Level	N	N2		P		MN		PI	
		M	SD	M	SD	M	SD	M	SD
Grade7-9	6	29.7	8.2	32.0	10.8	25.0	11.8	28.7	11.0
Grade10-12	7	27.4	9.0	29.1	8.6	30.0	7.7	33.4	12.1
Seniors	81	2.5	10.7	33.3	12.4	28.2	10.6	29.4	11.0
MS Degree	18	36.3	11.0	36.9	10.8	26.6	11.4	26.9	9.7

Table 6. DIT Means and Standard Deviations for Schema Scores and N2 Scores by Education Level for Chinese participants

3.5. Differences in P, MN, PI and N2 Scores from American participants

A comparison in moral judgment developmental indices between Chinese participants and American equivalents is conducted, which shows that Chinese participants score lower in P items by 9% and N2 by 10% than their American counterparts by education (see Table 7). The Chinese participants score higher by 14.7% in Personal Interest, while they score 17.9% lower in Maintaining Norms than the American counterparts.

	Chinese Participants		American Participants	
	Mean	Deviation	Mean	Deviation
PI	29.2	10.8	24.9	12.6
MN	27.9	10.5	32.9	14.1
P	33.5	11.9	36.7	16.1
N2	32.6	10.6	36.0	15.5

Table 7. A Comparison between Chinese participants and Americans in Mean and Standard Deviation

3.6. Differences in Meaningless Items frequency from American participants

Another interesting finding in this research lies in the difference in the choice of meaningless items between Chinese participants and American counterparts.

In DIT2 test, meaningless items refer to those that are lofty sounding, using complex style or verbiage, but essentially without any meaning (Bebeau & Thoma 2003). To investigate participants' choice of meaningless items is a good way to check whether they get the same understanding towards the same items in the two versions of DIT2.

A comparison between Chinese participants and American counterparts clearly shows that among the five meaningless items in the DIT2 test (f4, f8, r9, s4, c5), American participants have the higher percentage to correctly identify them as meaningless items (rate them as "No") than Chinese participants.

Firstly, American participants can identify more meaningless items than the Chinese counterparts. This difference is revealed by the comparison between the percentage of rating the meaningless item as "No" and that of rating it as "Great" by the participants from America and China. American

participants identify four out of five meaningless items (f4, r9, s4, c5), while Chinese participants identify three (f4, r9, c5).

Secondly, American participants distinguish the meaningful items more obviously than the Chinese counterparts. This difference is revealed by the distance between the correct rating (rating it as “No”) and the wrong rating (rating it as “Great”) of a meaningless item.

With regard to American participants, item f8 is the only one in which the percentage of wrong rating exceeds the percentage of correct rating. In this item, 18.7% participants rate it as “Great” while only 9.9% rate it as “No”, nearly twice of the correct rating. But as for the rest of the four items (f4, r9, s4, c5) in which the percentage of correct rating exceeds that of wrong rating, the percentage of correct rating is extraordinarily higher than that of wrong rating (see Table 8).

	Great	No
F4	3.2%	63.2%
F8	18.7%	9.9%
R9	4.2%	34.7%
S4	6.7%	27.4%
C5	4.3%	40.6%

Table 8. A comparison of percentage of frequency in rating meaningless items as “Great” and “No” among American participants

In contrast, with Chinese participants, the result is quite reverse. When they make the right identification and the correct rating exceeds the wrong rating (item f4, r9 and c5), the distance of the percentage between rating them as meaningless and rating them as most important is not very big for two items as item f4 and item r9 (23.9% vs. 10.6%; 15.9% vs. 7.1%);, while with items in which the wrong rating is higher than the right rating (item f8 and s4), there shows a very big distance (20.4% vs. 5.3%; 31.9% vs. 3.5%) (see Table 9).

	Great	No
F4	10.6%	23.9%
F8	20.4%	5.3%
R9	7.1%	15.9%
S4	31.9%	3.5%
C5	4.4%	30.1%

Table 9. A comparison of percentage of frequency in rating meaningless items as “Great” and “No” among Chinese participants

This comparison implies that Chinese participants have difficulty in identifying the meaningless items in the western culture in DIT2.

3.7. Differences in preference stages between Chinese and American participants

A comparison in the frequency of the choice of each item between American participants and Chinese participants shows the items in which the difference between these two groups of participants exceeds 5% and 10% respectively (see Table 10 & 11). By grouping those items by stages, we find out American participants favour stage 4 items most while Chinese participants favour stage 3 most (see Table 12 & 13). This finding is in correlation with the finding that Chinese participants score higher in PI score and lower in MN score than their American counterparts earlier in this paper.

Pref item_Ch	Pref stage_Ch	Pref item_Am	Pref stage_Am
F12	5	F2	2
R8	3	F11	5
C4	3	R3	2
C11	4	R11	4
D6	3	S3	2
F1	2	S8	4
F7	6	S12	5
F9	3	C1	4
R7	4	C6	5
C3	2	C10	4
C5	4	D1	4
C8	4	D11	6
D3	2	D10	4
		R1	4
		S1	4
		S7	5
		S11	3
		C9	3
		D4	4
		D8	5
		D9	6
		D12	4

Table 10. More than 5% in difference in items and stages between American participants and Chinese counterparts

Pref item_Ch	Pref stage_Ch	Pref item_Am	Pref stage_Am
F12	5	F2	2
R8	3	F11	5
C4	3	R3	2
C11	4	R11	4
D6	3	S3	2
		S8	4
		S12	5
		C1	4
		C6	5
		C10	4
		D1	4
		D11	6
		D10	4

Table 11. More than 10% in difference in items and stages between American participants and Chinese counterparts

	Frequency	Percentage	Valid Percentage	Total Percentage
Valid 2.00	6	17.1	17.1	17.1
3.00	2	5.7	5.7	22.9
4.00	16	45.7	45.7	68.6
5.00	8	22.9	22.9	91.4
6.00	3	8.6	8.6	100.0
Total	35	100.0	100.0	

Table 12. Frequency of Preference Stage among American Participants

	Frequency	Percentage	Valid Percentage	Total Percentage
Valid 2.00	3	8.6	16.7	16.7
3.00	7	20.0	38.9	55.6
4.00	5	14.3	27.8	83.3
5.00	2	5.7	11.1	94.4
6.00	1	2.9	5.6	100.0
Total	35	100.0	100.0	

Table 13. Frequency of Preference Stage among Chinese Participants

3.8. Differences in types between Chinese and American participants

In DIT2, type indicator is an important index based on schema preference and whether the profile is consolidated or transitional (Bebeau & Thoma 2003) to indicate the different reasoning type of the participant. There are seven types. Type 1 means that the participant is dominant in personal interests schema and consolidated; type 2 means that personal interest still dominates the participant but there shows a transition to maintaining norms schema. Type 3 refers to a dominance in maintaining norms schema with personal interest schema in a secondary place. Type 4 refers to a consolidated dominance in maintaining norms schema. Type 5 indicates a dominance in maintaining norms schema but a tendency to transit to postconventional schema. Type 6 refers to a predominance in postconventional schema with maintaining norms

schema in a secondary place. Type 7 refers to consolidated predominance in postconventional schema.

From the characteristics of each type, we can roughly match the seven types with schemas. In type 1 and 2, either in consolidated or transitional state, personal interest is in dominant place in moral reasoning process, therefore, we can regard participants in these two types to use personal interest schema and match these two types with Personal Interest schema. Type 3, 4 and 5 take maintaining norms as most important, therefore, they can be matched with Maintaining Norms schema. Type 6 and 7 have postconventional as dominant in moral reasoning, therefore, they can be matched into Postconventional schema.

Descriptive statistics analysis was conducted to compare the frequency of type indicator between Chinese and American participants (see Table 14 & 15).

	Frequency	Percentage	Valid Percentage	Schema Type	Percentage
Type 1	1	.9	.9	PI	31
Type 2	34	30.1	30.1		
Type 3	17	15.0	15.0	MN	30
Type 4	4	3.5	3.5		
Type 5	13	11.5	11.5		
Type 6	30	26.5	26.5	P	38.9
Type 7	14	12.4	12.4		
Total	113	100.0	100.0		

Table 14. Frequency of Type Indicator among Chinese Participants

	Frequency	Percentage	Valid Percentage	Schema Type	Percentage
Type 1	256	1.9	1.9	PI	17.7
Type 2	2131	15.8	15.8		
Type 3	1652	12.3	12.3	MN	35.3
Type 4	1836	13.6	13.6		
Type 5	1269	9.4	9.4		
Type 6	2524	18.7	18.7	P	47
Type 7	3810	28.3	28.3		
Total	13478	100.0	100.0		

Table 15. Frequency of Type Indicator among American Participants

From the two tables, we can clearly see that American participants' scores show an obvious climbing tendency in frequency of type indicators in terms of schema. They score lowest in personal interest schema (Type 1 & 2), higher in maintaining norms schema (Type 3, 4, & 5), highest in postconventional schema (Type 6 & 7). On the contrary, Chinese participants show an even tendency. They score almost the same in these three schemas,

with personal interest (Type 1 & 2) slightly higher than maintaining norms schema (Type 3, 4, and 5) and postconventional schema scores highest with slight advantage (Type 6 & 7) (see Figure 5 and 6).

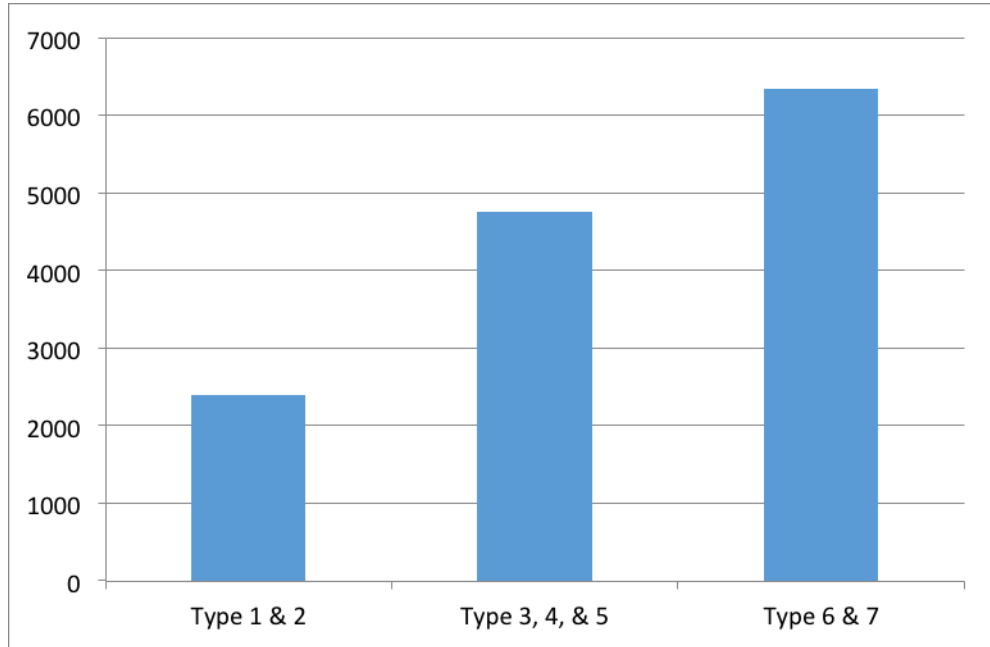


Figure 5. Percentage of Type Indicator among American Participants

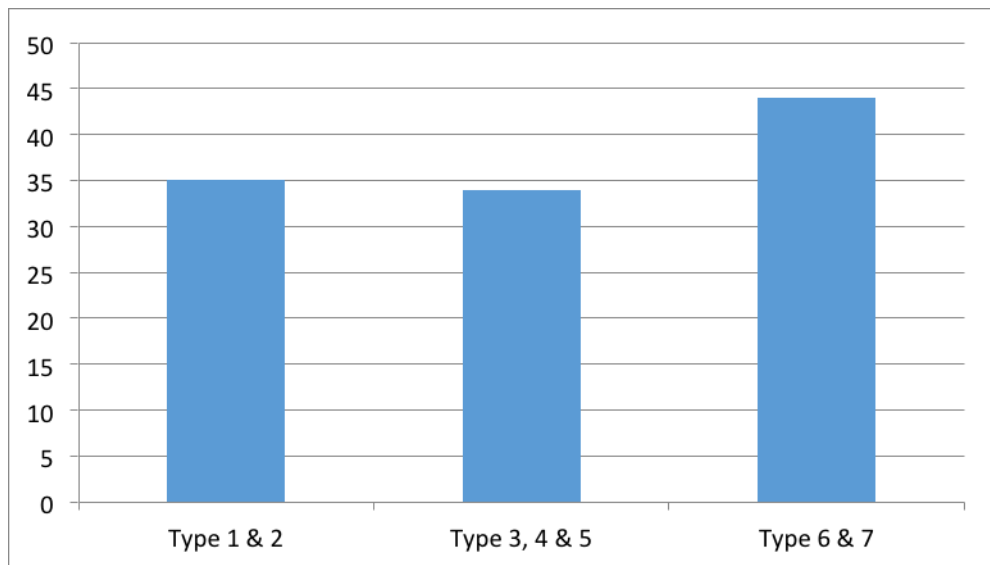


Figure 6. Percentage of Type Indicator among Chinese Participants

This comparison supports the earlier finding with the stage preference in this paper that Chinese participants have a more favor in stage 4 than the American participants.

4. Discussion

The present study tests the general pattern of DIT2 with the participants in China, and compares the differences between Chinese participants and American counterparts with regard to the variables.

Firstly, the results of the current research show the same trend with the general pattern of DIT2 in the following:

1. There is a significant relationship between education and moral judgment developmental index scores. The higher education one has, the higher N2 score and P score he has, and the lower score in Personal Interest he has.

2. There is a significant relationship between sex and moral judgment developmental index scores. Females show higher P score and N2 score than males by 10%, which is consistent with the research result of mega sample of Americans.

Secondly, the result of the current research shows that Chinese participants have a different pattern from the American equivalents as follows:

1. There shows no significant relationship between religious orthodoxy and moral judgment developmental index scores.

The Religious Orthodoxy variable assesses the extent to which one endorses the notion that only God should be the one who controls whether or not someone lives or dies, a notion evaluated by the rating and ranking of the participant in the single item 9 in DIT-2 Cancer dilemma (Bebeau & Thoma 2003; Thoma 2006). Since its score is decided by the participant's rating and ranking of one single item (item 9 in Cancer story), it's necessary for the researcher to study this item carefully to find out the reason for the inconsistent outcome with the general pattern of DIT2.

Item 9: Should only God decide when a person's life should end?

There can be two possible factors that may contribute to the result: translation problem and concept problem.

To exclude the interference of the translation factor, we first invite the two people mentioned at the beginning of the paper, who are quite familiar with both Chinese and English, to review the translation of this item. They do not think that there is any mistake in translation.

To exclude the concept factor, we interview some participants to see whether there is any difference in understanding this item. Interview reveals that God, the concept in American Christian culture, is foreign to the Chinese people and its ethics are not unanimously upheld in Chinese culture, which is built on Confucianism rather than Judeo-Christianism (Dien 1982). Confucianism containing the most influential ethics for Chinese culture and Chinese people advocates for self-reflection and self-cultivation, which is independent of the power of God. In that situation, if the concept of God is translated literally, it fails to arouse the same empathy on Chinese participants as it is on American counterparts. Therefore, the Chinese participants rate and rank it quite low, which accounts for the low scoring in Religious Orthodoxy. As

a result, we need to find out another concept that Chinese people are quite familiar with in indigenous culture to replace “God” in this item.

2. There shows no significant relationship between Humanitarian/Liberalism and part of moral judgment developmental index scores.

The current research shows that Humanitarian/Liberalism has a significant relationship with Personal Interest score and Maintaining Norms score. The higher Personal Interest score is, and the lower Maintaining Norms score is, the more pronounced Humanitarian/Liberalism is, which is consistent with the pattern of DIT2. But different from the American participants whose Humanitarian/Liberalism score is positively related with their P/N2 score, Chinese participants do not show any significant relationship between these two factors.

Thirdly, results obviously show that it is difficult for Chinese participants to identify meaningless items in DIT2.

Meaningless items are designed to purge the questionnaire from further analysis. Interview reveals that the Chinese participants fail to get the implied meaning of the meaningless item and show a high likeliness to take them as related to the story. In that case, it is possible that a high percentage of questionnaire will be regarded as invalid, which will influence the result.

Fourthly, from the research we can see that Chinese participants favor stage 3 items while American participants favor stage 4.

This result is consistent with the finding of some researchers by using DIT test. “Chinese tend to have a more pervasive and persistent emphasis on Stage 3 morality than Western people.” (Ma & Cheugn, 1996), “[The results reveal that]Taiwanese procurement executives demonstrate higher stage scores for the conventional level than for postconventional level of moral judgment development.”(Lin, 2009). Those studies contribute the difference between Chinese and Westerners in stage preference to the Chinese culture which emphasizes harmony and personal relationship. But there can be other possible reasons. There is possibility that Chinese people construe issue statements of different moral stages in a way different from Western people, which is shown by the different understanding of meaningless items in DIT2 test.

5. Limitations of the Present Research

The current research has some limitations as the followings:

Firstly, compared with the mega sample of American participants (n=13487), the number of the subjects in this research is quite limited (n=113). A larger data base is expected to be set up, which will make the result more objective and persuasive.

Secondly, in American mega sample, there are subjects coming from students in school, while in the current research, all the subjects have graduated from university and been working for more than 10 years. The working environment is different from the school environment, which may affect the result. Therefore, researchers should take in more subjects who are studying in the school to compare with the American counterparts.

6. Future Prospects

To say that, the future research would focus on:

1. Revise the present translated version of DIT2 to make sure that there is no barrier in understanding. For example, the current research reveals that the concept "God" in item 9 in story 3 cannot arouse the same feeling in Chinese participants as the original version in American participants. Therefore, researchers need to work out a better translation of "God" in the Chinese version.

2. Change the story and the item that is not suitable for the Chinese participants due to the cultural shock with what is suitable for them to achieve the same effect with the original version. For example, the current research shows that Chinese participants are not likely to tell the meaningless items in DIT2, therefore, the researchers should consider to adjust the meaningless items for Chinese cultural psychology to achieve the same effect with the original version on the participants.

3. Find out the social and educational reasons in China for the lower moral judgment competency, based on which the researcher puts forward the effective strategies to improve it.

References

- Bebeau, M. J. & S. J. Thoma. 2003. *Guide for DIT-2: A Guide for Using the Defining Issues Test, Version 2 (DIT-2) and the Scoring Service of the Center for the Study of Ethical Development*. Tuscaloosa, AL: Center for the Study of Ethical Development.
- Colby, A. & L. Kohlberg. 1987. *The Measurement of Moral Judgment: Vol. 1. Theoretical Foundations and Research Validation*. Cambridge, UK: Cambridge University Press.
- Davison, M. L., D. Robbins, & D. Swanson. 1978. "Stage Structure in Objective Moral Judgments." *Developmental Psychology*, Vol. 14: 137-146.
- Dien, D. S. 1982. "A Chinese Perspective on Kohlberg's Theory of Moral Development." *Developmental Review*, Vol. 2: 331-341.
- Chieh-Yu, Lin. 2009. "An Empirical Investigation of the Moral Judgment Development of Taiwanese Procurement Executives." *Social Behavior and Personality*, Vol. 37(1): 102-108.

- Haigen Gu, & Li Boshu. 1997. "Constructing the Chinese Moral Judgment Competence Rating Scale for 12-17-year-old School Children." *Psychology Science*, Vol. 20: 197-201.
- Hing-Keung, Ma. 1988. "Objective Moral Judgment in Hong Kong, Mainland China, and England." *Journal of Cross-Cultural Psychology*, Vol. 19(1): 78-95.
- Hing Keung, Ma & Ch. Chau-Kiu. 1996. "A Cross-Cultural Study of Moral Stage Structure in Hong Kong Chinese, English, and Americans." *Journal of Cross-Cultural Psychology*, Vol. 27: 700-713.
- Hing-Keung, Ma & Ch. Wing-Shing. 1987. "The Moral Judgments of Chinese Students." *The Journal of Social Psychology*, Vol. 127(5): 491-497.
- Irene Hau-Siu, Chou & D. Z. Q. Ding. 2002. "Moral Judgment and Conflict Handling Styles Among Chinese in Hong Kong and PRC." *Journal of Management Development*, Vol. 21(9): 666-679.
- Li Yao. 2004. "Reflection and Reconstruction of Individual's Moral Cognitive Development - Studies of Neo-Kohlbergian Theory" (Master's thesis). URL: http://epub.cnki.net/kns/brief/default_result.aspx. Retrieved on: 15.06.2016.
- Martin R. M., M. Shafto, & W. VanDeinse. 1977. "The Reliability, Validity and Design of the Defining Issues Test." *Developmental Psychology*, Vol. 13: 460-468.
- Rest, J. R. 1979. *Development in Judging Moral Issues*. Minneapolis: University of Minnesota Press.
- Rest, J. R. & D. Narvaez. 1998. *Defining Issues Test 2*. Minneapolis, MN: University of Minnesota.
- Rest, J. R., D. Narvaez, M. J. Bebeau, & S. J. Thoma. 1999. "A Neo-Kohlbergian Approach: The DIT and Schema Theory." *Educational Psychology Review*, Vol. 11: 291-324.
- Rest, J. R., D. Narvaez, M. J. Bebeau, & S. J. Thoma. 1999. *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Rest, J. R., D. Narvaez, & S. J. Thoma. 1999. "DIT2: Devising and Testing a Revised Instrument of Moral Judgment." *Journal of Educational Psychology*, Vol. 91: 644-659.
- Shaogang, Yang. 2006. "Defining Issue Test and Neo-Kohlbergian Theory." *Thoughts, Theory and Education*, Vol. 7: 26-29.
- Shaogang, Yang & W. Huihong. 2004. "Defining Issue Test and Analysis of the Construction Components in Moral Psychology." *Educational Science*, Vol. 20(6): 56-59.
- Thoma, S.J. 2006. "Research Using the Defining Issues Test." In M. Killen & J. Smetana (eds.) *Handbook of Moral Development* (pp. 67-91). Hillsdale, NJ: Lawrence Erlbaum.

- Xinyin, Chen. 1988. "Defining Issues Test of Moral Judgment and Its Application." *Psychology Science*, Vol 3: 19-24.
- Yi-Hui, Ho & L. Cheh-Yu. 2008. "Cultural Values and Cognitive Moral Development of Accounting Ethics: A Cross-Cultural Study." *Social Behavior and Personality*, Vol 36(7): 883-892.

Qian Zhang

(Guangdong University of Foreign Studies, gwzhangqian@126.com)

Stephen J. Thoma

(University of Alabama, sthoma@bamaed.ua.edu)

An Empirical Cross-Cultural Study of Moral Judgment Development in Mainland China

Abstract: The Chinese version of Rest's Defining Issues Test II was administered to 113 subjects in Mainland China (n=113, average age=34.7). The scores on development of moral judgment were compared with those of the online mega sample of American participants from 2011 to 2014. Results are as followings: 1. Chinese participants show the same pattern with Americans by both sex and education. 2. Chinese participants show different pattern from Americans by religious orthodoxy and humanitarian. 3. Chinese participants score higher in meaningless items than Americans. 4. Chinese participants score higher in stage 3 while Americans score higher in stage 4. The authors draw the conclusions as follows: with Chinese participants, 1. There is a significant relationship between education and moral judgment developmental index scores. 2. There is also a significant relationship between sex and moral judgment developmental index scores. 3. There shows no significant relationship between religious orthodoxy and moral judgment developmental index scores. 4. It is more difficult for them to tell the meaningless items in DIT2. 5. Since Chinese culture thinks less of laws and norms, Chinese participants favour personal interest schema more than maintaining norms schema.

Keywords: DIT2, Chinese participants, morality, personal interest, maintaining norms, postconventional judgment

Ethics in Progress (ISSN 2084-9257). Vol. 7 (2016). No. 1, Art. #16, pp. 228-245.

doi: 10.14746/eip.2016.1.15