

Running head: EVALUATING THE BEHAVIORAL DIT

Measuring Moral Reasoning using Moral Dilemmas: Evaluating Reliability, Validity, and  
Differential Item Functioning of the Behavioral Defining Issues Test (bDIT)

Youn-Jeng Choi\*      Hyemin Han\*      Kelsie J. Dawson

Stephen J. Thoma      Andrea L. Glenn

University of Alabama

Author Note

Youn-Jeng Choi, Educational Research Program, University of Alabama, Tuscaloosa, AL  
35487, USA.

Hyemin Han, Kelsie J. Dawson, and Stephen J. Thoma, Educational Psychology Program,  
University of Alabama, Tuscaloosa, AL 35487, USA.

Andrea L. Glenn, Center for the Prevention of Youth Behavior Problems, University of  
Alabama, Tuscaloosa, AL 35487, USA.

Correspondence concerning this article should be addressed to Hyemin Han, Educational  
Psychology Program, University of Alabama, Tuscaloosa, AL 35487, USA.

Contact: [hyemin.han@ua.edu](mailto:hyemin.han@ua.edu). Tel: 1-205-348-0746. Fax: 1-205-348-1191.

\*: These authors contributed equally to this work and are listed alphabetically.

**Measuring Moral Reasoning using Moral Dilemmas: Evaluating Reliability, Validity, and  
Differential Item Functioning of the Behavioral Defining Issues Test (bDIT)**

## **Abstract**

We evaluated the reliability, validity, and differential item functioning (DIF) of a shorter version of the Defining Issues Test-1 (DIT-1), the behavioral DIT (bDIT), measuring the development of moral reasoning. 353 college students (81 males, 271 females, 1 not reported; age  $M = 18.64$  years,  $SD = 1.20$  years) who were taking introductory psychology classes at a public University in a suburb area in the Southern United States participated in the present study. First, we examined the reliability of the bDIT using Cronbach's  $\alpha$  and its concurrent validity with the original DIT-1 using disattenuated correlation. Second, we compared the test duration between the two measures. Third, we tested the DIF of each question between males and females. Findings reported that first, the bDIT showed acceptable reliability and good concurrent validity. Second, the test duration could be significantly shortened by employing the bDIT. Third, DIF results indicated that the bDIT items did not favour any gender. Practical implications of the present study based on the reported findings are discussed.

*Keywords:* moral reasoning, moral development, Defining Issues Test, Differential Item Functioning, gender difference

## **Introduction**

Moral reasoning refers to a psychological process that enables one to decide what one ought to do, morally, based on values or sets of standards (Richardson, 2013). This process is necessary to identify appropriate moral behavioural solutions within a dilemmatic situation, and developed moral reasoning is positively associated with moral behaviour according to previous research (Rest & Narvaez, 1994). Neo-Kohlbergian framework, which is a mainstream theoretical framework in the field of moral development, explains that the development of moral

reasoning is associated with whether one employs sophisticated moral philosophical rationale while making moral decisions (Han, 2014; Rest, Narvaez, Bebeau, & Thoma, 1999).

According to Neo-Kohlbergian framework, there are three schemes of moral reasoning based on different moral perspectives: personal interest (PI), maintaining norms (MN), and post-conventional (PC) schemes. The degree of one's moral reasoning development can be explained by which schema is most preferred during moral decision-making. Endorsing the PI schema is related to a tendency to benefit one's own personal interest or relationship. Endorsing the MN schema is associated with an intention to maintain social norms and laws. The endorsement of the PC schema, the most sophisticated schema, results in the endorsement of human rights and universal moral principles, and critical evaluation of laws and social norms during moral decision-making (Rest, Narvaez, Bebeau, et al., 1999). Previous research has shown that the PI schema is most likely to be preferred during childhood and the MN schema during adolescence, and the development of the PC schema occurs during the transition from adolescence to adulthood (Rest, Narvaez, Bebeau, et al., 1999).

Traditionally, the Defining Issues Test-1 (DIT-1) has been utilized to assess one's moral reasoning based on Neo-Kohlbergian framework. The DIT-1 measures whether participants can make a moral judgment based on more sophisticated moral philosophical perspectives, particularly the PC schema, instead of the PI or MN schema (Rest, Narvaez, Bebeau, et al., 1999). A P-score, which is the likelihood of the endorsement of the PC schema while solving presented moral dilemmas in the DIT-1, has been used as an indicator of one's moral reasoning ability (Rest & Narvaez, 1994). The measurement for moral reasoning, the DIT-1, has been widely utilized in the fields of moral psychology and education, and its reliability and validity has been established using accumulated large datasets (Rest & Narvaez, 1994).

Recent research about measures for contextual moral reasoning suggests that the DIT possesses good ecological validity as well. Moral psychologists have developed the Intermediate Concepts Measures (ICMs), which measure moral reasoning within more specific contexts (Thoma, Derryberry, & Crowson, 2013). Unlike the DIT, which was initially invented to assess universal moral reasoning with hypothetical moral philosophical dilemmas, the ICMs were designed to assess contextual moral reasoning with concrete dilemmas; for example, different versions of ICMs were invented to measure moral reasoning within the context of dentistry or school life (Thoma et al., 2013). Previous research has shown universal moral reasoning assessed by the DIT and contextual moral reasoning assessed by the ICMs have at least moderate correlation (Thoma et al., 2013); this may suggest that the DIT possesses good ecological validity across different situations and contexts.

However, several issues related to the measurement have been discussed. First, a P-score is calculated from rating and ranking participants' responses (Rest, 1990). Such a complex scoring method makes it difficult to use the DIT-1 in behavioural experiments that aim to measure immediate responses from participants. Second, it takes at least twenty minutes to complete the test, so the long test duration could be a practical issue (Latif, 2001; Teal & Carroll, 1999). Third, Gilligan (1982) stated that the reasoning-based approach might devalue the orientation of care, which is frequently endorsed by women, and so the assessment of moral developmental level could be gender-biased. Although Thoma (1986) reported that there was no meaningful difference in the mean P-score as a quantified DIT score between males and females, such a difference has not been examined at the item level (e.g., whether items and texts favour a certain gender group).

The bDIT was developed by simplifying the DIT-1 format to address the aforementioned issues (Han, Dawson, Thoma, & Glenn, 2019). It was designed to be suitable for experiments assessing participants' behavioural responses, such as reaction time. The bDIT simply requires participants to select one of presented behavioural options within a relatively shorter timeframe, unlike the DIT-1 that requires individuals to complete complicated scoring tasks. Moreover, its simpler structure enables us to evaluate its psychometrical qualities at the item level. Although the DIT-2, a revised version of the DIT-1, has been developed (Rest, Narvaez, Thoma, & Bebeau, 1999), the bDIT was created from the DIT-1. The DIT-2 is mainly used to calculate an N2-score, which is more complicated to calculate compared with a P-score and could not be simply quantified from individual selected options (Rest, Narvaez, Thoma, et al., 1999). Because of this, it would not be appropriate to utilize the DIT-2 to develop a simplified test for behavioural experiments.

In the present study, we evaluated the psychometrical qualities of the bDIT to examine whether it can address the aforementioned issues. First, we tested its reliability and validity to examine whether it can be used to measure the development of moral reasoning within the contexts of behavioural experiments similar to the DIT-1. Second, we measured participants' response time to see whether the bDIT can save time. Third, we examined the DIF to explore gender bias in participants' responses.

## **Method**

### **Participants**

We analysed a dataset collected with computer survey forms (Han et al., 2018). A total of 353 American college students (81 males, 271 females, 1 not reported; age  $M = 18.64$  years,  $SD = 1.20$  years) completed the survey forms. All participants were taking undergraduate level

introductory psychology classes at a public University in a suburb area in the Southern United States. They were registered in the psychology subject pool for recruitment. Among them, 245 completed a Qualtrics survey online at their own convenient time and location (46 males, 199 females; age  $M = 18.51$  years,  $SD = .92$  years), and 108 visited our lab to complete a behavioural experiment programmed with E-Prime 2.0 in a lab space without anyone else present (35 males, 72 females, 1 not reported; age  $M = 19.20$  years,  $SD = 1.33$  years, 31 not reported) (Psychology Software Tools, 2016). Collected data is shared via the Open Science Framework at <https://osf.io/jtk42/>.

## **Measures**

### **Behavioral DIT**

We presented the bDIT consisting of three stories to each participant (See supplementary materials for the survey form and further methodological details). We calculated the P-score as well as the percentage scores of the PI and MN subscales for DIF tests. The bDIT was presented via Qualtrics to online participants, and via E-Prime to in-lab participants. Behavioural responses of the in-lab participants were recorded by a keyboard.

### **DIT-1**

We administered the original DIT-1 to examine the concurrent validity of the bDIT (Rest, 1990). Each participant's P-score was calculated from the three-story version of the DIT-1. Invalid responses were screened out following the scoring manual and were excluded from correlation analysis (Rest, 1990). Online participants were presented with the DIT-1 via Qualtrics, and in-lab participants were given hard copies. During online survey sessions, the bDIT and DIT-1 forms were presented in a randomized order. During lab experimental sessions, the bDIT was administered first and then the DIT-1 was administered.

## Analysis

### Reliability, validity, test duration, mode effect, and association with demographics

The basic analyses of the collected data were performed with STATA 14 and R. The reliability of the DIT-1 was estimated by Cronbach's  $\alpha$ . Because the measured variables of the bDIT were dichotomous, we calculated the  $\alpha$  value based on tetrachoric correlation with a customized R script following Napolitano, Callina, and Mueller (2013) (the script is available via <https://osf.io/jtk42/>). For the concurrent validity check, we examined the correlation between P-scores measured by the bDIT and DIT-1. We calculated a disattenuated correlation coefficient to remove possible measurement error (Osborne, 2003). It was calculated as follows:

$$r_{xy}^* = \frac{r_{xy}}{\sqrt{r_{xx}r_{yy}}} \quad (r_{xy}^*: \text{disattenuated correlation coefficient}; r_{xy}: \text{original correlation}; r_{xx}, r_{yy}:$$

measurement reliability)

Moreover, the test duration data was collected from the 108 responses recorded by E-Prime 2.0. This data could not be collected from online samples due to technical restrictions and uncountability of participants' behaviour. We compared the mean test duration with the previously reported minimum test duration of the DIT-1, twenty minutes (Latif, 2001), by performing one-sample t-test. The test mode effect was examined by performing ANOVA while setting the bDIT score as the dependent variable and the test type (online vs. in-lab experiment) as the independent variable following the method used by (Clariana & Wallace, 2002).

Furthermore, since a previous study reported statistical significance among the association between demographical variables, gender and age, and DIT outcomes, we performed ANOVA to examine this association (Caravita, Giardino, Lenzi, Salvaterra, & Antonietti, 2012). The bDIT or DIT-1 P score was used for the dependent variable, and gender or age was used for

the independent variable. For the DIT-1 score analysis, we excluded responses that did not pass the screening procedure.

### **Differential item functioning analysis**

The validity of the bDIT was examined by Differential item functioning analysis (DIF) using Mantel-Haenszel chi-square test (MH- $X^2$ ) and logistic regression (LR) using SAS 9.4 (De Ayala, 2009). AERA, APA, and NCME (2014) introduced the sources of validity evidence and determined that DIF is the psychometric method to find the validity evidence based on internal structure. DIF happens when different groups (e.g., gender, ethnics, culture, etc.) of participants with similar ability have different responses on a particular item. Although the framework of DIF has been mainly utilized in the field of educational tests, it can also be applied to studies in human development to examine the potential group-biasedness of a developmental measure (Bingenheimer, Raudenbush, Leventhal, & Brooks-Gunn, 2005). There are two types of DIF: nonuniform and uniform. Nonuniform DIF exists when the difference in probabilities is not uniform across ability levels. Uniform DIF exists when the difference in probabilities of success is uniform for the two groups across all ability levels. The MH- $X^2$  statistic is used to detect only uniform DIF, and LR DIF is used to detect both uniform and nonuniform DIF (Camilli & Shepard, 1994). Following Holland and Thayer (1988) and Mantel and Haenszel (1959), we can calculate MH- $X^2$  and  $\alpha_{MH}$  (see supplementary methods).

The application of LR to DIF analysis requires model comparisons using likelihood ratio test ( $\Delta G^2$ ):

$$\Delta G^2 = -2 \ln \left( \frac{L_R}{L_F} \right) = -2 \ln(L_R) - (-2 \ln(L_F))$$

where  $L_r$  and  $L_f$  are the maximum likelihood for the reduced model and for the full model, respectively and the  $df$  for evaluating the significance of  $\Delta G^2$  is the difference in the number of parameters in the full and the reduced model (see supplementary methods).

## Results

### Reliability, Validity, Mode Effect, and Association with Demographics

Descriptive statistics and calculated  $\alpha$  values are presented in Table 1. A correlational coefficient between bDIT's and DIT-1's P-scores was .71 before disattenuation and .87 after disattenuation. The mean test duration recorded by E-Prime 2.0 was 6.52 minutes ( $N = 108$ ,  $SD = 2.06$  minutes, median = 6.13 minutes), which is significantly shorter than the minimum duration of the DIT-1 reported in the previous review article (Latif, 2001), twenty minutes,  $t(107) = -67.86$ ,  $p < .001$ ,  $d = -6.53$ .

*< Table 1 about here >*

No significant main effect of test type was found, so no significant test mode effect was found either,  $F(1, 351) = 2.74$ ,  $p = .10$ ,  $\omega^2 = .00$ . The main effects of demographical factors on test scores were also non-significant. First, gender was not significantly associated the DIT-1 score,  $F(1, 248) = .33$ ,  $p = .56$ ,  $\omega^2 = .00$ , nor the bDIT score,  $F(1, 350) = .00$ ,  $p = .95$ ,  $\omega^2 = .00$ . Second, there was no significant main effect of age on either the DIT-1 score,  $F(1, 220) = 1.76$ ,  $p = .19$ ,  $\omega^2 = .00$ , or bDIT score,  $F(1, 317) = .69$ ,  $p = .41$ ,  $\omega^2 = .00$ .

### Differential Item Functioning

To detect DIF items, we applied a multiple comparison correction using the Bonferroni correction method to control inflated Type I error. Table S2 shows the MH- $X^2$  DIF results. No DIF item was found when Bonferroni correction was applied ( $\alpha = .002$ ). Similarly, when LR

DIF was applied, no DIF item was detected using  $\alpha = .002$  (see Tables S3 and S4). Furthermore, the practical effects of DIFs are negligible because of their small effect size (De Ayala, 2009).

### **Discussion**

We evaluated the reliability, validity, and DIF of the bDIT. First, the bDIT showed acceptable reliability and good concurrent validity with the original DIT-1. Second, the test duration became significantly shorter. Third, the test mode or demographics were not significantly associated with the outcome scores. Fourth, the DIF demonstrated that the bDIT did not especially favour one gender group, males in particular, despite concerns raised by Gilligan (1982) about developmental methods derived from the Kohlberg model. These results suggest that the bDIT can be used in future studies that require recording participants' immediate behavioural responses during a shorter test period due to practical issues. The bDIT will provide moral educators and researchers with a reliable, valid, and practical tool to assess participants' behavioural outcomes associated with moral reasoning within the contexts of educational program evaluations and developmental studies.

However, several limitations will need to be addressed by future research. First, other types of reliability (e.g., test-retest reliability) and validity (e.g., predictive validity) indicators should also be tested. Second, in addition to the gender difference, other types of group differences, such as cross-cultural differences, may need to be examined to validate the use of the bDIT across diverse groups. Third, although we examined differences in responses between two genders, we collected the data from more female participants due to the nature of the subject pool, the psychology subject pool, so male participants were underrepresented in the present study. Related to this point, all participants were college students, so the age range of the sample was also limited, although the DIT has been developed to assess moral reasoning across diverse

age groups. Since the main purpose of the DIT is to study the development of moral reasoning, future studies with the bDIT should be conducted with more diverse age groups. Fourth, we could not measure the duration of the DIT-1 in the lab experiment because it was administered as a paper-and-pencil test. Thus, we had to compare the duration of the bDIT with the minimum required timeframe for the DIT-1 in the previous review article (Latif, 2001), instead of the real DIT-1 duration, so the result of the duration comparison should be interpreted with caution. Fifth, given that the DIT employs philosophical and hypothetical dilemmas, its ecological validity is a possible issue. Although previous research showed a significant association between the P-score, ICM score, and moral behaviour (Rest & Narvaez, 1994; Thoma et al., 2013), ideally, researchers should consider employing more realistic dilemmas in their future research to address this issue. In addition, as an alternative, the comparison between the bDIT and ICM can also be considered.

### References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Bingenheimer, J. B., Raudenbush, S. W., Leventhal, T., & Brooks-Gunn, J. (2005). Measurement equivalence and differential item functioning in family psychology. *Journal of Family Psychology, 19*(3), 441–455. doi:10.1037/0893-3200.19.3.441
- Camilli, G., & Shepard, L. A. (1994). *Methods for Identifying Biased Test Items*. Thousand Oaks, CA: Sage.
- Caravita, S. C. S., Giardino, S., Lenzi, L., Salvaterra, M., & Antonietti, A. (2012). Socio-economic factors related to moral reasoning in childhood and adolescence: the missing link between

- brain and behavior. *Frontiers in Human Neuroscience*, 6. doi:10.3389/fnhum.2012.00262
- Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, 33(5), 593–602. doi:10.1111/1467-8535.00294
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford Press.
- Gilligan, C. (1982). *In a different voice*. Harvard University Press. Cambridge, MA: Harvard University Press. doi:10.2307/2067520
- Han, H. (2014). Analyzing theoretical frameworks of moral education through Lakatos's philosophy of science. *Journal of Moral Education*, 43(1), 32–53. doi:10.1080/03057240.2014.893422
- Han, H., Dawson, K. J., Thoma, S. J., & Glenn, A. L. (2019). Developmental level of moral judgment influences behavioral patterns during moral decision-making. *Journal of Experimental Education*. doi:10.1080/00220973.2019.1574701
- Holland, P. W., & Thayer, D. T. (1988). Differential item functioning and the Mantel–Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test Validity* (pp. 129–145). Hillsdale, NJ: Erlbaum.
- Latif, D. A. (2001). The relationship between ethical reasoning and the perception of difficulty with ethical dilemmas in pharmacy students: Implications for teaching professional ethics. *Teaching Business Ethics*, 5(1), 107–117. doi:10.1023/A:1026502902003
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–748.

- Napolitano, C. M., Callina, K. S., & Mueller, M. K. (2013). Comparing alternate approaches to calculating reliability for dichotomous data: The sample case of adolescent selection, optimization, and compensation. *Applied Developmental Science, 17*(3), 148–151.  
doi:10.1080/10888691.2013.804372
- Osborne, J. W. (2003). Effect sizes and the disattenuation of correlation and regression coefficients: lessons from educational psychology. *Practical Assessment, Research & Evaluation, 8*(11), 1–5.
- Psychology Software Tools. (2016). E-Prime 2.0. Pittsburgh, PA: Psychology Software Tools.
- Rest, J. R. (1990). *DIT: Manual for the Defining Issues Test*. Minneapolis, MN: University of Minnesota Center for the Study of Ethical Development.
- Rest, J. R., & Narvaez, D. (1994). *Moral development in the professions: Psychology and applied ethics*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Rest, J. R., Narvaez, D., Bebeau, M., & Thoma, S. (1999). A Neo-Kohlbergian approach: The DIT and schema theory. *Educational Psychology Review, 11*(4), 291–324.  
doi:10.1023/a:1022053215271
- Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology, 91*(4), 644–659.  
doi:10.1037/0022-0663.91.4.644
- Richardson, H. S. (2013). Moral reasoning. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from <http://plato.stanford.edu/entries/reasoning-moral/>
- Teal, E. J., & Carroll, A. B. (1999). Moral Reasoning Skills: Are Entrepreneurs Different? *Journal of Business Ethics, 19*(3), 229–240. doi:10.1023/A:1006037510932

The Center for the Study of Ethical Development. (2018). The Center for The Study of Ethical Development. Retrieved from <http://ethicaldevelopment.ua.edu/>

Thoma, S., Derryberry, W. P., & Crowson, H. M. (2013). Describing and testing an intermediate concept measure of adolescent moral thinking. *European Journal of Developmental Psychology, 10*(2), 239–252. doi:10.1080/17405629.2013.766129

Thoma, S. J. (1986). Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review, 6*(2), 165–180. doi:10.1016/0273-2297(86)90010-9

## Tables

Table 1. Descriptive statistics and Cronbach's  $\alpha$

	<i>N</i>	<i>M</i>	<i>SD</i>	Reliability*
bDIT (P-score)	353	49.15	20.25	.74
DIT-1 (P-score)	250**	35.29	19.06	.79

Note: \* The reliability of the bDIT was estimated with tetrachoric correlation. The DIT-1

reliability was estimated with ordinary Cronbach's  $\alpha$ . \*\* The data collected from 103

participants (86/245 online and 17/108 in-lab participants) were excluded from the DIT-1 data

analysis because they could not pass the DIT-1 screening procedure

## Supplementary Materials

### Supplementary Methods

#### *Behavioral DIT*

The bDIT contained three moral dilemmas, i.e., Heinz and the drug, Escaped prisoner, and Newspaper, imported from the three-story (short) version DIT-1; the full version DIT-1 consists of six stories: Doctor, Webster, Students, and the aforementioned three stories. The aforementioned three stories were selected and tested in prior studies to shorten the length of the form while minimizing the decrease in reliability and validity (The Center for the Study of Ethical Development, 2018).

First of all, each dilemma story was presented to participants. After reading each dilemma, the participants were asked to make a behavioural decision. Then, they were presented with eight questions asking the rationale of their behavioural decision-making for each story (24 questions total). For each question, participants were asked to select the most important philosophical criterion among three options. One option represented the PI schema, one the MN schema, and one the PC schema (see Han, Dawson, Thoma, & Glenn (2019) and its supplementary materials for further details about how questions and criterion were imported from the original DIT-1). A P-score (in percentage) was calculated from how many PC schema options were selected out of 24 questions.

We selected two philosophical rationale items for each schema in each story based on the correlation between each item's score and total schema level score. While selecting items, we selected two items that showed the greatest correlation with the related schema level score per schema per story. This correlation analysis was performed by the Center for the Study of Ethical Development with the large DIT-1 dataset collected from 58,449 participants. Table S6 reports

the correlation between each item's score and total schema score. Selected items are highlighted for readers' information.

For instance, in the case of the post-conventional level in "Heinz and drug," we selected "what values are going to be the basis for governing how people act towards each other ( $r = 35$ )" and "whether the law in this case is getting in the way of the most basic claim of any member of society ( $r = 31$ )," but excluded "would stealing in such a case bring about more total good for the whole society or not ( $r = 24$ )" from the bDIT. We applied the same selection processes to other schema levels and other stories.

By presenting one of the two items at each schema level in each question, we created 8 questions per story, 24 questions in the whole test (3 stories  $\times$   $2C_1 \times 2C_1 \times 2C_1$ ) = 24 questions). Here is one same question from "Heinz and drug" dilemma:

*Which issue is more important in making a decision?*

- Isn't it only natural for a loving husband to care so much for his wife that he'd steal? → **PI item**
- Whether a community's laws are going to be upheld. → **MN item**
- What values are going to be the basis for governing how people act towards each other. → **PC item**

Items were presented in a random order in the test.

### ***Differential item functioning analysis***

Following Holland and Thayer (1988) and Mantel and Haenszel (1959), we can calculate MH- $X^2$  and  $\alpha_{MH}$  by

$$MH_{x^2} = \frac{\{|\sum_{t=1}^{(L-1)} [A_t - \varepsilon(A_t)] - 0.5|\}^2}{\sum_{t=1}^{(L-1)} var(A_t)}, \text{ where } var(A_t) = \frac{n_{R_t} n_{F_t} m_{1_t} m_{0_t}}{T_t^2 (T_t - 1)} \text{ and } \varepsilon(A_t) = \frac{n_{R_t} m_{1_t}}{T_t}$$

$$\alpha_{MH} = \frac{\sum_{t=1}^{(L-1)} \frac{A_t D_t}{T_t}}{\sum_{t=1}^{(L-1)} \frac{B_t C_t}{T_t}}$$

When  $\alpha_{MH}$  equals 1, no DIF on the item. When  $0 < \alpha_{MH} < 1$ , the item favours the focal group. A value of  $\alpha_{MH}$  greater than 1 indicates the reference group members performed better than comparable focal group members on the item (Holland & Thayer, 1988).

The application of LR to DIF analysis requires likelihood ratio test statistic to determine whether the full model differed significantly from the reduced model (De Ayala, 2009).

The three models for our study are as follows:

Model 1 (Full Model):  $\text{Logit}(p) = \tau_0 + \tau_1(X) + \tau_2(\text{Gender}) + \tau_3(X \times \text{Gender})$

Model 2 (1<sup>st</sup> Reduced Model):  $\text{Logit}(p) = \tau_0 + \tau_1(X) + \tau_2(\text{Gender})$

Model 3 (2<sup>nd</sup> Reduced Model):  $\text{Logit}(p) = \tau_0 + \tau_1(X)$

where  $p$  is the probability to get correct response of an item and  $X$  is a measure of an individual's ability parameter in IRT or sum of all items. In our study, total score is used. Gender = 0 for male (focal group) and Gender = 1 for female (reference group). The term  $\tau_1$  indicates the relationship between the performance on the item and the person's total score,  $\tau_2$  reflects the mean gender difference in performance on the item, and  $\tau_3$  shows the interaction between gender and total score.

There are two hypothesis tests to be tested using likelihood ratio test: 1)  $\tau_3 = 0$  for non-uniform DIF test and 2)  $\tau_2 = 0$  for uniform DIF test. For first hypothesis, Model 1 and Model 2

are compared and Model 2 and Model 3 are compared for second hypothesis. When the null hypothesis was rejected, the evidence of DIF exists.

### ***References***

Han, H., Dawson, K. J., Thoma, S. J., & Glenn, A. L. (2019). Developmental Level of Moral Judgment Influences Behavioral Patterns during Moral Decision-making. *Journal of Experimental Education*. doi:10.1080/00220973.2019.1574701

Holland, P. W., & Thayer, D. T. (1988). Differential item functioning and the Mantel–Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test Validity* (pp. 129–145). Hillsdale, NJ: Erlbaum.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–748.

The Center for the Study of Ethical Development. (2018). The Center for The Study of Ethical Development. Retrieved from <http://ethicaldevelopment.ua.edu/>

## Supplementary Tables

Table S1. Mantel-Haenszel Chi-square Results

	Personal Interest			Maintaining Norms			Post Conventional		
	MH- $X^2$	p value	$\alpha_{MH}$	MH- $X^2$	p value	$\alpha_{MH}$	MH- $X^2$	p value	$\alpha_{MH}$
Item 1	3.24	.07	.55	.49	.48	1.26	2.46	.12	1.54
Item 2	1.14	.29	.70	.93	.33	1.40	.93	.34	1.31
Item 3	.33	.56	.82	.20	.65	1.20	1.22	.27	1.37
Item 4	2.51	.11	.61	.08	.78	1.10	4.19	.04	1.83
Item 5	.90	.34	1.37	1.60	.21	1.50	2.51	.11	.65
Item 6	.03	.86	1.06	1.05	.31	1.38	.56	.46	.82
Item 7	.05	.82	.93	.09	.76	.90	1.75	.19	1.44
Item 8	.00	.98	.99	2.41	.12	.61	7.34	.01	2.18
Item 9	.01	.92	1.03	.06	.81	.92	.18	.67	1.13
Item 10	1.47	.23	.68	.34	.56	1.22	.37	.54	1.20
Item 11	.49	.48	1.27	.39	.53	1.24	.98	.32	.76
Item 12	.36	.55	1.21	1.03	.31	1.45	2.65	.10	.63
Item 13	.21	.65	.83	.78	.38	1.37	.11	.74	.90
Item 14	1.15	.28	.59	1.10	.29	.70	.78	.38	1.32
Item 15	.04	.83	.92	.29	.59	1.20	.68	.41	.78
Item 16	.38	.54	1.40	.27	.61	1.20	2.49	.11	.62
Item 17	1.31	.25	1.43	.75	.39	.76	.19	.66	.89
Item 18	6.55	.01	2.18	.51	.47	.80	3.33	.07	.60
Item 19	.93	.33	1.36	.13	.72	.89	1.98	.16	.65

Item 20	3.59	.06	1.83	2.90	.09	.59	.68	.41	.79
Item 21	.10	.75	1.16	.06	.81	.93	.19	.66	.89
Item 22	.89	.34	.69	.35	.55	.83	.17	.68	1.13
Item 23	7.09	.01	.29	.00	.97	1.01	.06	.81	1.07
Item 24	.02	.89	1.07	1.28	.26	.71	.04	.84	1.06

---

Table S2. Logistic Regression Analysis for Nonuniform DIF

	Personal Interest		Maintaining Norms		Post Conventional	
	$\Delta G^2$	$\Delta R^2$	$\Delta G^2$	$\Delta R^2$	$\Delta G^2$	$\Delta R^2$
Item 1	.15	.000	1.90	-.007	.00	.000
Item 2	1.26	-.004	.57	-.002	.68	-.002
Item 3	.18	-.001	.00	.000	.64	-.002
Item 4	.44	-.001	1.14	-.005	.00	.000
Item 5	2.22	-.008	.10	.000	.66	-.002
Item 6	3.93	-.014	.03	.000	.05	.000
Item 7	.05	.000	3.28	-.013	.09	.000
Item 8	3.33	-.012	.62	-.003	.10	.000
Item 9	.28	-.001	.03	.000	.26	-.001
Item 10	3.57	-.014	.12	.000	1.70	-.005
Item 11	.25	-.001	.04	.000	1.45	-.004
Item 12	1.20	-.005	2.97	-.007	1.79	-.005
Item 13	.26	-.001	2.99	-.009	.43	-.001
Item 14	.07	.000	.41	-.001	.96	-.003
Item 15	.03	.000	.69	-.002	.09	.000
Item 16	.00	.000	.03	.000	.55	-.002
Item 17	.28	-.001	.76	-.003	.70	-.002
Item 18	.45	-.002	.20	-.001	.64	-.002
Item 19	.00	.000	.00	.000	.31	-.001
Item 20	.06	.000	.01	.000	.36	-.001

Item 21	.19	-.001	.30	-.001	.06	.000
Item 22	.15	-.001	.55	-.002	.08	.000
Item 23	.06	.000	.73	-.003	1.28	-.005
Item 24	.00	.000	.65	-.002	.90	-.003

---

Table S3. Logistic Regression Analysis for Uniform DIF

	Personal Interest		Maintaining Norms		Post Conventional	
	$\Delta G^2$	$\Delta R^2$	$\Delta G^2$	$\Delta R^2$	$\Delta G^2$	$\Delta R^2$
Item 1	2.44	-.007	.40	-.002	4.02	-.013
Item 2	1.83	-.006	.88	-.003	2.09	-.007
Item 3	.10	.000	.37	-.002	1.88	-.006
Item 4	2.06	-.006	.39	-.002	4.56	-.015
Item 5	1.38	-.005	1.74	-.007	1.54	-.005
Item 6	.04	.000	1.51	-.006	.43	-.001
Item 7	.17	-.001	.12	.000	2.17	-.007
Item 8	.15	-.001	2.38	-.010	5.68	-.018
Item 9	.06	.000	.02	.000	.08	.000
Item 10	.77	-.003	.16	.000	.13	.000
Item 11	.79	-.003	.55	-.002	1.18	-.004
Item 12	.62	-.003	1.07	-.003	2.77	-.008
Item 13	.07	.000	1.22	-.004	.65	-.002
Item 14	.29	-.002	1.32	-.004	.49	-.002
Item 15	.00	.000	.48	-.001	.76	-.002
Item 16	.30	-.002	.19	.000	2.50	-.007
Item 17	2.17	-.008	.38	-.002	.08	.000
Item 18	8.41	-.031	.75	-.003	3.20	-.011
Item 19	1.58	-.006	.04	.000	.94	-.003
Item 20	4.13	-.016	2.16	-.008	.71	-.002

Item 21	.06	.000	.05	.000	.10	.000
Item 22	.66	-.003	.10	.000	.31	-.001
Item 23	4.07	-.025	.36	-.001	.09	.000
Item 24	.03	.000	.54	-.002	.02	.000

---

Table S4. Logistic Regression Coefficients for Personal Interest

	Model 1				Model 2			Model 3	
	$\tau_0$	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_0$	$\tau_1$	$\tau_2$	$\tau_0$	$\tau_1$
Item 1	-2.74**	.40**	-.24	-.04	-2.55**	.37**	-.49	-2.97**	.38**
Item 2	-2.85**	.43**	.31	-.12	-2.29**	.35**	-.41	-2.66**	.35**
Item 3	-2.69**	.41**	-.39	.05	-2.90**	.44**	-.10	-3.00**	.44**
Item 4	-2.70**	.42**	-.02	-.07	-2.38**	.37**	-.44	-2.76**	.38**
Item 5	-2.85**	.24**	-.77	.16	-3.74**	.35**	.41	-3.35**	.34**
Item 6	-4.74**	.50**	1.92	-.24	-3.17**	.31**	.06	-3.11**	.31**
Item 7	-2.91**	.33**	-.30	.02	-3.04**	.34**	-.13	-3.16**	.35**
Item 8	-4.29**	.48**	1.44	-.21	-3.00**	.32**	-.12	-3.11**	.32**
Item 9	-2.99**	.30**	-.32	.05	-3.29**	.34**	.08	-3.22**	.33**
Item 10	-2.36**	.19*	-1.70*	.19	-3.32**	.32**	-.30	-3.58**	.33**
Item 11	-3.17**	.30**	-.09	.05	-3.47**	.34**	.30	-3.19**	.33**
Item 12	-3.00**	.23**	-.62	.12	-3.69**	.31**	.28	-3.43**	.30**
Item 13	-4.15**	.34**	-.68	.07	-4.58**	.39**	-.11	-4.68**	.39**
Item 14	-3.69**	.25*	.03	-.03	-3.50**	.23**	-.22	-3.69**	.23**
Item 15	-4.25**	.35*	.20	-.02	-4.10**	.34**	.01	-4.09**	.34**
Item 16	-4.28**	.25*	.26	.00	-4.28**	.25**	.26	-4.05**	.25**
Item 17	-2.24**	.19*	.13	.05	-2.49**	.22**	.45	-2.08**	.21**
Item 18	-1.97**	.12	.48	.06	-2.28**	.16**	.86	-1.49**	.14**
Item 19	-2.17**	.19*	.39	.00	-2.15**	.18**	.37	-1.82**	.18**
Item 20	-2.05**	.12	.48	.02	-2.16**	.14**	.61	-1.60**	.12**

Item 21	-2.78**	.09	-.26	.05	-3.07**	.12*	.10	-2.97**	.12*
Item 22	-2.30**	.11	-.55	.04	-2.49**	.14**	-.30	-2.74**	.14**
Item 23	-2.72**	.15	-1.04	.03	-2.85**	.17**	-.84*	-3.52**	.19**
Item 24	-2.98**	.13	.06	.00	-3.00**	.13**	.07	-2.93**	.13**

---

Note: \*  $p < .05$ , \*\*  $p < .01$ , Model 1 = full model, Model 2 = 1<sup>st</sup> reduced model, and Model 3 = 2<sup>nd</sup> reduced model.

Table S5. Logistic Regression Coefficients for Maintaining Norms

	Model 1				Model 2			Model 3	
	$\tau_0$	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_0$	$\tau_1$	$\tau_2$	$\tau_0$	$\tau_1$
Item 1	-2.09**	.09	-.61	.11	-2.71**	.18**	.21	-2.56**	.18**
Item 2	-2.73**	.17*	-.19	.06	-3.13**	.22**	.32	-2.90**	.22**
Item 3	-3.42**	.18*	.22	.00	-3.44**	.19**	.24	-3.27**	.19**
Item 4	-2.01**	.07	-.40	.08	-2.49**	.14**	.21	-2.34**	.14**
Item 5	-2.18**	.11	.25	.02	-2.32**	.13**	.42	-2.02**	.13**
Item 6	-2.57**	.17*	.29	.01	-2.65**	.18**	.39	-2.36**	.18**
Item 7	-3.74**	.32**	1.21	-.16	-2.66**	.19**	-.11	-2.74**	.19**
Item 8	-2.09**	.14*	-.97	.06	-2.42**	.18**	-.51	-2.75**	.17**
Item 9	-3.25**	.32**	-.16	.02	-3.34**	.33**	-.04	-3.37**	.33**
Item 10	-3.70**	.41**	-.16	.04	-3.92**	.44**	.14	-3.82**	.44**
Item 11	-3.50**	.34**	.09	.02	-3.62**	.36**	.25	-3.43**	.36**
Item 12	-5.81**	.67**	2.33	-.25	-4.14**	.46**	.36	-3.85**	.46**
Item 13	-4.78**	.47**	1.95	-.19	-3.44**	.31**	.37	-3.16**	.31**
Item 14	-3.66**	.43**	.13	-.07	-3.27**	.38**	-.38	-3.53**	.37**
Item 15	-3.86**	.45**	.89	-.09	-3.32**	.37**	.23	-3.16**	.38**
Item 16	-3.39**	.42**	.02	.02	-3.48**	.43**	.14	-3.37**	.43**
Item 17	-2.58**	.19**	.31	-.07	-2.18**	.14**	-.20	-2.32**	.14**
Item 18	-2.07**	.16*	-.03	-.03	-1.89**	.14**	-.26	-2.07**	.14**
Item 19	-3.03**	.25**	-.07	.00	-3.04**	.25**	-.07	-3.08**	.25**
Item 20	-2.46**	.23**	-.52	.01	-2.50**	.23**	-.47	-2.81**	.23**

Item 21	-2.27**	.18**	.23	-.04	-2.04**	.15**	-.06	-2.08**	.15**
Item 22	-2.55**	.23**	.33	-.06	-2.22**	.18**	-.09	-2.28**	.18**
Item 23	-2.42**	.23**	.64	-.06	-2.05**	.18**	.17	-1.93**	.18**
Item 24	-2.10**	.22**	.20	-.06	-1.79**	.18**	-.21	-1.93**	.18**

---

Note: \*  $p < .05$ , \*\*  $p < .01$ , Model 1 = full model, Model 2 = 1<sup>st</sup> reduced model, and Model 3 = 2<sup>nd</sup> reduced model.

Table S6. Logistic Regression Coefficients for Post Conventional

	Model 1				Model 2			Model 3	
	$\tau_0$	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_0$	$\tau_1$	$\tau_2$	$\tau_0$	$\tau_1$
Item 1	-2.78**	.18**	.58	.00	-2.77**	.18**	.56*	-2.30**	.17**
Item 2	-2.29**	.14**	-.24	.05	-2.78**	.18**	.41	-2.45**	.18**
Item 3	-2.02**	.14**	-.22	.05	-2.47**	.18**	.38	-2.16**	.18**
Item 4	-3.05**	.19**	.61	.00	-3.05**	.19**	.61*	-2.54**	.19**
Item 5	-1.85**	.19**	.24	-.05	-1.41**	.15**	-.34	-1.66**	.15**
Item 6	-1.76**	.16**	-.02	-.01	-1.65**	.15**	-.18	-1.78**	.15**
Item 7	-2.39**	.18**	.63	-.02	-2.22**	.17**	.40	-1.89**	.17**
Item 8	-2.62**	.20**	.91	-.02	-2.43**	.19**	.66*	-1.87**	.18**
Item 9	-2.45**	.19**	-.32	.03	-2.75**	.21**	.08	-2.69**	.21**
Item 10	-2.60**	.19**	-1.01	.09	-3.40**	.26**	.11	-3.32**	.26**
Item 11	-2.42**	.20**	-1.33	.08	-3.14**	.26**	-.32	-3.38**	.26**
Item 12	-2.09**	.18**	-1.56	.09	-2.86**	.24**	-.48	-3.22**	.24**
Item 13	-2.82**	.28**	.31	-.05	-2.40**	.24**	-.23	-2.57**	.24**
Item 14	-3.25**	.29**	1.03	-.07	-2.60**	.23**	.20	-2.44**	.23**
Item 15	-3.44**	.30**	.02	-.02	-3.24**	.28**	-.26	-3.44**	.28**
Item 16	-2.38**	.23**	-1.08	.05	-2.83**	.27**	-.47	-3.17**	.26**
Item 17	-2.94**	.25**	.59	-.06	-2.42**	.21**	-.08	-2.48**	.21**
Item 18	-1.58*	.13**	-1.07	.05	-2.01**	.17**	-.49	-2.37**	.17**
Item 19	-2.95**	.25**	.18	-.04	-2.60**	.22**	-.28	-2.81**	.22**
Item 20	-2.33**	.19**	-.72	.04	-2.68**	.22**	-.24	-2.86**	.22**

Item 21	-1.11	.15**	-.27	.02	-1.24**	.16**	-.09	-1.31**	.16**
Item 22	-1.89**	.18**	-.05	.02	-2.04**	.20**	.16	-1.92**	.20**
Item 23	-2.07**	.19**	.88	-.07	-1.44**	.14**	.08	-1.38**	.14**
Item 24	-2.50**	.23**	.76	-.06	-1.94**	.18**	.04	-1.91**	.18**

---

Note: \*  $p < .05$ , \*\*  $p < .01$ , Model 1 = full model, Model 2 = 1<sup>st</sup> reduced model, and Model 3 = 2<sup>nd</sup> reduced model.

Table S6

Correlation between each individual philosophical rationale item and the overall scheme score

Items	PI	MN	PC
<b>Heinz and the drug</b>			
Whether a community's laws are going to be upheld. (MN)	-.21	.23	.03
Isn't it only natural for a loving husband to care so much for his wife that he'd steal? (PI)	.29	-.08	-.17
Is Heinz willing to risk getting shot as a burglar or going to jail for the chance that stealing the drug might help?	.24	-.01	-.16
Whether Heinz is stealing for himself or doing this solely to help someone else.	.19	-.17	-.02
Whether the pharmacist's rights to his invention have to be respected. (MN)	-.16	.18	.01
What values are going to be the basis for governing how people act towards each other. (PC)	-.38	-.07	.35
Whether the law in this case is getting in the way of the most basic claim of any member of society. (PC)	-.20	-.19	.31
Whether the pharmacist deserves to be robbed for being so greedy and cruel. (PI)	.30	-.08	-.20
Would stealing in such a case bring about more total good for the whole society or not.	-.10	-.16	.24
<b>Escaped prisoner</b>			
Hasn't Mr. Thompson been good enough for such a long time to prove he isn't a bad person?	.25	-.23	.00

(PI)			
Everytime someone escapes punishment for a crime, doesn't that just encourage more crime?			
(NM)	-0.03	.26	-.17
Has Mr. Thompson really paid his debt to society?	-0.06	-.05	.13
Would society be failing what Mr. Thompson should fairly expect?	-.05	.02	.00
How could anyone be so cruel and heartless as to send Mr. Thompson to prison? (PI)	.28	-.10	-.18
Would it be fair to all the prisoners who had to serve out their full sentences if Mr. Thompson was let off?	-.08	.21	-.06
Was Mrs. Jones a good friend of Mr. Thompson?	.24	-.09	-.11
Wouldn't it be a citizen's duty to report an escaped criminal, regardless of the circumstances?			
(NM)	-.12	.38	-.17
How would the will of the people and the public good best be served? (PC)	-.30	-.13	.38
Would going to prison do any good for Mr. Thompson or protect anybody? (PC)	-.11	-.26	.34
<b>Newspaper</b>			
Is the principal more responsible to students or parents?	.02	.10	-.07
Did the principal give his word that the newspaper could be published for a long time, or did he just promise to approve the newspaper one issue at a time? (NM)	.11	.16	-.21

Would the students start protesting even more if the principal stopped the newspaper? (PI)	.31	.05	-.26
When the welfare of the school is threatened, does the principal have the right to give orders to students? (NM)	-.10	.22	-.07
If the principal stopped the newspaper would he be preventing full discussion of important problems? (PC)	-.12	-.22	.29
Whether the principal's order would make Fred lose faith in the principal.	.15	-.05	-.12
Whether Fred was really loyal to his school and patriotic to his country. (PI)	.17	-.03	-.16
What effect would stopping the paper have on the student's education in critical thinking and judgment? (PC)	-.17	-.25	.35
Whether Fred was in any way violating the rights of others in publishing his own opinions.	-.13	-.10	.24
Whether the principal should be influenced by some angry parents when it is the principal that knows best what is going on in school.	.07	.09	-.10
Whether Fred was using the newspaper to stir up hatred and discontent.	.02	-.06	.09

*Note.* Letters in parentheses indicate to which schema the item is assigned. (PI): personal interest; (MN): maintaining norms; (PC): post-conventional schema. Selected items are color-highlighted (PI: yellow, MN: orange, PC: red).

## SAMPLE bDIT FORM

All questions are presented in the same way as they were presented during actual experiments for readers' information. Hence, all items and questions are presented in a random order as they were presented to participants.

It is strongly recommended to contact the Center for the Study of Ethical Development (<https://ethicaldevelopment.ua.edu/>) for the user's guide including the scoring guide.

**Frank and the Car** Frank Jones has been thinking about buying a car. He is married, has two small children and earns an average income. The car he buys will be his family's only car. It will be used mostly to get to work and drive around town, but sometimes for vacation trips also. In trying to decide what car to buy, Frank Jones realized that there were a lot of questions to consider. For instance, should he buy a larger used car or a smaller new car for about the same amount of money?

- Should buy a larger used car
- Can't decide
- Should buy a smaller new car

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Whether the car dealer was in the same block as where Frank lives.
- Whether the color was green, Frank's favorite color.
- Would a large, roomy car be better than a compact car.

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Whether the car dealer was in the same block as where Frank lives.
- Whether the color was green, Frank's favorite color.
- Whether the front connibilities were differential.

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Whether the car dealer was in the same block as where Frank lives.
- Whether the cubic inch displacement was at least 200.
- Would a large, roomy car be better than a compact car.

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Whether the car dealer was in the same block as where Frank lives.
- Whether the cubic inch displacement was at least 200.
- Whether the front connibilities were differential.

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Would a used car be more economical in the long run than a new car.
- Whether the color was green, Frank's favorite color.
- Would a large, roomy car be better than a compact car.

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Would a used car be more economical in the long run than a new car.
- Whether the color was green, Frank's favorite color.
- Whether the front connibilities were differential.

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Would a used car be more economical in the long run than a new car.
- Whether the cubic inch displacement was at least 200.
- Would a large, roomy car be better than a compact car.

(Current dilemma: Should he buy a larger used car or a smaller new car for about the same amount of money?) Which issue is more important in making a decision?

- Would a used car be more economical in the long run than a new car.
- Whether the cubic inch displacement was at least 200.
- Whether the front connibilities were differential.

**Heinz and the drug** In Europe, a woman was near death from a special kind of cancer.

There was one drug that doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost to make. He paid \$200 for the radium and charged \$2,000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about \$1,000, which is half of what it cost. He told the druggist that his wife was dying, and asked him to sell it cheaper or let him pay later. But the druggist said, "No, I discovered the drug and I'm going to make money from it." So Heinz got desperate and began to think about breaking into the man's store to steal the drug for his wife. Should Heinz steal the drug?

- Should steal
- Can't decide
- Should not steal

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- Isn't it only natural for a loving husband to care so much for his wife that he'd steal?
- What values are going to be the basis for governing how people act towards each other.
- Whether a community's laws are going to be upheld.

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- Whether a community's laws are going to be upheld.
- Whether the law in this case is getting in the way of the most basic claim of any member of society.
- Isn't it only natural for a loving husband to care so much for his wife that he'd steal?

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- What values are going to be the basis for governing how people act towards each other.
- Isn't it only natural for a loving husband to care so much for his wife that he'd steal?
- Whether the pharmacist's rights to his invention have to be respected.

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- Isn't it only natural for a loving husband to care so much for his wife that he'd steal?
- Whether the law in this case is getting in the way of the most basic claim of any member of society.
- Whether the pharmacist's rights to his invention have to be respected.

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- Whether a community's laws are going to be upheld.
- What values are going to be the basis for governing how people act towards each other.
- Whether the pharmacist deserves to be robbed for being so greedy and cruel.

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- Whether the pharmacist deserves to be robbed for being so greedy and cruel.
- Whether the law in this case is getting in the way of the most basic claim of any member of society.
- Whether a community's laws are going to be upheld.

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- Whether the pharmacist's rights to his invention have to be respected.
- What values are going to be the basis for governing how people act towards each other.
- Whether the pharmacist deserves to be robbed for being so greedy and cruel.

(Current dilemma: Should Heinz steal the drug?) Which issue is more important in making a decision?

- Whether the pharmacist's rights to his invention have to be respected.
- Whether the pharmacist deserves to be robbed for being so greedy and cruel.
- Whether the law in this case is getting in the way of the most basic claim of any member of society.

**Newspaper** Fred, a senior in high school, wanted to publish a mimeographed newspaper for students so that he could express many of his opinions. He wanted to speak out against the use of military in international disputes and to speak out against some of the school's rules, like the rule forbidding boys to wear long hair. When Fred started his newspaper, he asked his principal for permission. The principal said it would be all right if before every publication Fred would turn in all his articles for the principal's approval. Fred agreed and turned in several articles for approval. The principal approved all of them and Fred published two issues of the paper in the next two weeks. But the principal had not expected that Fred's newspaper would receive so much attention. Students were so excited by the paper that they began to organize protests against the hair regulation and other school rules. Angry parents objected to Fred's opinions. They phoned the principal telling him that the newspaper was unpatriotic and should not be published. As a result of the rising excitement, the principal ordered Fred to stop publishing. He gave as a reason that Fred's activities were disruptive to the operation of the school. Should the principal stop the newspaper?

- Should stop it
- Can't decide
- Should not stop it

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- Did the principal give his word that the newspaper could be published for a long time, or did he just promise to approve the newspaper one issue at a time?
- Would the students start protesting even more if the principal stopped the newspaper?
- If the principal stopped the newspaper would he be preventing full discussion of important problems?

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- Would the students start protesting even more if the principal stopped the newspaper?
- What effect would stopping the paper have on the student's education in critical thinking and judgment?
- Did the principal give his word that the newspaper could be published for a long time, or did he just promise to approve the newspaper one issue at a time?

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- Would the students start protesting even more if the principal stopped the newspaper?
- When the welfare of the school is threatened, does the principal have the right to give orders to students?
- If the principal stopped the newspaper would he be preventing full discussion of important problems?

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- When the welfare of the school is threatened, does the principal have the right to give orders to students?
- Would the students start protesting even more if the principal stopped the newspaper?
- What effect would stopping the paper have on the student's education in critical thinking and judgment?

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- If the principal stopped the newspaper would he be preventing full discussion of important problems?
- Whether Fred was really loyal to his school and patriotic to his country.
- Did the principal give his word that the newspaper could be published for a long time, or did he just promise to approve the newspaper one issue at a time?

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- Did the principal give his word that the newspaper could be published for a long time, or did he just promise to approve the newspaper one issue at a time?
- What effect would stopping the paper have on the student's education in critical thinking and judgment?
- Whether Fred was really loyal to his school and patriotic to his country.

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- Whether Fred was really loyal to his school and patriotic to his country.
- If the principal stopped the newspaper would he be preventing full discussion of important problems?
- When the welfare of the school is threatened, does the principal have the right to give orders to students?

(Current dilemma: Should the principal stop the newspaper?) Which issue is more important in making a decision?

- When the welfare of the school is threatened, does the principal have the right to give orders to students?
- Whether Fred was really loyal to his school and patriotic to his country.
- What effect would stopping the paper have on the student's education in critical thinking and judgment?

**Escaped prisoner** A man had been sentenced to prison for 10 years. After one year, however, he escaped from prison, moved to a new area of the country, and took on the name of Thompson. For eight years he worked hard, and gradually he saved enough money to buy his own business. He was fair to his customers, gave his employees top wages, and gave most of his own profits to charity. Then one day, Mrs. Jones, an old neighbor, recognized him as the man who had escaped from prison eight years before, and whom the police had been looking for. **Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?**

- Should report him
- Can't decide
- Should not report him

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Everytime someone escapes punishment for a crime, doesn't that just encourage more crime?
- Hasn't Mr. Thompson been good enough for such a long time to prove he isn't a bad person?
- How would the will of the people and the public good best be served?

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Would going to prison do any good for Mr. Thompson or protect anybody?
- Hasn't Mr. Thompson been good enough for such a long time to prove he isn't a bad person?
- Everytime someone escapes punishment for a crime, doesn't that just encourage more crime?

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Hasn't Mr. Thompson been good enough for such a long time to prove he isn't a bad person?
- How would the will of the people and the public good best be served?
- Wouldn't it be a citizen's duty to report an escaped criminal, regardless of the circumstances?

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Hasn't Mr. Thompson been good enough for such a long time to prove he isn't a bad person?
- Wouldn't it be a citizen's duty to report an escaped criminal, regardless of the circumstances?
- Would going to prison do any good for Mr. Thompson or protect anybody?

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Everytime someone escapes punishment for a crime, doesn't that just encourage more crime?
- How could anyone be so cruel and heartless as to send Mr. Thompson to prison?
- How would the will of the people and the public good best be served?

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Everytime someone escapes punishment for a crime, doesn't that just encourage more crime?
- Would going to prison do any good for Mr. Thompson or protect anybody?
- How could anyone be so cruel and heartless as to send Mr. Thompson to prison?

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Wouldn't it be a citizen's duty to report an escaped criminal, regardless of the circumstances?
- How would the will of the people and the public good best be served?
- How could anyone be so cruel and heartless as to send Mr. Thompson to prison?

(Current dilemma: Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison?) Which issue is more important in making a decision?

- Would going to prison do any good for Mr. Thompson or protect anybody?
- Wouldn't it be a citizen's duty to report an escaped criminal, regardless of the circumstances?
- How could anyone be so cruel and heartless as to send Mr. Thompson to prison?