مجلة
بحوث كلية الآداب
جامعة المنوفية

1

The Development of Moral Reasoning Judgment in Samples of Intermediate, Secondary School and University Male and Female Students in Kuwait

By

Ramadan A.Ahmed

Department of Psychology College of Social Sciences Kuwait University

محكمة تصدرها كلية آداب المنوفية
العدد الثانى والخمسون يناير ٢٠٠٣

Setting:

Kuwait is a rapidly changing country which at the time of the study (2001-2002) had about two million predominantly Arab inhabitants. The country has large financial resources based upon its oil and natural gas industries, a very high per capita income, a comprehensive welfare system, and free compulsory education for all boys and girls between the ages of 6-14. The country's population includes 40% Kuwaitis and 60% foreigners, among which, Arabs, Persians, Indians, Sri Lankans, Philippines and Pakistanis, are especially numerous.

Moral value system in Kuwait centers on the Islamic cultural heritage. Islam assumes that Qur'an contains the true, ultimate and unchanging word of God as revealed to the Prophet and messenger Mohammed. It contains a corpus of religious beliefs, prescriptions and laws supporting social and political institutions. Together with the religious traditions, the Qur'an orders social relationships and institutions with a special emphasis on the family. Male and female roles are sharply differentiated. The social ethic of Islam stresses humility, solidarity, equality and God piety, charity towards the less fortunate, and rejection of sin.

The Development of Moral Judgment:

A basic but controversial claim of Kohlberg's cognitive-developmental theory states that moral reasoning develops according to a universal sequence of cultural invariant stages (Gielen, Comunian, & Antoni, 1994). This meant that in all societies the development of moral reasoning follows the same fixed order without stage regression or stage

skipping, although the rate and endpoint of moral development may vary from one society to the next.

Research suggests that in diverse societies moral reasoning develops in more or less stage-like form as a function of age, education, intelligence, and sociocultural experience (Comunian and Gielen, 1995).

Kohlberg, (1964, 1984) has developed a theory that conceptualizes the structures and process underlying the development of moral reasoning. As a follower of Piaget, he assumes that development occurs through a series of cumulative stages that vary in complexity. Development is said to follow a sequence of stages and to occur in the same manner in all societies regardless of cultural differences. The speed and endpoint of development through the stages vary from person to person and from society to society, but the basic sequence of stages is assumed to be universal. Kohlberg emphasizes that his assertions apply only to the structures of moral reasoning whereas the specific contents of moral decision-making may vary both from person to person and from culture to culture.

Kohlberg's six stage model traces development from the preconventional stages 1 and 2, to the conventional stages 3 and 4, and finally to the principles stages 5 and 6. At the preconventional stages a person makes moral decisions from the point of view of an individual act or who tries pragmatically to avoid trouble, to satisfy his or her needs, and to follow practical rules of fairness, exchange, and social acceptability. At the conventional stages a person has internalized interpersonal expectations of societal virtues, religious prescriptions, and the duties, obligations, and rights prevailing in his or her culture or

subculture. The person tries to be morally good by responding to the justified expectations of others and those of society. At the principled level of morality a person has developed inner abstract of principles of justice that are used to make decisions in difficult situations involving moral conflict. Such moral principles appeal to notions of human dignity, ultimate equality under the moral law, moral autonomy, and conceptions of an ideal just society. Moral decision-making is expected to be sharable or universalizable, representing an effort to reach consensus based on nonarbitrary social cooperation (Ahmed and Gielen, 2002).

Kohlberg has developed an instrument designed to assess the moral stages or reasoning, the Moral Judgment Interview (MJI). It is based on the interview method, relies on a very complex coding manual, and is quite cumbersome to use (Kohlberg, 1964, 1984).

Given the complexities of the measurement instrument, his student James Rest (1979) has developed a new technique to assess the development of moral reasoning stages, the Defining Issues Test (DIT). The DIT is based on Kohlberg's developmental theory of moral reasoning and aims at measuring an individual's preference hierarchy for stage-typical moral arguments.

The DIT is an objective test containing six brief moral-political dilemmas which are similar to those employed in Kohlberg's MJI. After each dilemma, there follow 12 moral arguments that could be used to solve the dilemmas. The arguments vary in structural complexity and represent Kohlberg's moral reasoning stages 2, 3, 4, 4½, 5A, 5B, and 6. In addition, the test contains three so-called consistency checks in order to discover whether the respondent understands the meaning of the

arguments and rates and ranks the arguments in a consistent and meaningful way. Included in the consistency check are some meaningless but lofty sounding arguments. If a person chooses several of these arguments he or she is eliminated from the subject pool because it is unclear whether the person fully understands the moral arguments contained in the test. The respondent is also asked to rate and to rank the same series of moral arguments. If a person rates and ranks the same arguments in a very different way he or she is also eliminated from the subject pool because it is unclear whether the person fully understands the test. Alternatively, such a person may simply be careless in making moral choices. The "consistency checks" are of special importance in cross-cultural research because they help the researcher to decide whether the DIT is an appropriate test in a given culture (Rest, 1979; Rest, Narvaez, Bebeau and Thoma, 1999).

After the test-taken has rated and ranked all moral arguments contained in the test, his or her moral choices are integrated across all dilemmas and moral arguments by using an objective scoring procedure. The scoring procedure determines the percentage of the moral arguments chosen by the respondent for each of the moral stages. In this way, one can establish for each respondent a "moral judgment profile" representing all moral stages preferred by the respondent. However, the moral stages are unequally represented on the test. Conventional moral arguments (stages 3 and 4) are most frequently represented followed by post conventional arguments (stages 5A, 5B, 6) and preconventional arguments (stage 2).

Rest (1979) recommends the so-called P%-Score as the best overall indicator or a person's moral judgment maturity. The P%-Score refers to

the combined percentage of moral arguments chosen by a person who represents stages 5A, 5B, and 6. In other words, the P% -Score indicates how much the person prefers principles moral arguments over arguments at the lower preconventional (stage 2 in the DIT) and conventional (stages 3 and 4) stages.

Another overall indicator which can be obtained from the test is the D-Score. It integrates a person's ratings of all presented moral arguments into an overall developmental score (Ahmed and Gielen, 2002).

Critics of Kohlberg and Rest such as Vine, 1986 (cited in Ahmed and Gielen, 2002), have argued that Kohlberg's and Rest's conceptions of principled morality are biased toward a specifically Western, male, upper-class ideology of moral autonomy and individualism.

The critics claim that Kohlberg's and Rest's tests of moral reasoning set up an ethnocentric standard or moral maturity that should not be applied to non-Western cultures. However, there exists a broad array of cross-cultural studies whose results are at variance with most claims of the critics. These studies have been reviewed by Gielen (1996a), Ahmed and Gielen (2002), Gielen and Markoulis (2001), and Naito, Wen-Ying and Gielen (2001). A special issue published by World Psychology provides an up-to-date overview of the strengths and weaknesses of Kohlbergian theory and research (Gielen, 1996b).

A review of DIT studies conducted in Arab countries:

In the following, the author reviews the results of research projects conducted in various Arab countries that are, at times, compared to the results of research projects in Western and East Asian countries (Ahmed and Gielen, 2002; Gielen, 1996a and b; Gielen and Markoulis, 2001;

Gielen, Miao and Avellani, 1990; Naito et al., 2001). Such cross-cultural comparisons must be done in a critical manner. Specifically, the DIT was developed in the United States and it is therefore important to ask whether the test provides comparable estimates of moral reasoning development among Arab and non-Arab respondents.

In a pioneering, small-scale, cross- cultural study, Ismail (1976) found that cultural differences do exist in the moral development of American and Saudi Arabian Muslim respondents. It was noted by Bouhmama (1984) that Ismail's findings may reflect some combination of religious as well as general cultural differences between the two nation groups.

In 1983, Webb and Steentsma administered the DIT to 259 male and female Christian and Muslim Palestinian university students. They found that correlation between years of schooling and the P%-Score was very low, though statistically significant ($r=.16$, $p < .05$). Age and P%-Score did not correlate significantly with each other. Practitioners of religion received lower scores than non-practitioners. The Palestinian students received an overall P%-Score of 29.35% which is much lower than the mean P%-Score typically reported for comparing North American samples which have been summarized by Rest (cited in Ahmed, Gielen and Avellani, 1987).

Jessup (1984) administered the DIT to 40 multinational students in 9th grade English classes at an American school in Kuwait. They were around 15 years of age. The purpose of the study was to investigate the moral development of multinational students in Kuwait in a culturally disparate educational setting. The study also aimed at determining both effects of knowledge of the Kohlbergian theory of cognitive moral development gained through a short training period and of a request to

"fake high" on the DIT's principled moral thinking items (P%-Score). It was found that the respondents had a P% mean of 17.06 on the DIT pretest. This score falls within the range of the standard deviation for American peer groups although at the lower end. It was also found that the discussion of Kohlberg's moral stage theory did not significantly increase the participants' ability to select principled items. No significant advantages for the experimental group were found when examining the means of their post-test scores. As for the possible effect of sex, nationality, and language, no significant differences emerged in the statistical tests, although the relatively small number of respondents should be kept in mind in this context.

In 1985, El-Sheikh administered the DIT to 333 Egyptian male and female intermediate, secondary school, and college of education students ranging in age from 15 to 32.5 years. Results showed that participants preferred conventional moral judgments (stages 3 and 4=50%) over principled arguments (stages 5 and 6=20.30%) and preconventional arguments (stage 2=12%). Significant differences between males and females concerning stage 3 arguments in favor of the female subjects have been found. A significant difference ($p < .01$) was found between males and females at stages 5 and 6 indicating that males preferred principled arguments more frequently than females. Level of education appeared to contribute positively to the students' moral judgment level. El-Sheikh reports that the difference between intermediate and secondary school students on one hand and the university students on the other hand was significant at the .01 level.

In the 1985, Eissa used an adapted Arabic version of the DIT to study the impact of education on the development of moral judgment. He

compared the responses of his small sample (n=15) of senior males, majoring in Arabic, aged from 23-27 years, attending the College of Education, Tanta University, Egypt, and having a pre-college education at a teacher's institute (secondary level) with those of their peers (n=16) having no college level training. Both groups were matched for the variables of sex, grade point average, and years of pre-college certification study. Eissa proposed that most adults reach a general plateau in the development of their moral judgment after their formal schooling and that pervious research results in Western studies demonstrates a strong positive relationship between level of moral judgment and years of education (Rest, 1979). A comparison between the two means of the P%-Score shows a significant difference in favor of the college group ($P < .001$). Their mean P%-Score was 20.13 (SD=4.33) as contrasted with the mean P%-Score of the non-college group which was 10.56 (SD=3.14). A further analysis of other stage scores indicated a significant change in the level of moral judgment according to the number of years spent in formal education at the College of Education. It should be added that the P%-Scores reported by Eissa are usually low when compared to scores typically obtained in the United States and elsewhere (Gielen and Markoulis, 2001).

In a later study, Eissa (1993) administered a locally adapted version of the DIT to 37 male and 40 female Egyptian secondary school students ranging in age between 15.4 and 17.3 years. The students also responded to Barron's Ego-Strength Score, a scale that had shown good psychometric properties in pervious Arab research. The correlation between the P%-Score and the Ego-Strength scale was r=.59 pointing to a strong relationship between the two variables. In addition, Eissa administered the DIT to 43 participants on two separate occasions. The

test-retest reliability coefficient over a time period of two weeks reached a very convincing r=.97.

Ahmed, et al., (1987) administered the DIT to 679 Sudanese male and female intermediate, secondary school, and university students. Sixty-two percent of the protocols did not pass the standard inconsistency check recommended by Rest, 1986. These results suggest that the DIT was a very difficult and unusual test for most of the Sudanese respondents. P%-Scores and D-Scores showed low but statistically significant correlation with age for the remaining 147 male students, but the remaining 102 female students the correlations were not significant. The results for the Sudanese students indicated that they strongly preferred conventional moral judgments (stages 3+4=57%) over principled arguments (stages 5A, 5B, and 6=26.6%) and preconventional arguments (stage 2=5.7%).

Compared to the results typically in American and East Asian studies, the Sudanese participants received much lower P%-Scores than usual (Gielen & Markoulis, 2001). For instance, the Sudanese college/university students in their junior and senior years received an average P%-Score of 24.5% whereas college students from Hong Kong received an average P%-Score of 37.9%, from South Korea a P%-Score of 41.5%, and from Taiwan a P%-Score of 41.4% (Gielen, Miao, & Avellani, 1990).

Ahmed, et al., also investigated the relationships between parental behavior (as perceived by the students) and moral reasoning skills. For male students, the father's acceptance was positively related to moral judgment development, while the father's aggression and indifference

were negatively related to moral judgment. For female students, the mother's warmth and acceptance were negatively related to moral judgment development, while the mother's aggression, indifference, and rejection were positively related to moral judgment development. The results for the Sudanese females were surprising and difficult to interpret.

In a second study, Gielen, Ahmed, and Avellani (1992) administered the DIT to 685 Kuwaiti male and female intermediate, secondary, and university students. 54.1% of the protocols were rated as inconsistent. Consequently, the data analysis included only protocols from 314 participants who successfully passed the consistency check.

The results for the Kuwaiti students indicate that they preferred conventional moral arguments (stages 3 & 4=57.32%) over principled arguments (stages 5A, 5B, and 6=26.57%) and preconventional arguments (stage 2=5.20%). It should be noted in this context that the DIT contained more conventional than principled or preconventional arguments. The obtained results, therefore, partially reflect the overall construction of the DIT.

A comparison between the samples from Kuwait and Sudan demonstrates a highly similar overall distribution of moral stage scores. The average P%-Scores for the two countries are almost identical (Sudan=26.64%; Kuwait=26.57). These P%-Scores are, however, only slightly lower than the average P%-Scores of 29.35% that Webb and Steentsma (1983) reported for a sample of 259 male and female Palestinians. In 1985, along the same lines El-Sheikh reported average P%-Scores ranging from 20.23% to 29.93% for 333 intermediate, secondary school, and college and graduate level male and female

students from Cairo, Egypt. In 1985 Eissa also, found in his Egyptian samples an average P%-Scores of 10.56% (SD=4.1) for the non-college group and an average P% -Score of 20.13% (SD=4.33) for the college group.

It was expected that the Kuwaiti sample would show a somewhat higher P%-Score than their Sudanese counterparts due to the influence of modernization. But, contrary to the authors' expectations, the results appeared to reflect similar rather than different socio-cultural influence for the two samples. Compared to the Taiwanese study (Gielen, et al., 1990), the Kuwaiti and Sudanese students received much lower P%-Scores, but higher stage 4 scores; this is consistent with the idea that the Kuwaiti and Sudanese students tended to adopt a conservative ideology.

In a subsequent study, Bouhmama (1988) administered, the DIT to 104 Muslim students, ranging in age from 18 to 43 years, who resided in England. The students came from many Arab countries and non-Arab countries. 90 of Bouhmama's 104 students (or 86.54%) passed to the consistency check. It should be added that the students were allowed to take the test home, a procedure that provided them with an unlimited amount of time to complete the questionnaire and presumably also with the opportunity to ask others about the meaning of different items. Bouhmama divided the students according to four educational levels: senior high school level, bachelor's degree level, master's degree level, and Ph.D. level. However, an ANOVA indicated no significant differences between the moral judgment scores of the four educational groups. This finding is important since it suggests that the moral reasoning of Muslim students residing in England did not become similar in its structure to that of their English peers.

In Saudi Arabia, Ben Abdel-Aziz (1989) administered the DIT to 260 male and female Saudi university students. While students majoring in different academic subjects did not differ from each other in their moral judgments levels, there was a significant effect due to level of education: Seniors received higher scores on stages 5A and 5B while freshmen were more likely to endorse moral arguments representing stages 2 and 3. There were no gender differences except that male students favored stage 5A arguments more frequently than female students.

Bouhmama (1989) administered the DIT to 100 Algerian male and female psychology students aged between 19 and 25 years old. The overall distribution of stage scores was difficult to interpret. Bouhmama does not report his DIT data according to standard procedures. Adding all mean stage scores, gives a result of 116%, a percentage that is clearly impossible (The total of stage score does not add up to 100% as it should). A similar problem arises when one compares Bouhmama's reported P%-Scores against the summed-up scores for stages 5A, 5B, and 6 separately for males and for females. For males, the reported average P%-Score is 15.27%, but the summed-up scores for stages 5A+5B+6 are equal to 19.17%. For females, the reported P%-Score is 24.83% whereas the summed-up scores fro stages 5A+5B+6 are equal to 36.57 (compare to numbers reported by Bouhmama, 1989). Another question arises when one compares Bouhmama's reported P%-Scores (P%=25.52) to the summed-up scores of stages 5A, 5B, and 6 (Combined equals 24-66%). In reality, these two scores should be equal. Therefore Bouhmama's scores cannot be compared to scores obtained in other DIT studies.

Habib (1991) sought to identify various characteristics of moral development among 100 male and 100 female university students in Egypt. In addition to the DIT and a set of general background questions, the author administered Arabic adaptations of various self-concept, social competence, and shyness scales. The internal validity for- the DIT was assessed on the basis of intercorrelations matrices between the various stage scores separately for males, females, and the overall sample. The resulting patterns of intercorrelations provided modest statistical support for the structural coherence of the DIT. The male and female students did not differ from each other in their stage scores.

In a recent cross-cultural study, Kamel and El-Shouny (2000) administered the DIT to samples of Muslim university male and female students in Egypt and Saudi Arabia. Unlike the researchers' assumptions, results showed that Egyptian respondents achieved significantly higher P%-Score (principled morality) compared with their Saudi counterparts (P%-Score for Egyptians was 14.65=24.42%, and for the Saudis was 12.78=21.30%). The Saudi sample however, scored significantly higher, than the Egyptian samples on stages 3 and 4 (conventional level of morality). As for gender differences, Egyptian and Saudi males, achieved significantly higher P%-Scores than their female peers. Yet, Egyptian and Saudi females outnumbered significantly males on stage 3. No significant gender differences have been found on stage 4.

In Lebanon, Ghusani and El-Hassan (2001) investigated the levels of moral judgment development using the DIT in 210 intermediate and secondary school male and female students. Students were divided into two sub samples according to their age (13-15 years, and 16-18 years old). The younger age group scored significantly higher on the

conventional level stages (2, 3, 4), while the eldest age group achieved higher moral scores on the post conventional level, or the principled morality, e.g. stages 5 and 6 and P%-Scores (P%-Scores were 24.3 and 28.9 respectively), which reflects the effect of age in the development of moral judgment.

Conclusions:

It appears that the DIT contains some moral issues that are more or less unfamiliar to many Arab respondents, reflecting cross-cultural differences in religion, politics, cultural traditions, and socialization patterns. In this context, the following two moral dilemmas contained in the DIT may be especially inappropriate for an Arab adaptation of the test: 1) The Doctor's Dilemma: Euthanasia is not known in Arab societies and it is against Islam; 2) The Newspaper Dilemma: The issue of freedom of speech in high school newspapers is an unfamiliar topic in Arab school systems. (This dilemma was not used by either Ahmed, et al., 1987; or Gielen, et al., 1992 in their studies.)

It is noted that only men are represented as the main characters in the DIT dilemmas, and this factor may have affected the moral evaluations by both male and female participants (Eissa, 1985, 1993). However, many DIT studies have found slight gender differences on the P%-Scores favoring female over male respondents (cited in Ahmed et al., 1987) suggesting that the gender of the DIT characters is not of crucial importance. It is also noted that a person's performance on the DIT, like that person's performance on, other objective tests, may be strongly influenced by his or her intellectual level, verbal comprehension, and social class background (Shafie, 1994).

The above mentioned survey of the studies using the DIT with Arab respondents has underlined that a number of methodological problems makes it difficult to arrive at clear-cup conclusions about the usefulness of the DIT in Arab cultural settings. Most importantly, the high rates, especially of Kuwaiti and Sudanese students that failed the three standard inconsistency checks, suggests to us that the DIT may not be a reliable and valid assessment instrument of the moral judgment skills of many Arab students at various educational and age levels. Various other studies reviewed above did not employ the inconsistency checks and/or contained other methodological shortcomings. In addition, as reported by Gielen et al. (1992), informal discussions with some Kuwaiti respondents showed that the students found many of the DIT's moral arguments to be "strange" and not easily understandable.

Taken together, the results of the studies under review suggest to the author that other Kohlbergian measures may prove to be more useful than the DIT for studies attempting to investigate the moral development of Arab respondents. For instance, Gibbs, Basinger, and Fuller (1992) have recently introduced a relatively brief production measure of moral judgment, the Sociomoral Reflection Measure-Short Form (SRM-SF).

In the following, a study is reported that investigates the development of moral judgment in samples of intermediate, secondary school and university students from Kuwait. The study focuses on four broadly defined goals: (1) To study the development of moral judgment in students from Kuwait; (2) To test the cross-cultural usefulness and validity of the Defining Issues Test (DIT) in Kuwaiti setting; (3) To assess the sex differences in the development of moral judgment and (4) To investigate the cross-cultural usefulness of some other tests for

assessing the development of moral judgment developed by various researchers (e.g. Sentence Completion Test [SCT], and Sociomoral Reflection Measure-Short Form (SRM-SF].

Method

Subjects

The Arabic versions of the Defining Issues Test (DIT); Sentence Completion Test (SCT), and Sociomoral Measure-Short Form (SRM-SF) were administered to 530 subjects, 246 males and 284 females, ranging in age from 13 to 35 years (mean age: 18.48). The sample included four groups representing different ages and educational levels. There were 52 eighth grade students (27 males and 25 females; mean age: 13.23 years), 82 tenth graders (45 males and 37 females; mean age: 14.44), 86 eleventh and twelfth graders (45 males and 41 females; mean age: 15.45), 40 adults with no more than ten to twelve years of education (15 males and 25 females; mean age: 22.48), and 220 university students (105 males and 115 females; mean age: 22.62).

Instruments

The development of moral judgment has been assessed by using four scales and measures, and they were:

1) The Defining Issues Test (DIT):

The DIT has been developed by J. Rest (1979), and contains six dilemmas. Example of the dilemmas is the poor husband. Should a poor husband steal a drug in order to save life of his very sick wife if he cannot get the drug any other way?

Following each dilemma, twelve arguments were provided that could be used to solve the conflict. They arguments reflected different

moral stages. Respondents were asked to rate the important of each argument. In addition, the subjects were asked to select the four most important arguments.

The DIT was objectively scored following procedure recommended by Rest. The test provided moral stage scores for stages 2, 3, 4, 4½ (A-Score), 5A, 5B, and 6. Preference for principled thinking (stages 5A, 5B, 6 combined) was expressed by the P%-Score which indicated the percentage of respondents' rankings falling within the principled stage.

Following standard procedures, respondents were removed from the sample if their selection of moral arguments were too inconsistent, if they appeared to be careless in their selections, or if they endorsed meaningless dummy items (Ahmed et al., 1987; Gielen et al., 1992).

2) Sentence Completion Test (SCT). This scale was developed by J. Loevinger (Loevinger, 1976, and through a personal communication with Professor U.P. Gielen, St. Francis College, New York, USA). The scale has two forms one for men and the other for women. Each form contains 36 items paper-pencil test measuring six aspects of personality: Self-concept, parental attitude, peer attitude, need for achievement, attitude towards law and order and body-image. Items are sentence stems, which the subject completes. The scale was translated into Arabic by the present researcher and back translated through existence discussions between the present researcher and Professor U. P. Gielen.

3) The Sociomoral Reflection Measure-Short-Form "SRM-SF" (Gibbs, Basinger and Fuller, 1992). This measure is a group administrable, pencil-an-paper production test designed to assess the development of moral judgment. The test is a shorthand adaptation of the Sociomoral Reflection Measure which was developed by Gibbs 1982, and which was derived from Kohlberg's (1984) Moral Judgment Interview (MJI). For Gibbs, Basinger and Fuller's (1992) measure acceptable psychometric properties with American samples have been reported, as proved true in same previous research with the Italian adaptation (Comunian and Gielen, 1995, 2000; Gielen et al., 1994). The measure contains 11 items addressing sociomoral values. Items 1 to 4 address the values of contract and truth; Items 5 and 6 pertain to the values of affiliation, Items 7 and 8 to life, and Items 9 to 11 to property, law and legal justice. Respondents are asked to evaluate and justify the importance of each value. The justificatory responses are then scored for stage of moral development.

The primary score is the Sociomoral Reflection Maturity Score, the mean of these ratings. These ratings range from 100 to 400, representing Stages 1 to 4. Stages of moral judgment may be also represented by a Global Stage Score indicating the development level within which Sociomoral Reflection Maturity Scores falls. Finally, the questionnaires may be scored for the presence of Type B responses when balanced, internal, and universalistic forms of moral reflection are present in a protocol. Gibbs et al., (1992) refer to Type B reasoning as moral ideality

The Arabic version of the SRM-SF, which was translated into Arabic by the present researcher, was obtained by the integration of two

3) The Sociomoral Reflection Measure-Short-Form "SRM-SF" (Gibbs, Basinger and Fuller, 1992). This measure is a group administrable, pencil-an-paper production test designed to assess the development of moral judgment. The test is a shorthand adaptation of the Sociomoral Reflection Measure which was developed by Gibbs 1982, and which was derived from Kohlberg's (1984) Moral Judgment Interview (MJI). For Gibbs, Basinger and Fuller's (1992) measure acceptable psychometric properties with American samples have been reported, as proved true in same previous research with the Italian adaptation (Comunian and Gielen, 1995, 2000; Gielen et al., 1994). The measure contains 11 items addressing sociomoral values. Items 1 to 4 address the values of contract and truth; Items 5 and 6 pertain to the values of affiliation, Items 7 and 8 to life, and Items 9 to 11 to property, law and legal justice. Respondents are asked to evaluate and justify the importance of each value. The justificatory responses are then scored for stage of moral development.

The primary score is the Sociomoral Reflection Maturity Score, the mean of these ratings. These ratings range from 100 to 400, representing Stages 1 to 4. Stages of moral judgment may be also represented by a Global Stage Score indicating the development level within which. Sociomoral Reflection Maturity Scores falls. Finally, the questionnaires may be scored for the presence of Type B responses when balanced, internal, and universalistic forms of moral reflection are present in a protocol. Gibbs et al., (1992) refer to Type B reasoning as moral ideality

The Arabic version of the SRM-SF, which was translated into Arabic by the present researcher, was obtained by the integration of two

methods of translation: the back-translation procedure, and the bilingual technique. An Italian version of the SRM-SF evidenced good levels of reliability and validity (Cronbach Alpha=.86; test-retest r=.83) which are comparable to the levels obtained with the original American version of the SRM-SF (cited in Comunian & Gielen, 1995, 2000).

An early form of the Social Reflection Measure has been used in very few Arab studies (such as Al-Ghamdi, 2001).

Scoring of the SRM-SF Protocols:

All SRM-SF protocols will be scored blindly by K. Bassinger, USA, trained by the process of study and practice in the manual. Interrater reliability between K. Bassinger and another American expert rater, based on 30 protocols randomly selected from a sample, will be performed.

Procedure

The Arabic versions of the DIT, SCT and SRM-SF were administered to eighth to twelve grade students, drawn classwise from some intermediate and secondary schools in Kuwait, during regular school hours in March and April and also in October 2001.

The university group was composed of psychology, sociology, biology and computer sciences students attending the second, third and fourth years of their university course. The protocols were handed out in the class and the application of the all three tests was divided into two sessions. Each session took 40-50 minutes.

The subjects of the less-educated group work comprise factories, governmental sector, police, and army in Kuwait. They were contacted by the trained researchers who indicated the research purposes and assured the subjects that their participation was anonymous.

This report reviews only results of a conducted study on moral development in Kuwait by using the Defining Issues Test (DIT). For several reasons the present report will not review results based on the Sentence Completion Test (SCT), and the Sociomoral Measure-Short Form (SRM-SF) data.

Reasons for not including Loevinger's SCT and Gibbś SRM-SF data in the present paper:

The scoring of the responses to the Loevinger's Sentence Completion Test (SCT) and to the Gibbs SRM-SF questionnaire and has to be done by highly trained scorers of whom very few are available. Furthermore, it is a very time consuming process. Mrs. K. Basinger, who was expected to score the SRM-SF questionnaires, underwent a serious eye operation and has subsequently become unavailable for scoring the questionnaires. Dr. Gielen is now training one student in the SRM-SF scoring procedures. In addition, he is now searching for a second person to be trained in the Loevinger (SCT) methodology.

Two papers will be prepared by 2004. The first paper -- to be finished by spring 2004 -- will contain a detailed statistical analysis of the SRM-SF data. The second paper -- to be completed by summer 2004 -- will focus on the SCT data and include both quantitative and qualitative analyses.

Sample of the study reported here:

Unlike, our previous study using the DIT in Kuwait (Gielen et al., 1992), where 84% of the total numbers of the respondents were Kuwaitis and the other 16% were from other Arab countries who resided in Kuwait at the time of the study, the sample of the present study included only Kuwaiti respondents.

Consistency Check

530 Muslim and Arabic speaker male and female students, aged between 13 and 25 years (Yet very few ages were older than 25 years), residing in Kuwait responded to the DIT and to background items. However, 110 students (or 26% of the total number of the respondents) give inconsistent or erratic responses to the DIT, and they were eliminated from further analysis.

A preliminary analysis of 420 completed research protocols (198 males and 222 females) from Kuwait shows that 87 respondents fall into the 13-16 year age range, 103 respondents in the 17-18 year old age range, 176 respondents in the 19-22 year age range, and 54 respondents are 23 or more years old (Table 1).

The remaining 420 subjects consisted of 198 males and 222 females intermediate, high school, college and university students aged between 13 and 25 years, with few subjects aged between 26 and 35 years. Table 1 illustrates the sample's characteristics.

Table 1: Sample classified by sex and age

Sex / Age groups	Males	Females	Males and Females
13-16 years	57	30	87
17-18 Years	64	39	103
19-22 Years	54	122	176
23+ years	23	31	54
Total	198	222	420

Mean of age= 18.85 years.

Results

Distribution of Moral Stage Scores

Table 2, depicts the distribution of DIT moral stage scores for the 420 students in this study. The results for the Kuwaiti students indicated that they preferred conventional moral arguments (stages 3 and 4=52.06) over principled arguments (stages 5A, 5B, and 6=31.08) and preconventional arguments (stage 2=7.13). It should be noted in this context that the DIT contains more conventional than principled or preconventional arguments. The obtained results, therefore, reflect to a significant degree the overall construction of the DIT.

Insert Table 2 about here

Table 2 also shows that moral judgment development (represented by the P%-Scores or the principled morality) increases by the increasing of age. The differences between age groups were significant, except in the

case of the difference between 17-18 years group and the 19-22 years group.

Table 3, contains the distributions of the DIT moral stage scores separately for the males and females sample (n=198 and n=222) and for the combined sample (n=420) the overall distribution for the male and female subjects are highly similar to each other. Results of the present study showed that males outnumbered female on stages 2, 3, 5A while females scored higher on stages 4, 5B, 6. P%-Scores (principled morality) were almost identical in both sexes. Yet these gender differences did not reach the levels of significance and were relatively unimportant. These findings agree fairly well with conclusions that Walker, 1984, 1991 and Thoma, 1986 (cited in Gielen et al., 1994) arrived at in their surveys of gender differences in structurally oriented moral reasoning (thinking) studies. Thoma analyzed a large number of studies employing the DIT and found that gender differences, although very small in size, tended to favor females over males.

Insert Table 3 about here

A comparison with older data based upon Arab respondents to the DIT (for instance, those published by Ahmed et al., 1987; Gielen et al., 1992; Kamel and El-Shouny, 2000; Ghusani and El-Hassen, 2001) suggests the following conclusions:

A significantly higher percentage of the subjects passed the consistency/validity checks in 2003 than in 1992 (Data was collected in 1991). This finding suggests greater familiarity of the respondents in

2003 with both the general testing format of the DIT and the moral arguments presented in the DIT.

The average P%-Score in the present study (31.08%) is higher than the score typically found in other studies conducted in the Arab world (e.g., 26.63% for Sudan in 1987 (Ahmed et al., 1987); 26.57% for Kuwait in 1992 (Gielen et al., 1992); 24.42% found for Egypt in 2000; 21.3% for Saudi Arabia in 2000 (Kamel and El-Shouny, 2000), and 24.3% - 28.9% for Lebanon (Ghusani and El-Hassan, 2001)

There small but statistically highly significant correlations between educational level and P%-Score ($r = .176$, $p < .01$) and between age and P%-Score ($r = .166$, $p < .01$).

Taken together the findings suggest that the Kuwaiti sample in the present study displayed a higher moral maturity level than was found for Kuwaiti respondents in 1992 (Gielen et al., 1992). In addition, modest but meaningful positive associations were found between increasing educational level and moral maturity.

To sum up, the results of the present study provide support of the use of the DIT as a measure of cross-cultural moral thinking development. Results of analysis of variance suggest that the Arabic version of the DIT reflects moral development stages as interpreted by Kohlberg (1964, 1984), and discriminates well between high and low scores. Significant and positive relationships are found between both age and educational level on one hand, and the level of the moral thinking development (As represented by P%-Score, principled morality) on the other hand. Kuwaiti students in the present study, generally, showed

higher P%-Scores (or principled morality) compared with their peers who were participated in an early similar study conducted in Kuwait in 1991 and published in 1992 (Gielen et al., 1992). Moreover, Kuwaiti students in the present study showed higher P%-Scores compared with similar age and educational levels Arab respondents who participated in several Arab previous studies by using the DIT such as the studies of Ahmed et al., 1987 in Sudan; Bouhmama, 1984, 1988, 1989 in Algeria; Kamel and El-Shouny, 2000 in Saudi Arabia; Ghusani and El-Hassan, 2001 in Lebanon; Eissa, 1985, 1993; El-Sheikh, 1985; Habib, 1991; Ibrahim, 1992; and Shafie, 1994 in Egypt.

It is very important to see the possible correlations between the performance of present study subjects on the DIT, and their performance on the other two tests (SRM-SF and SCT) which have been administered in the present study and their results are not included in the present report. These possible correlations will be discussed in later reports upon the completion of coding process and statistical analysis of the data based on SRM-SF and SCT.

Further research on the development of moral judgment is still very much needed. For example, the relationship between the development of moral judgment and some other related variables (such as self-esteem, religiosity, and I. E. locus of control) should be sufficiently investigated. Moreover, previous research studies have focused on school and university students, the future research studies should be directed to employ samples other than school and university students. Samples of the suggested future research on the development of moral judgment have to include subjects of different age, gender, socioeconomic status, and educational and vocational levels.

References

Al-Ghamdi, H. A. (2001). The relationship between moral reasoning development and ego-identity formation in a sample male adolescents and youth in the west area in the Saudi Arabia. The Egyptian Journal for Psychological Studies (Egypt), 8 (29), 221-255 (in Arabic).

Ahmed, R. A. & Gielen, U. P. (2002). A critical review of studies on moral judgment development using the Defining Issues Test in Arab countries. Arab Journal for the Humanities (Kuwait), 20 (77), 261-281.

Ahmed, R. A., Gielen, U. P., & Avellani, J. (1987). Perceptions of parental behavior development of moral reasoning in Sudanese Students. In C. Kagitcibasi (Ed.). Growth and progress in cross-cultural psychology. (pp. 196-206) Lisse (The Netherlands): Swets and Zeitlinger.

Ben Abdel-Aziz, B. H. S. (1989). Level of moral judgment in male and female students at the University of Umm el-Qura, Mecca. Unpublished master's thesis, University of Umm el-Oura, Saudi Arabia (in Arabic).

Bouhmama, D. (1984). Assessment of Kohlberg's stages of moral development in two cultures. Journal of Moral Education, 13 (2), 124-132.

Bouhmama, D. (1988). The relation of formal education to moral judgment development. The Journal of Psychology, 122 (2), 155-158.

Bouhmama, D. (1989). Moral judgment of a sample of psychology students at the University of Oran (Algeria). The Educational Journal (Kuwait), 6(21), 107-133 (in Arabic).

Comunian, A. L. and Gielen, U. P. (1995). Moral reasoning and prosocial action in Italian culture: Journal of Social Psychology. 135(6), 699-706.

Comunian, A. L. and Gielen, U. P. (2000). Sociomoral reflection and prosocial and antisocial behavior: Two Italian studies. Psychological Reports, 87, 161-175.

Eissa, M. R. (1985). The relationship between higher education and the level of moral judgment using a selected sample of students in the College of Education, Tanta University, Egypt. Journal of the Social Sciences (Kuwait), 13(2), 117-131 (in Arabic).

Eissa, M. R. (1993). The relationship between level of moral judgment and ego-strength. The Educational Journal (Kuwait), 8 (28), 61-115 (in Arabic).

El-Sheikh, S. A. (1985). A study of moral thinking in Egyptian adolescents and adults. In F.A. Abou Hatab (Ed.), Yearbook of Psychology (Vol. 4), (pp. 123-169), Cairo: The Anglo-Egyptian Bookshop (in Arabic).

Gibbs, J. C., Basinger, K. S., and Fuller, M. (1992). Moral maturity: Measuring the Development of Sociomoral Reflection. Hillsdale, NJ: Erlbaum.

Ghusani, R. and El-Hassan, K. (2001). Measuring the moral judgment. In R. Ghusani (Ed.), Values and education: The Third Yearbook (pp. 229-264), Beirut (Lebanon). The Lebanese Association for Educational Studies (in Arabic).

Gielen, U. P. (1996a). Moral reasoning in cross-cultural perspective: A review of Kohlbergian research. World Psychology, 2 (3-4), 313-333.

Gielen, U. P. (1996b). Focus on Lawrence Kohlberg. World Psychology. 2(3-4), I-VIII.

Gielen, U. P.; Ahmed, R. A. Avellani, J. (1992). The development of moral reasoning and perceptions of parental behavior in students from Kuwait. Moral Education Forum, 17(3), 20-37.

Gielen, U. P.; Comunian, A. L. and Antoni, G. (1994). An Italian cross-cultural study of Gibbs' sociomoral reflection measure short form. In A.L. Comunian and U.P. Gielen (Eds.), Advancing psychology

and its application: International perspectives (pp. 125-134), Milan (Italy): Franco Angeli.

Gielen, U. P.; Markoulis, D. C. (2001). Preference for principled and moral reasoning: A developmental and cross-cultural perspective. In L.L. Adler and U.P. Gielen (Eds.), Cross-cultural topics in psychology, 2nd ed. (pp. 81-101), Westport, CT: Praeger.

Gielen, U. P.; Miao, E. S. C. Y., & Avellani, J. (1990). Perceived parental behavior and the development of moral reasoning in students from Taiwan. Proceedings of CCU-ICP International Conference: Moral Values and Moral Reasoning in Chinese Societies, 1990, (pp. 464-506), Taipei: Chinese Culture University.

Habib, M. A. (1991). The relative importance of the characteristics of moral development: A social-cognitive approach for solving male and female adolescents' problems." Psychological and Educational Measurement and Evaluation (Gaza Strip, Palestine), 1(3), 45-92 (in Arabic).

Ibrahim, I. A. A. (1992). The relationship between level of moral judgment and level of dogmatism in samples of secondary school students. Unpublished M.A. thesis, Tanta University, Egypt, (in Arabic).

Ismail, M. A. (1976). A cross-cultural study of moral judgment: The relationship between American and Saudi Arabian university students on the Defining Issues Test. UMI, 77-5105, 1976.

Jessup, S. E. B. (1984). "The effects of knowledge of moral development stage theory and a request to fake on the Defining Issues Test scores of adolescents in a multi-national American school in Kuwait". DAI 45, 158A-159.

Kamel, M. M. and El-Shouny, M. E. (2001). Levels of moral judgment in male and female university students: A cross-cultural study on Egyptian and Saudi samples. Journal of Psychology (Egypt), 14(56), 106-130 (in Arabic).

Kohlberg, L. (1964). The development of moral character and ideology. In M. L. Hoffman and L. W. Hoffman (Eds.) Review of child development research. Vol. 1 (pp. 383-431). New York: Russell Sage Foundation.

Kohlberg, L. (1984). Essays on moral development, Vol. II: The psychology of moral development. New York: Harper & Row.

Loevinger, J. (1976). Ego development. San Francisco, CA: Jossey-Bass.

Naito, T.; Wen-Ying, L. and Gielen, U. P. (2001). Moral development in East Asian societies: A selective review of the cross-cultural literature. Psycholgia: An International Journal of Psychology in the Orient. 44 (2), 148-160.

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

Rest, J.; Narvaez, D.; Bebeau, M. J. and Thoma, S. J. (1999). Postconventional moral thinking: A neo-Kohlbergian approach. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Shafie, A. M. (1994). [Comparing] the Moral judgment of adolescents with general education and with Al-Azhar education. Unpublished M.A. thesis, Institute of Higher Studies on Childhood, Ain Shams University, Egypt (in Arabic).

Webb, R. J., and Steentsma. I. R. (1983). A study of moral reasoning by Palestinian Arabs. Paper presented at the 91st Annual Convention of the American Psychological Association, Anaheim, California, U.S.A.

Table 2: Distribution of DIT Stage Scores for Kuwaiti Adolescents and Adults

Stages / Age	N	Stage 2		Stage 3		Stage 4		Stage 5A		Stage 5B		Stage 6		P%		K		R	
		x	SD	x	SD	x	SD	x	SD	x	SD	x	SD	x	SD	x	SD	x	SD
13-16 yrs	87	8.76	8.18	24.50	9.92	28.99	9.19	15.65	8.03	5.10	4.63	6.69	5.02	27.24	11.40	4.62	2.59	5.79	4.13
17 – 18 yrs	103	6.77	4.93	23.43	9.17	27.82	8.52	18.33	7.64	6.24	4.58	7.02	5.57	31.44	9.72	4.37	2.55	6.05	4.93
19 – 22 yrs	176	6.73	5.45	20.83	9.07	31.23	9.80	17.72	7.49	6.57	4.52	7.51	5.35	31.70	9.75	3.38	2.98	5.92	4.56
23+ yrs	54	6.51	5.57	18.95	8.14	32.47	9.24	19.20	5.60	7.68	5.12	7.25	5.52	34.23	8.27	3.37	2.70	4.57	4.22
Total	420	7.13	6.06	21.99	9.32	30.07	9.43	17.63	7.49	6.33	4.68	7.19	5.35	31.08	10.11	3.88	2.81	5.75	4.54

Table 3: Distribution of DIT Stage Scores for Kuwaiti Males and Females

Stages / Age	N	Stage 2		Stage 3		Stage 4		Stage 5A		Stage 5B		Stage 6		P%		K		R	
		x	SD	x	SD	x	SD	x	SD	x	SD	x	SD	x	SD	x	SD	x	SD
Males	198	7.24	5.94	23.13	9.61	29.36	10.01	18.31	7.27	5.99	4.58	6.78	5.06	31.10	9.61	3.81	2.77	5.34	4.57
Females	222	7.03	6.17	20.97	8.96	30.70	8.85	17.03	7.64	6.64	4.76	7.55	5.59	31.07	10.56	3.94	2.85	6.12	4.49
Total	420	7.13	6.06	21.99	9.32	30.07	9.43	17.63	7.49	6.33	4.68	7.19	5.35	31.08	10.11	3.88	2.81	5.75	4.54