DOCUMENT RESUME

ED 260 121 TM 850 447

AUTHOR Moon, Yong Lin
TITLE A Review of Cross Cultural Studies on Moral Judgment Development Using the Defining Issues Test.
PUB DATE 85
NOTE 27p.; Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985). Small print in tables.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Cross Cultural Studies; *Cultural Differences; Evaluation Criteria; Literature Reviews; *Moral Development; Moral Values; Psychological Studies; Religious Differences; Rural Urban Differences; Sex Differences; *Test Validity; Trend Analysis; Value Judgment
IDENTIFIERS *Defining Issues Test

ABSTRACT

This paper reviews 20 cross-cultural studies conducted with the Defining Issues Test (DIT). All or part of the samples in the studies were non-Americans who had not been a part of the populations presented in the DIT manual. The studies were aimed at the comparison of moral reasoning structure and development across cultures. Several aspects of the findings were reviewed: (1) the psychometric properties of cultural versions of the DIT; (2) effects of the examinee's ethnic background; (3) age and education trends; (4) sex differences; (5) correlations with other psychological tests; (6) religious differences; (7) urban-rural milieu; (8) delinquent behavior; and (9) familial and social factors. The studies reviewed in this report satisfy the evaluation criterion of the generalizability of psychological theories developed in one culture to another culture. The DIT seems to have cross-cultural validity in detecting moral reasoning structure and its development in cultures outside the United States. The validation and explanation roles of the DIT in cross-cultural studies should be pursued together. (DWH)

A Review Of Cross-Cultural Studies On Moral Judgment Development Using The Defining Issues Test

Yong Lin Moon

Department of Educational Psychology
University of Minnesota
206B Burton Hall,
178 Pillsbury Drive,
Minneapolis, Minnesota 55455
(612-373-5539)

Paper presented at the division E (human development) of AERA Annual Meeting, Chicago, Illinois
March 31 - April 4, 1985

Running head : Moral development

A Review of Cross Cultural Studies On Moral Judgment Development Using The Defining Issues Test

I. Introduction

Hundreds of studies have been conducted using Rest's Defining Issues Test (Moon, 1984). Among the studies, cross-cultural studies are frequently found. In this paper, 20 cross-cultural DIT studies which were available to the author will be reviewed. They were regarded as "cross-cultural" in that all or part of the samples in the studies were non-American who had not been a part of the populations presented in the DIT manual(Rest,1979b).

The samples of the 20 studies represent 15 cultures or countries: Australia(3), Brazil(1), Greece(1), Hong Kong(2), Iceland(1), India(1), Israel(1), Japan(2), Korea(1), Mexico(1), Philippines(1), Saudi Arabia(1), South Africa(1), Taiwan(2), and Trinidad and Tobago(1). All the studies but one were cross-sectional and "etic" in the sense that their aims were primarily at the comparison of moral reasoning structure and development across cultures.

Seven of the 19 etic studies employed two or more culturally different samples and attempted to test the hypothesis related to cultural background.The other 12 studies employed samples of the same cultural background and attempted to compare the results to other already established results such as shown in the DIT manual (Rest, 1979b, pp.7.1 - 7.18), for example. Their objectives vary from validation studies of the DIT to correlation and intervention studies. A variety of factors were related to the DIT scores: for example, age, education, sex, religion, region, scores from various psychological tests, familial/societal factors, and so on. This review will proceed to discuss the cultural validation issues first and the factors later. Table 1 below presents a summary overview of the 20 studies.

Insert Table 1 here

II. Review of the findings

A. Psychometric properties of cultural versions of the DIT

The equivalence of test stimuli is a basic requirement for a comparison to be made and is the most vexing problem in cross-cultural research (Fredericsen, 1977). In DIT studies,

comparisons of the DIT scores across cultures make sense only if cultural versions of the original DIT measure the same thing that the original DIT has measured in the American studies. Such equivalence could be addressed in terms of psychometric properties, translation, and cultural adaptation.

One of the important ways of establishing the equivalence is to see if cultural versions of the original DIT have the same psychometric properties (e.g., reliability, validity, item structures,etc.) and the same relationship with other convergent/divergent variables that the original DIT has shown so far. First of all, the psychometric properties of the culturally adapted and/or translated versions of the DIT will be examined before looking into other results of the tests.

The original DIT has quite high test-retest reliability (.82 and .77 for 6- and 3-story form respectively) and internal consistency (.77 and .76 for 6- and 3-story form respectively in terms of Cronbach's alpha). The 3-story form tends to show lower levels of reliability and internal consistency. Four out of 19 studies report reliability and/or internal consistency measures. By and large, test-retest stability and internal consistency of the translated DITs are somewhat low except the Australian DIT (Dickinson, 1979,r=.98-.99) which is almost identical to the original DIT except for a few changes in several words. In Korea, Park, J and Johnson,R,s study report a quite comparable (but a little lower) level of reliability(r=.69).In the studies of Bzuneck (1978,r=.39) and Hau (1983,r=.32), the reliabilities are quite low. Therefore these measures may not be successful adaptations of the original DIT. Only two studies report internal consistency measures which are somewhat lower bound(Dickinson, 1979, alpha=.66; Hau,1983, alpha=.50)

The original DIT has shown consistently significant correlations with Moral Maturity Score(MMS) of Kohlberg's test at .43 - .70 level and with the Law and Order Test at -.60 - -.58 level. Considering these measures are supposed to relate by definition to the construct of P-score of the DIT, similar relationship can be expected in the translatedetic DITs if they are really equivalent to the original DIT. The P-score of the Icelandic DIT (Thorlindsson, 1978) has negative correlation with Law and Order Test at -.4498 level which compares to -.58 in the American studies (Rest, 1979b). Two Chinese studies (Tsaing, 1980,r=.486; Ma, 1980r=.20 - .29) show that P-score of their Chinese DITs have moderate correlations with a Kohlberg's measure. In Ma's study (1980), the correlation was much lower than those of the American and Tsaing's sample. But the high homogeniety of Ma's sample may be a cause.

In Hau's study (1983), the psychometric property of his Chinese DIT is extensively explored. It, therefore, is worth mentioning more in detail. First, he replicated McGeorge's fakability experiment(1975). The result confirmed that his

subjects were unable to fake good. Second, he replicated some of Davison's analysis of the DIT's stage and internal structure (Davison, 1979), using a correlational matrix of stage scores and a factor analytic method. He found a simplex pattern of intercorrelations and a two factor structure as demonstrated in Davison's study (even though there was a minor reversal of stage ordering). Third, he conducted an item analysis. A Good item should have the highest correlation with its own stage score and low correlations with other stage scores. His "results show that in 61 out of 72 items, the correlations with their own stage scores were the highest. For the remaining 11 wrongly keyed items, nine items did not have significant correlations with any of the 6 stages, thus precluding them from being reassigned to other stages"(p.70).

He also checked the Alpha level when each item is deleted from its subscale.The result showed that all but 11 items have the highest correlation with its own subscale and 66 out of 72 items show decreased Alpha when they are deleted. It shows, by and large, that his DIT has comparable psychometric properties to that of the original DIT(p.70)

Translation and cultural adaptation are also important factors affecting the equivalence of the translated and the original DIT. The insensitivity to translation error is well reflected in the studies, because information about translation is rarely reported and, when reported, very superficially. Only three studies describe their translation method or procedure briefly: multiple translation (Thorlindsson, 1978), team and back translation (Benor et al, 1978), and the bilingual method (Hau,1983). The first two studies did not attempt to identify translation error systematically. They just "tried their best". But in Hau's study (1983), translation error or its effect was detected through the experimental method as suggested by Prince & Mombour (1967) : "subjects were randomly divided into two groups. The first group took the first half of The DIT in English and the second half in Chinese ; the order was reversed for the second group. After two weeks, they were retested, using the test of the other group (p.34)." The result shows no significant difference in all scores (2, 3, 4, 5a, 5b, 6, and D-score) but the P-score. It was interpreted that to the Chinese student, the P-score in the English version was an underestimation whereas the scores in the Chinese version are the subject's actual level of moral judgment. Therefore, it was contended that the difference in P-score is at least in part due to subject's deficiency in English.

In the course of translation, M-items (check items for faking good) were found difficult to translate. For example, Thorlinsson reports that the M-items (especially #5 item in "Newspaper" and #6 item in "Prisoner" dilemma story) were perceived by subjects as a Stage 4 or a Principled stage item. Such difficulties were also reported in Gendron's study. M-items

have another problem besides translation. According to Gendron, if he dropped out subjects who had M-score of 4 (the cutoff point in the 3-story form), he would have lost 37.2% of total subjects (usually 5-15% in American sample). A slightly higher than expected dropout rate was also reported in Benor et al's study (1982, 17.1%) and Hau's study (1983, 16%). The details of attempts at making cultural adaptation are rarey reported . In Dickinson's study (1979) in Australia, several American words were changed into Australian equivalents. In Benor et al's study (1982), the name, nationality, occupation of the actors in the story were changed to fit the Israeli setting. The effects of such changes were not systematically examined.

B. Effects of ethnic background

According to Kohlberg(1971), "all individuals in all cultures go through the same order or sequence of gross stages of moral development, though varying in rate and terminal point of development" (p.175). In the DIT studies, the effects of cultural background have been investigated in several ways.One type of study selects two or more samples having different subcultural backgrounds and treat them as representing levels of a cultural dimension as an independent variable -- for instance, Anglo-and Greek Australian (Watson,1983) and American and Japanese- American (Jacobson, 1977). Another type of study compares the scores of subjects from one culture to those of another culture(usually, the USA) -- for instance, Hau(1982) and Park & Johnson(1983).. Seven of the 20 studies belong to the former category like Watson and Jacobson's study.

Insert Table 2 here

As shown in Table 2, the studies treat cultural background as an independent variable while holding the age and education level constant. Considering that the effect of age-education factor on moral judgment development(P-score) approximates 38% of the total variance (Rest, 1979a, p.110), the effect of cultural background could be well delineated if they succeed in holding them constant. This control is not clear in Jacobson's(1977) mother subjects and Deyoung's(1980) teacher samples. In Ma's study (1980), subjects are not controlled in age.

In the seven studies, ten comparisons were made between ethnic groups: 6 comparisons of Western versus non-Western culture, 4 of Western versus Western cultures. Five out of six Western versus non-Western comparisons showed significant difference in P-score, while none of Western versus Western comparisons showed significant differences. In general, the DIT P-score is more likely to diverge between Western and non-Western samples than between Western and Western samples. In

most of the comparisons, the closer the sample to American culture, the higher the P-score tends to be. That is, American subjects scored higher in P-scores than their respective counterparts in Mexico, Japan, and Saudi Arabia. English adolescents did better than Greek agemates. Anglo-Australians scored higher than miscellaneous Australian groups, Greek-Australians, and Asian-Chinese (See Table 2 for the explanation of the four ethnic groups).

C. Age/education trends

Nothing is more crucial to a cognitive developmental construct than evidence of change over time from less advanced forms of thinking to more advanced forms of thinking. Age trends in moral development have been used by Piaget and Kohlberg as a major justification of their claims. The cross sectional trends have been well shown in many DIT studies as well, as is shown in Figure 1 for the American sample in the DIT manual (Rest, 1979b).

Insert Figure 1 here.

In the case of the American sample, one-way ANOVA of the four student groups showed an extraordinary strong developmental trends ($p<.0001$) and age-education was found to account for 38% of the total variance of moral judgment. The American's average DIT score tends to increase about 10 points with each increase in level of education (Rest, 1979, p.110). Comparing the P-scores of American samples to those of other cultures, it is interesting that the American samples are not always the most advanced as shown in Figure 1.

All six studies of non-American samples show clear developmental trends in P-score. This means that older and better educated subjects are likely to attribute more importance to the higher stage issue statements (items). In Hau's study(1983), age/education accounts 12% and 23% (Omega-square) of the variance of P and D scores respectively, suggesting somewhat lesser effect of age-education on moral reasoning development than in the American study. Even though statistical test were not performed, the general trends shown in Figure 1 appear that the P-score increase with age-education, and that the rate of development seems different between Western (Americans, Icelandics, and Australians) and Eastern sample (Koreans and Chinese). It seems that the older and better educated the Eastern subjects, the more they lag behind their American agemates. As shown in Figure 1, in junior high school years the Eastern subjects' P -scores are either higher or quite close to that of American counterparts. But as they get older and more educated, the rate of increase in P-score appears to

slow down.This seems to be not the case with Icelandic and Australian samples. Such slow-down pattern was also indicated by Lei(1981) in his research on Chinese sample. However, this may be an overinterpretation of the data, because it is not certain that the difference of the rate is significant or not without any test of significance between Western and non-Western.

In the developmental trends, age and education are confounded each other especially in student samples, since the more educationally advanced subjects are invariably older. In American samples, education was found to be more related than chronological ageto the increase in P-score. Hau's(1983) and Watson's(1983) studies directly compare the effects of age and education on the DIT scores. In Hau's study, when the education was controlled, the correlation of DIT scores(P and D score) with age were very low ranging from -.17 to .22 ($p < .001$).This means that most of the contribution of age-education to the DIT P-scores(12% in P-score ,23% in D score) are due to education and not age (p.43).

In Watson's study (1983) also,educational attainment was more strongly related to moral judgment level than age. He found a significant interaction effect between age and education suggesting that the picture is not as clear as that proposed by Rest (1979a). There, Rest suggests that adults in general do not show much advance in moral reasoning beyond that accounted for by their level of education (p. 112). Watson(1983) interprets the interaction as suggesting that a P-score in the low 30s represents the minimal level achieved by adults with at least some high-school education. Further advances in the level of moral judgment beyond this point appears to be related to the experience of higher education" (p.60).

Some variables related to education have been explored to see if they have any consistent relationship to the DIT scores. Two studies address the relation of school type to the DIT scores (Beddoe, 1980; Deyoung, 1982). Benor, Notzer, Sheehan, & Norman(1982) investigated the relationship of interview score gathered from a medical school entrance examination with the DIT scores. Clarke(1978) studied the relative potency of teacher-related variables as a predictors of the development of the children.Career interest and school speciality were also related to the DIT scores by Miller (1980) and Prahallada (1983). In those studies, DIT scores were found to relate significantly or at least marginally to the factors such as school type, medical school examination, teacher potency, and career interest of students.

D. Sex difference

The charge of sex bias in Kohlberg's test has drawn much attention to the sex variable. While some researchers have claimed that Kohlberg's model is biased against women because a

morality of justice is not attuned to the thinking patterns of women (Gilligan, 1977), other researchers have argued that there is little evidence for the alleged sex-bias claim , because there are not significant and consistent sex difference found in Kohlberg's test (Walker, 1985) and in the DIT (Rest, 1979). Recently Thoma (in press) conducted a meta-analysis on the accummulated DIT data of 56 samples comprising 6000 male and female subjects. He found that " overall, and at every age/education level, females score significantly higher than males. But the magnitude of the difference is small"(p.1).

Seven studies directly address this issue. Most of the Chinese and Korean subjects except Ma's study (1980) show sex difference in P-score and Stage 4 score with females scoring higher P-scores, and lower Stage 4 than male subjects. In Tsaing's study (1980), a sex difference is shown consistently through age 13 to 17 except at age 16. Two studies with Western samples (Jacobson, 1980; Watson, 1983) found no sex difference.

It is interesting that Stage 4 score is higher in male subjects in Hau's Chinese and Park & Johnson's Korean samples,. One more interesting fact is that the sensitivity of P-score and D-score to sex difference is mixed . In Hau's study (1983), the P-score has more discriminating power over the D-score in detecting a sex difference. In Ma's study (1980), however, the P-score was found less sensitive than the D-score in his English sample(n=108), but later when the test was refined with more subjects(N=272), the result was the opposite($p < .100$). It was not explained whether this was due to just a measurement error or due to a true difference.

In Fox's study, sex differences were shown in the M-score in both Greek and English subjects with male subjects having almost double the score of females. Other studies do not have information on sex differences in M-score. Even though there are significant sex difference in many studies, the actual magnitude of the difference is not large. By and large, the result of the seven studies seem to be consistent with the conclusion of Thoma (in press).

E. Correlations with other psychological tests

The DIT has been found to have a consistent relationship with some psychological variables as predicted from theory. For example, the P-score is hypothesized to have a moderate correlation with IQ measures, while it is hypothesized to have a negative or inconsistent correlation with dogmatism and F-scale. With American samples (Rest, 1979,pp.198-201), the DIT was found to correlate in the .20 to .50 range with measures of IQ, aptitude, and achievemement and to have moderate correlation with some scales of the California Personality Inventory (e.g.,

(Achievement via independence,.48;Intellectual efficiency, .42) and Omnibus Personality Inventory (e.g., Autonomy, .47;Complexity, .45).

Four of 20 studies carry some information on the relationship of DIT scores to other psychological tests. The trends are quite similar to those of American studies. The Chinese sample of Hau's study shows very similar correlation between the P and IQ score(Raven's Progressive Matrix, hence RPM) to that of the American sample. In Prahallada's study (1982) in India, the P-score has no positive correlation with IQ measures of RPM at a significant level($p<.05$). Two cognitive measures of Thorlindsson's Icelandic sample(Elaborate Language Test and Role taking ability) show somewhat low but significant correlation with P-score. In the case of personality variables, Ma's(two of California Personality Inventory subscales) and Prahallada's studies(Bell's Personality Adjustment Inventory) show similar but a little lower correlation between DIT scores and personality scores than reported for American samples.

F. Religious affiliation.

No significant difference has been found in the moral reasoning development among Protestants, Catholics, Jews, Buddhists, Moslems, and atheists in Kohlbergian studies (Kohlberg, 1971, p.174; Kohlberg and Kramer, 1976). The studies using the DIT also report no significant difference in DIT P-score among religious affilfation (Getz,1983). But as shown in Lawrence's study(1978; Ernsberger, 1976), idiological(or religious) commitments can override conceptual adequacy in making moral judgments. In this sense, the issue of religion to moral judgment development is not settled.

Three studies investigated the relationship of religious affiliation and the DIT scores. In two studies (Watson, 1983;Beddoe, 1980), religious affiliation made no significant difference in P-scores. This is consistent with other studies in Getz's review (1983). In Gendron's study (1981), direct comparison between the non-Christian and the Catholic group is not warranted because they are different in age-education level. Of interest is the seminarian and nun group (Group 4 in Table 1), who are undergraduate Catholic students from the department of thelogy and religious studies. Their DIT P-scores are lower than those of other Catholic students (Group 3), the Catholic young professionals (group 5), and are not significantly higher than the senior high students (Group 1 and 2). Cohort effects may be a possible explanation because the group covers a wide age range (23-63 years old). As Gendron suggests, another plausible explanation may be sought from the background and personalities of this sample: "most of the older subjects in the group are nuns, who were brought in a conservative environment, where Catholic moral teaching had a

clear and unquestioned answer for any moral problem. Strict adherence to official church teaching was required (p.3)". If the inference is correct, the religious factor should be regarded as having profound effects on moral judgment development. His study has an implication that the non-significant result of the previous studies on religious affiliation and moral development may be due to its superficial or misguided categorization of the religious variable.

G. Urban-rural milieu

As Rest(1983) suggests, "the lack of paticipation in secondary institution (e.g., the national legal system and government, bureaucracies of industry) makes it less likely that those people will evolve the conceptual schemes of stage 4, 5, and 6 that deal especially with moral problems at this more abstract level" (p.58). The urban-rural dimension may be a proxy indicator for the degree of participation in such institutions. Therefore, villgers are hypothesized to score lower in moral reasoning development than town dwellers.

Three studies (Thorlindsson, 1978; Park & Johnson, 1983; Ismail, 1976) directly compare the difference in DIT scores between urban and rural subjects. Only Korean subjects in Park & Johnson's study show significant difference between the two groups in the P- and Stage 3- score. Urban subjects score high in P score, but rural subjects score higher in Stage 3. The difference was attributed to lack of cultural relativism and greater shame orientation of rural subjects.

In Thorlinsson's Icelandic study (1978), significant difference was found only in stage 4 score between male and female subjects of the rural sample. The researcher interprets it as indicating that "the recent change in women's role has not spread to the village community". The female's high stage 4 score is in contrast to the claim that stage 4 is more likely to be a male stage (Holstein 1976).

The result of the three studies can not be counted equally in weight because the samples used in the studies are quite different. In Thorlinsson's study, the subjects are 14-15 years old students living in a village or city at the time of study. In Ismail's study (1976), the subjects were 20-29 year old college students who had come to the United States for advanced study . Park & Johnson's subjects include students ranging from 6 graders to college students living in a village or in a city at the time of study. Therefore as far as the samples are concerned, Park and Johnson's finding could be given more credit.

H. Delinquent behavior

The relation of moral judgment to behavior is complicated and mediated by many factors. According to Rest's four Component Model (1983), moral judgment is part of Component II : "therefore moral judgment is one player in a large cast of players, and even if it is a star, it is not the whole show" (p.600). As Blasi(1980) concludes, nevertheless, the significant, but not very strong relationship can hardly be denied.

Delinquent behavior has been found to have a significant relationship with measures of social cognitive development such as Kohlberg's test, role-taking skills, and Piaget's Golden Rule task (Rest, 1983). In DIT studies, delinquents (male prison inmates) were found to score lower in the P-score than the general adult group (see Rest, 1979a, p.188). In case of adolescent subjects (16 year olds), McColgan(1975) found that they were also significantly lower in P-score(18.8) than the normal group(28.7).

Two of the 20 studies investigated delinquent adolescents in Brazil or South Africa. Some of the DIT scores were found to be related significantly to delinquency variables. It is interesting that Stage 4 score seems to have more consistent relationship with delinquency variables more than the P-score.

Bzuneck (1978) assessed the discriminating power of the DIT between delinquent and non-delinquent adolescents from either intact or disrupted families. The average P-score of the four groups were not found significantly different. However, Stage 2, Stage 4 and the A index showed significant difference among the four groups. Delinquents from a disrupted (father absent) family scored higher on stage 2 than the other three groups. However, this group was lower in the Stage 4 score than other groups. On theA-index, delinquent groups were higher in the P-score than non-delinquent groups ($p<.05$).

The average P-score of the Brazilian sample was quite low. The non-delinquent's P-score(18.7%) was lower than that of delinquent's(20.1%). Although the difference was not statistically significant, it needs close examination. Maybe, the low P-score of non-delinquent group was due to their sample characteristics, for instance. In other words, the non-delinquents might not well represent the non-delinquent population of Brazilian adolescents. A good evidence is the forced match of age and education level of the non-delinquent group. Another reason can be sought in the low reading level of the both groups(the author assumed their subjects' education level was grade 4 to 5).

Heyns et al.(1981) studied the relation of moral reasoning level to behavioral dimensions of juvenile delinquency. The subjects were 57 delinquent boys in a reformatory catering for severe cases. Their average P-score was 21.06(%). The

researchers related DIT scores to four dimensions of delinquency which were factored out from the Quay measures and subject's biographic data. P-score were found to correlate negatively with the social inadequacy dimension. Stage 4 was significantly associated with social orientation dimension. Such results were regarded as validating information for both the dimension of delinquency and the DIT. The non-significant correlation between anti-social egocentric dimension and the P-score was, however, unexpected contradictory to the previous findings such as reported by Fordor (1973).

I. Familial/social factors

Many studies have the kinds of familial and societal factors related to moral development (c.f., Edwards, 1981;Rest, 1979). Although Piaget emphasized the influence of peers on chilren's moral development more than that of parents, Kohlberg acknowledges both. Peers or parents are important because they are the main source of role-taking opportunities. This opportunity varies with family and society. Therefore, "Kohlberg postulates that differences in role-taking opportunities among cultures and socio-economic class can speed up or slow down development through the stages. Lower socio-economic classes in several countries seem to go through the stages slower than high SES groups (Kohlberg, 1969; Rest, 1983, p.598)."

Dickinson(1979) investigated the relation of moral development to sample characteristics of family factors in the Australian setting. Eleven sample characteristics (father's occupation, parent's education, country of origin, number of children, types of school attended, religious denomination, etc.) account for only 5% of the total P-score variance (p.128). Best predictors were father's occupation, religion, sex, and father's education. The correlation of P-score among family members were significantly high(adjusted R-square=.79).

Tsaing(1980) explored the influences of child rearing pattern, family size, birth order, and SES on P-score. All but child-rearing pattern showed non-significant differences. Four types of child-rearing pattern were analyzed in relation to sex and four age levels. A significant correlation was found only at "high love-high induction" at age 14 male subjects ($r = .25$, $p < .05$).

The interaction among family members was investigated in ralation to DIT scores. According to Dickinson (1979), Style A-type verbal utterance (rational and persuasive) has significant correlation with the P-score over Style B-type verbal utterance (imperative and authoritative). In addition, Thorlindsson (1979) studied the interaction of mother and child in relation to the P-score. Mother-child person oriented interaction (reasonable and reciprocal) has significant positive correlation with P-score over mother child position oriented

interaction(power and authority). By and large, the findings are consistent with general expectation as Edwards(1980) indicated: "the most successful parent are expected to be those verbally and overtly rational who encourage warm and close relation with children and who promote a democratic style of life (p.31)".

J. An emic DIT

Only one out of the 20 studies can be regarded as using an emic DIT in the sense that it was not just a translated version of the original DIT for cross-cultural comparison but a full adaptation of the DIT into the target culture with no direct intention of cross-cultural comparison. Villanueva (1982) constructed a measure of moral judgment (The Exercise in Evaluating Issues, hence EEI), that is patterned after the DIT but has moral dilemmas familiar to Filipinos and incorporates values commonly accepted by them. Therefore, the EEI can not be considered just as a culturally adapted and/or a translated version of the original DIT. It is interesting to see how similar or different the two measures are. The sample comprised 500 students from diverse educational levels ranging from 4th year high school students to graduate and seminarian students. The EEI scores correlate highly with the scores of the original DIT, with correlation coefficients ranging from .63 to .91 (average r=.8347, p<.01). The seminarians obtained the highest E-score (EEI's equivalent of P-score in the DIT) among all the groups. However, though the average E-score increases as educational level goes up, the differences in mean E-score among the high school students, college students, and graduate students were not significant. Only the seminarian group differed significantly from the other groups.

The reliability of EEI turned out quite high (r=.74-.91 in test-retest measure). The most significant single factor which influences the EEI score proves to be education (r=.4166). The next factor was father's education (r=.1296). When considered jointly, age and urban residence were the most significant pair. Though age did not appear as a significant single factor, it appeared significant in interacting with other variables. Based on these findings, the researcher concludes that the EEI has very similar properties to that of the DIT, even though the average principled morality level was quite low compared to the American's P-score. It is somewhat surprising that only 13% of seminarians and only 2% of graduate students are predominantly using principled moral reasoning. Although, all in all, the EEI reveals its close similarity to the original DIT, the lower discriminating power among various educational-level groups may be a serious drawback in detecting and tracing the developmental trends of Filipinos' moral judgment development.

III. Conclusion and perspectives

There are two major kinds of questions of interest in cross-cultural studies. The first question concerns whether psychological theories developed in one culture generalize to other cultures (e.g., is Kohlberg's stage theory universal?). The second interest is in using cross-cultural variation to explore and illuminate relationships among psychological variables (e.g., do different social roles of women in different cultures effect their moral judgment development?).

In terms of the first question of interest, cross-cultural studies are supposed to be "proving grounds". Because the main theme of the study is to see if the theoretical assumptions and its methodology work or not in another culture, no specific hypothesis related to the target culture is required. Just as in replication studies, the findings of these studies are compared and evaluated in terms of its similarity or difference to the findings of the original studies.

In terms of the second question of interest, however, cross-cultural studies have to adopt any specific hypotheses concerning the relation of some cultural factors or aspects to moral reasoning structure or level. In other words, because the main theme of the studies here is to see what is the effect of the culture on moral reasoning development, specific testable hypothesis are required. In this case, therefore, the cross-cultural studies are no longer "proving grounds". The similarity or difference across cultures are not just compared but explained in terms of what cultural factors makes the difference or similarity.

Applying these two questions to th studies reviewed here as criteria for evaluation, I regard the findings as moderately satisfying the first criterion but not the second one. Based on the findings above, it may be possible to conclude that all in all the cross-cultural findings are comparable to American findings and Kohlbergian findings. The similarity of the findings among them are a good demonstration that the DIT has cross-cultural validity in detecting moral reasoning structure and its development in other cultures outside the USA.

However, considering that the studies reviewed here are small in number(20), not exhaustive in the cultures covered(15), and by no means perfect in research design and analysis, the conclusions are not strong. The studies reviewed here can not be considered as having fully carried out the role as a " proving ground ". First, they did not cover so broad a spectrum of the age-education sample as in the original DIT studies. Second, they did not fully explore the relations of convergent and divergent variables to the measures of the DIT. Third, most of the studies did not show the psychometric properties of the translated DIT to be comparable to those of the original DIT.

For instance, if the DIT items had been biased against some cultural groups over the other cultural groups, then the direct comparisons of the DIT scores (e.g., P-score) might have resulted in errorneous conclusions. Fourth, most importantly there was no longitudinal evidence.

One of important goals of cross-cultural study lies in the explanation of cultural effects on moral reasoning structure and its development. Most of the studies reviewed here have not given much attention to this point. Few tried to explain the similarity or difference of their findings to that of American studies in terms of cultural variation. Considering that most of the studies are the first attempt of using the DIT in their own culture, it is no wonder that they had to be more involved in the issues of the first criterion.

I regard that the validation and explanation roles of the DIT in cross-cultural studies should be pursued together. To strengthen the explanation role, it is important to formulate specific, testable, and meaningful hypothesis concerning the relation of the cultural aspects to moral reasoning structure and its development. In formulating such hypotheses, it may be necessary to have interdisciplinary perspectives including those of anthropology, ecology, history, and sociology.

Reference

Beddoe,I.B. (1980). Assessing principled moral thinking among student teachers in Trinidad and Tobago. Unpublished manuscript, University of West Indies, Trinidad and Tobago.

Benor,D.E.,Notzer,N.,Sheehan,T.j., & Norman,C.R.(1982). Moral reasoning as a criterion for admission to medical school. Paper presented at the AERA annual conference,March.

Blasi,A.(1980). Bridging moral cognition and moral action: A critical review of the literature. Psychological Bulletin, 88,593-637.

Bzuneck,J.K.(1978) Moral judgment of delinquent and non-delinquent adolescents in relation to father absence. Unpublished doctoral dissertation,Brazil.

Clarke,J.(1978). An intervention study on moral judgment development in Australian Primary school. Unpublished doctoral dissertation, Macquarie University.

Davison,M.(1979). Internal structure and the psychometric properties of the Defining Issues Test. In J.R.Rest, Development in judging moral issues, Minneapolis,Minnesota: University of Minnesota Press,223-245.

Deyoung,A.M.(1982). A study of relationships between teacher and student levels of moral reasoning in a Japanese setting. Unpublished doctoral dissertation, Michgan State University.

Dickinson,/./.(1979). A study on the development in relation to commonly used sample characteristics and family factors in Australian setting. Unpublished doctoral dissertation, Macquarie University,Australia.

Edwards,C.P.(1980) The comparative study of moral judgment and reasoning. In R.L. Munroe,R.Munroe, & B.B. Whiting (Eds.), Handbook of Cross-cultural human development. New York: Garland Publishing Company.

Ernsberger,D.J..(1976). Intrinsic-extrinsic religious identification and level of moral development. Unpubiished Doctoral dissertation, University of Texas.

Fordor, E.M. (1973). Moral development and parent behavior antecedents in adolescent psychopaths. Journal of Genetic Psychology, 122, 37-43.

Fox,P. (1982) Stages of moral development in Greek and English schools. Unpublished manuscript,England.

Fredericsen,N.(1977). How to tell if a measures the same thing in different cultures. In Y.P.E.Poortinga(Eds.),Basic problems in cross-cultural psychology, Tilbug: Swets and Zeitlinger B.V., 12-18.

Gendron,L. (1981). An empirical study of the Defining Issues Test in Taiwan. UnpublishedManuscript,Fujen Catholic University,Taiwan.

Getz,I.R.(1983). The relation of moral reasoning and religion. Unpublished manuscript,University of minnesota.

Gilligan,C.F.(1977). In a different voice: women's conception of self and morality. Harvard Educational Review, 47,481-517.

Hau,Kit Tai.(1983) A cross-cultural study of a moral judgment test(The D.I.T.). Unpublished master's thesis, Chinese University, Hong Kong.

Heyns,P.M., Niekerk & Rouk,J.A. (1981) Moral judgment and behavioral dimensions of juvenile delinquency. International Journal of Advanced Counseling,4,139-151.
Holstein,C.B. (1976). Irreversible,stepwise sequence in the development of moral judgment: a longitudinal study of males and females. Child Development,47,51-61.

Ismail,M.A.(1976). Across-cultural study of moral judgment: The relationship between American and Saudi Arabic university students in the Defining Issues Test. Unpublished doctoral dissertation, Okalahoma University.

Jacobson,L.T. (1977). A study of relationships among mother,students,and teacher levels of moral reasoning in a Department of Defense Middle School. Unpublished doctoral dissertation, Michigan State University.

Kohlberg,L.(1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study od moral development. In T. Mischel(Ed.),Cognitive development and epistemology. New York: Academic Press.

Kohlberg,L. (1969). stage and sequence: The cognitive-developemtal approach to socialization. In D. Goslin(Ed.), Handbook of Socialization theory and research. Chicago: Rand McNally.

Kohlberg,L. & Kramer,R. (1969). Continuities and discontinuities in child and adult moral development. Human Development, 12, 93-120.

Lei,T. (1981). The development of moral,political and legal, reasoning in Chinese society. Unpublished master' thesis. University of Minnesota.

Ma,H.K. (1980). A study of the moral development of adolescents. Unpublished master's thesis, University of London.

Martin,R.M., Shafto,M.,& Van. (1977). The reliability, validity and design of the DIT, Developmental Psychology,13,460-468.

McColgan,E.B.(1975). Social cognition in delinquents,predelinquents, and non-delinquents. Unpublished doctoral dissertation, University of Minnesota.

McGeorge,C.(1975). The sussceptibility to faking of the Defining Issues Test of moral development. Developmental Psychology,11,108.

Miller,D.G.(1979). A cross-cultural study of moral development in Mexican and American adolescents. Unpublished master's thesis, Universidad de la Americas,Mexico.

Moon,Yong-Lin.(1983). A bibliography of DIT studies on moral judgment development. Technical Reports, Minnesota Moral Research Projects(MMRP), University of Minnesota.

Park,J.Y.& Johnson,R.C. (1983). Moraldevelopment in rural and urban Korea. Unpublished manuscript, Hankook University of Foreign Studies,Seoul,Korea.

Prahallada,N.N. (1982). An investigation of the moral judgments of junior college students and their relationship with the socio-economic status, intelligence and personality adjustment. Unpublished doctoral dissertation. University of Mysore, Mysore, India.

Prince r. & Mombour,W.(1967). A technique for improving linguistics equivalence in cross-cultural surveys. International Journal of Social Psyciatry,13,229-237.

Raven,T.C.(1975). Guide to the Standard Progressive Matrices: Set A,B,C,D, and E. London: H.K. Lewis & Co.

Rest,J.R. (1983). Morality. in P. Mussen(Ed.) Carmichael's Manual of Child Psychology. New York: John Wiley & Sons, 556-629.

Rest,J.R. (1979a). Development in judging moral issues. Minneapolis,Minnesota: University of Minnesota Press.

Rest,J.R. (1979b). Revised manual for the Defining Issues Test: An objective test of moral judgment development. Minnespolis,Minnesota: Minnesota Moral Research Projects

Rest,J.R.(1976). New approaches in the assessment off moral judgment. In L.Lickona(Ed.), Moral development and behavior, New York: Holt,Rinehart & Winston,198-220.

Thoma,S. (in press). Estimating Ginder difference in the

comprehension and preference of moral issues. Developmental Psychology.

Thornlidsson,T. (1978). Social organizattion, role-taking, elaborated language and moral judgment in an Icelandic setting. Unpublished Doctoral dissertation, University of Iwoa.

Tsaing,W.C. (1980) Moral judgment development and familial factors. Unppublished Master's thesis, National Taiwan Normal University.

Villanueva,E.S. (1982). Validation of a moral judgment instrumen for Filipino students. Unpublished doctoral dissertation, University of the Philippines system, Quexon City,Philippines.

Walker,L.(in 1984). Sex difference in the eveolpment of moral reasoning :A critical review, Child Development.

Watson,W. (1983). A study of Factors effecting the Development of moral judgment. Unpublished manuscript, Monash Chirering, Clayton, Victoria, Australia.

Table 1
Summary of cross-cultural studies using the DIT

Culture(country), Authors & Design	Sample description	N	Age Grade	P(%) Score	S.D	Major variables	Forms of DIT, Language, Methods of translation	Validity/ Reliability	Cultural adaptation
1. AUSTRALIA									
Dickinson(1979)	HS students	14	15			religion,	long,	Test-retest	words
	HS students	761	16	32.2	12.7	sex,	English,	r=.98~.99	
etic, cross-sectional	HS students	334	17			family &	———		
	HS students	19	18			friends		Cronbach	
								α=.66	
Watson (1983)									
	Study I :	10	14.7	19.2	6.2	age, sex,	long,	hi	dilemmas,
		10	16.8	32.7	11.2	education,	English,		phrases,
	Anglo-Australian (age 14-58)	10	18.7	43.8	7.0	non-students	———		words
		10	21.2	47.5	8.0				
	Study II :								
	Anglo-Australian	20	18.8	43.7	7.0				
	Miscellaneous Australian	10	18.6	41.8	13.5	culture,	long,	ni	names,
	Greek Australian	10	18.5	34.6	9.2	sex, religion	English, ———		concepts (e.g. God)
	Asian Chinese	20	19.11	32.6	12.7				
Clerke(1978)									
	PR pupils	617	5th	-	-	teacher, familial& societal& intervention effects	long, English, ———	-	-
etic, cross-sectional, intervention	Teachers	24	-	41.7	-				
2. Brazil									
Bzuneck(1978)	Adolescents:								
	Delinquents	40	12 - 18	20.2	9.3	Delinquency,	short,	test-retest (r=.39)	
etic, cross-sectional						father-absence	Portuguese, ———	r=-.13~.32	
	Non-delinquents	39	12 - 18	18.7	7.8			r=.0~.51	
3. Greece									
Fox (1982)	HS students:								
etic, cross-sectional	English	33	17	33.9	-	culture & sex	short, Greek, ni	ni	ni
	Greek	18	17	32.3	-				
4. Hong Kong									
Ma (1980)	HS students:					age, sex,			
etic, cross-sectional	English	108	15.2	26.1	14.0	culture, item-analysis	5-story form Chinese, ni	[illegible] r=.[illegible]	ni
	Chinese	78	17.1	27.9	12.5				
Hau (1983)	Students:								
	JR students	68	7-8th	25.2	-	age, educ, IQ, sex, fak-ability, item-analysis translation	long, Chinese, bilingual	test-retest r= .32	
	JR students	71	9-10th	29.3	-			= .50	
	SR students	69	11-12	34.5	-				
	Univ-students	34	-	37.9	-				

Table 1
Summary of cross-cultural studies using the DIT (Cont'd)

Culture(country), Authors & Design	Sample description	N	Age Grade	P(%) Score	S.D	Major variables	Forms of DIT, Language, Translation	Validity/ Reliability	Cultural adaptation
5. Iceland									
Thorlinsson (1978)	Students:					familial-societal,		P-score & Law & order r=.4498	
etic, cross-sectional	JR students	27	8 th	21.6	-	urban-rural,	short,	ni	M-items
	College & graduate-students	19	-	56.0	-	role-taking, language, mother-child interaction	Icelandic, multiple translation		
6. India									
Prahallada[5] (1982)	HS students:	16	-	-	-	IQ, SES,	ni, Indian (Kannada),	ni	ni
		17	-	-	-	personality(Bell Inventory)	ni		
etic, cross-sectional		18	-	-	-				
		19	-	-	-	school-speciality			
7. Israel									
Benor et al (1982)	Applicants for medical school								
	School I[a]					admission interview	short, Hebrew back-translation	ni	name, nationality, occupation
	accepted	44	-	39.4	12.4				
	rejected	135	-	40.0	13.2				
	School II[a]								
	accepted	38	-	50.0	17.0				
	rejected	101	-	39.4	12.8				
8. Japan									
Jacobson (1977)	Americans living in Japan								
etic, cross-sectional	US teachers[a]	30	-	42.6	10.0	sex, ethnic background	long, English,	ni	ni
	US mothers[a]	63	-	34.8	14.1				
	Japanese-born mothers[a]	24	-	28.6	12.0				
	Children of Japanese-born mothers	24	-	24.5	8.3				
	Children of US mothers	63	-	18.5	6.9				
Deyoung[b] (1982)	Japanese teachers & students & US teachers								
etic, cross-sectional	Jr college students	47	-	-	-	culture, school type, teacher/ students	ni, ni, ni	ni	ni
	University students	47	-	-	-				
	Jr college teachers	10	-	-	-				
	University teachers	17	-	-	-				
	US English teachers living in Japan	30	-	-	-				

Table 1
Summary of cross-cultural studies using the DIT (Cont'd)

Culture(country), Authors & Design	Sample description	N	Age Grade	P(%) Score	S.D	Major variables	Forms of DIT, Language, Translation	Validity/ Reliability	Cultural adaptation
9. Korea									
Park et al (1983)	Students:								
	JR students	60	6th	25.0	-	age, educ	short,	test-retest	ni
	JR students	60	8th	30.2	-	sex,	Koraan,	r= .69	
etic,	SR students	60	11th	37.4	-	urban-rural	ni		
cross-sectional	College students	60	-	41.5	-				
10. Mexico									
Miller(1979)	Mexican & US students:								
	Bilingual Mexican students	37	11 - 12 th	19.6	-	sex, age, career interests	long, English	ni	ni
	US students living in Mexico	35	11 - 12 th	22.5	-				
11. Philippines									
Villanueva (1983)	Students					age, educ	6-story;form	test-retest	emic
	HS students	77	13.8	21.3	7.5	sex,	ni,	r=.74 - .91	
emic,	HS students	70	15.3	21.8	9.6	urban-rural	ni	DIT & EEI[18] r=.83	
cross-sectional	College students	42	20.4	22.3	7.5	familial factor			
	Graduate students	16	32.9	23.5	7.2				
	Seminarian	23	26.0	31.6	8.9				
12. Saudi Arabia									
Ismail(1976)	US & Saudi students in US college								
etic,	American								
cross-sectional	Undergraduate	20	28	22.0	8.4	culture,	long,	ni	ni
	Graduate	20	28	29.5	7.5	length of stay in USA, urban-rural	English,		
	Saudi-Arabian								
	Undergraduate	21	28	15.7	4.7				
	Graduate	19	28	18.5	6.2				
13. South Africa									
Heyns et al (1981)	Delinquents (boys)	57	11-12 th	21.1	12.7	demographic data & Quay Scale	long, English,	ni	ni
etic, cross-sectional									
14. Taiwan									
Tsaing(1980)	Students:								
	JR students	160	13	20.4	6.8	age, educ	long,	[illegible]	ni
etic,	Jr students	165	14			sex, SES, Family size	Chinese,	r=.[illegible]	
cross-sectional	SR students	158	16	26.8	6.2	birth order child-rearing	Group discussion		
	SR students	172	17						

Table 1
Summary of cross-cultural studies using the DIT (Cont'd)

Culture(country), authors & design	Sample description	Ns	Age / Grade	P(%) Score	S.D	Major variables	Forms of DIT, Language, Translation	Validity/ Reliability	Cultural adaptation
Gendron (1981) etic, cross-sectional	Catholic students:								
	Group 1	41	16-18	31.7	11.5	age, educ, religion	short, Chinese, ni	ni	M-items
	Group 2	37	16-18	28.9	10.7				
	Group 3	36	20-26	37.9	10.7				
	Group 4	40	23-65	33.4	14.6				
	Group 5	16	22-30	44.4	15.8				
15. Trinidad & Tobago									
Beddoe(1981) etic, cross-sectional	4 teacher's college students								
	College[12] I	210	20 - 39	35.1	-	age, educ religion school type	short, ni, ni	ni	ni
	College[12] II			25.0	-				
	College[12] III			26.2	-				
	College[12] IV			26.7	-				

Note:

1. long: 6 dilemma story DIT
   short: 3 dilemma story DIT
2. PR : Primary school
   JR : Junior High School
   SR : Senior High School
   HS : High School
3. ni : No information or information not available
4. Anglo-Australian students: Students whose parents were born in Australia or who migrated from Britain.
   Greek-Australian students: Students parents were of Greek descendents and at least one parent grew up in Greece and came to Australia after eighteen years of age.
   Miscellaneous Australian : Students whose at least one of parents was of non-English speaking background and did not meet the criteria for inclusion in one of the Anglo- and Greek-Australian group.
   Asian- Chinese : Students whose both of the parents were of Chinese discent.
5. Only the synopsis or abstract of the study was used in this review.
6. Bell's Personality Adjustment Inventory
7. School I : The Sackler School of Medicine, T l Aviv University(STA), where applicants are selected only on the basis of previous scholastics achievement and performance on a psychometric test.
   School II : The faculty of Health Sciences, Ben-Gurion University(BGU),where applicants are selected on the basis of personal characteristics, interpersonal skills and orientation toward the community as expressed by previous behavior with down-played scholastic achievement.
8. US teachers: American teachers employed by a Department of Defense Overseas school located in Japan, a Middle School.
   US mothers : American-born women who are wives of American enlisted men or officers or Department of Defense civilians serving in Japan.
   Japanese-born mothers: Japanes-American women who are wives of American soldiers or DOD civilians serving in Japan.
9. emic: A term coined by Pike(1966) to refer to an anthropological approach in which cross-cultural comparison is ruled out.
10. EEI : The Exercise in Evaluating Issues
11. Quay Scales: 3 lists and questionaire were used
    a. The Behavior Problem Checklist(Quay & Parsons,1972)-B list
    b. The Checklist for the Analysis of Life History Data(Quay & Parsons 1972)-A list
    c. The Personal Opinion Study(Quay & Parsons,1972)- P list
12. College I: Catholic Women's Teacher's College(CWTC)
    College II: Government Traing College (GTC) -Coed
    College III: POSTC- coed
    College IV: Mausica-Coed

Table 1

Cultural(ethnic background) effect

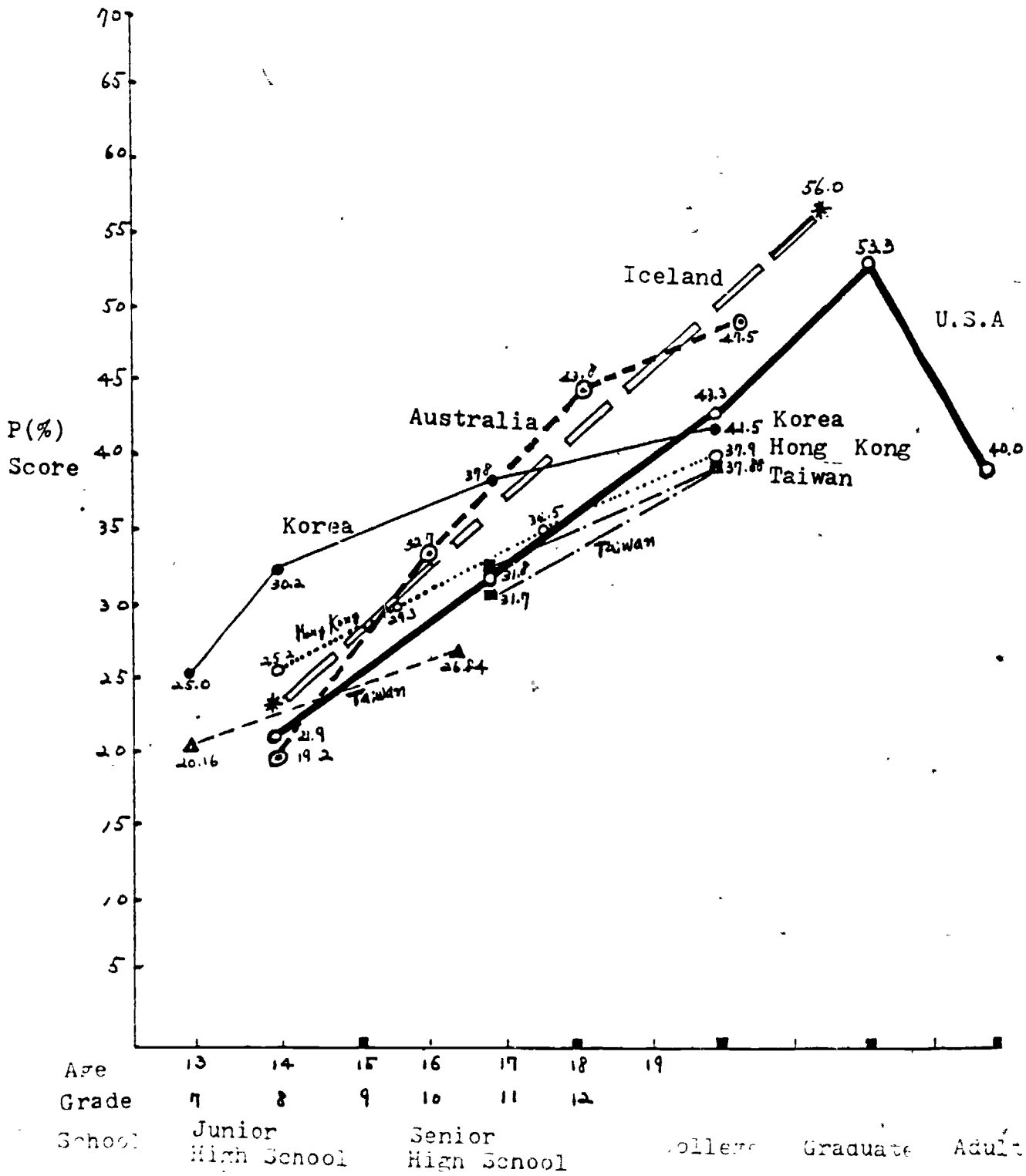
Study	Sample	age	N	M(%) score	SD	Significance
Watson (1983)	1 Anglo-Australian CS[2]	18.8	20	43.75	6.97	1 & 2 NS
	2 Misc- Australian CS	18.6	10	41.84	13.55	1 & 3 *
	3 Greek-Australian CS	18.5	10	34.00	9.17	1 & 4 *
	4 Asian Chinese CS	19.11	20	32.73	12.73	
Miller (1979)	American HS	17	35	37.6	-	NS
	Mexican HS	17	31	32.8	-	
Fox(1982)	English HS	17	33	33.9	-	NS
	Greek HS	17	18	32.3	-	
Jacobson (1977)	1 American-born mothers[3]	-	42	43.83	14.14	1 & 2 *
	2 Japanese-born mothers	-	15	28.26	12.0	
	3 Children of Japanese-born mothers MS	10-14	39	18.51	6.90	3 & 4 *
	4 Children of American-born mothers MS	10-14	21	24.57	8.36	
Ma(1980)	English HS	15.2	108	26.11(15.6)	14.0(4.2)	P-score: NS
	Chinese HS	17.1	78	27.9(18.3)	12.5(3.2)	(D-score): *
Ismail(1976)	US undergraduates	28	20	22.00	8.40	25.58(8.70) *
	US graduates	28	20	29.15	7.46	
	Saudi Undergraduates	28	21	15.71	4.74	16.95(5.64)
	Saud graduates	28	19	18.52	6.22	
Deyoung (1980)	US English teacher	-	30	-	-	NS
	Japanese college teacners	-	17	-	-	

1: See Table 1 for detailed explanation.
2: College students
*: $p < .05$

NS: Non-significance
3: See Table 1 for detailed explanation.

Figure 1

Age-education trends of P(%) score



A Review of Cross-cultural Studies on Moral Judgment Development Using the Defining Issues Test

Yong Lin Moon

University of Minnesota

(Abstract)

Twenty cross-cultural studies on moral judgment development using the Defining Issues Test (DIT) were reviewed with respect to cross-cultural validity, age/education trends, gender difference, correlations with other psychological factors, religion, urban-rural, delinquency, and familial/societal factors. By and large, results indicate that the DIT has similar psychometric properties (factor structure, internal consistency, and reliability) and construct validity in non American cultures. However, the rate of development and the strength of relationships with other variables were found to differ across cultures, especially between Western and non-Western cultures. It is suggested that fine grained research is needed to explain (instead of describing) the differences and similarities in terms of cultural causation.