Validation of the Moral Reasoning Questionnaire against Rasch Measurement Theory

Lyndon Lim1,2 and Elaine Chapman3

Abstract
To support teachers in facilitating students’ moral reasoning development as specified within the Singapore Ministry of Education Character and Citizenship Education curriculum, the Moral Reasoning Questionnaire (MRQ) was developed and underwent preliminary validation. Based upon expert reviews, cognitive interviews and a classical test theory-based factor analytic approach, the development and preliminary validation found evidence (i.e., content appropriateness, response processes and internal structure) to support the validity and reliability of the MRQ. This study aims to extend the validation by examining the purported MRQ items and scale at a deeper level on the Rasch Measurement Theory, given that it is the only model that presents appropriate properties of interval measurement on a log-linear scale. The Rasch analysis found anomalies including differential item functioning and disordered thresholds in the initial set of items. Upon remediation and a second Rasch analysis, the MRQ responses were consistent with that expressed by the Rasch model (i.e., an item with an endorsability higher than what a respondent would tend to endorse would have a lower probability of being endorsed than an item exhibiting an endorsability below what that respondent would tend to endorse) and hence, there was sufficient evidence to support measurement invariance, and that MRQ scores could be concluded to characterise persons invariantly across a continuum.

Keywords
Moral reasoning, assessment, secondary education, validity, reliability, Rasch measurement

Introduction
The Character and Citizenship (CCE) curriculum developed by the Singapore Ministry of Education (MOE) aims to develop students into confident people who are discerning in judgment and possess a strong sense of right and wrong (MOE, 2012, 2016, 2020). This supports the MOE desired outcomes of education, which include qualities embodied within a person who is: (1) confident, (2) a self-directed learner, (3) an active contributor and (4) a concerned citizen. Learning outcome eight (LO8: Reflect on and respond to community, national and global issues, as an informed and responsible citizen) of the CCE curriculum appropriates a charge in supporting the desired outcomes of education. LO8 is detailed by key stage outcomes, in part addressing the values of respect and responsibility, and the social awareness domain. Inextricably linked to moral reasoning, the intended key stage outcome is for students to: (1) be able to distinguish right from wrong at the primary level, (2) have moral integrity at the secondary level and (3) have the moral courage to stand up for what is right at the pre-tertiary level (MOE, 2012, 2016, 2020).

In part to achieve LO8, the CCE curriculum intends for students to progress through various levels of moral reasoning based on Kohlberg’s stage-based theory of moral development (Kohlberg, 1984) (Figure 1). To do this, curriculum documents have suggested several approaches that teachers can apply (e.g., discussing moral dilemmas on a clarify–sensitise–influence approach and modelling how decisions could be made in the context of these dilemmas). While personable and desirable, the approaches suggested are considerably resource intensive, given that teachers

1Teaching and Learning Centre, Singapore University of Social Sciences, Singapore
2Assessment and Research Group, Singapore Examinations and Assessment Board, Singapore
3Graduate School of Education, The University of Western Australia, Perth, Australia

Corresponding author:
Lyndon Lim, Teaching and Learning Centre, Singapore University of Social Sciences, Singapore.
Email: lyndonlimjk@suss.edu.sg
would have to record their discussions with each student as a form of tracking, without which they might not be conscious of progress made by each student. In light of this, the Moral Reasoning Questionnaire (MRQ) was developed upon an operational definition of moral reasoning proffered by Lim & Chapman (2021a) for use in Singapore schools on a large-scale basis for students aged between 12 and 18 (between grade 7 and 12), after an extensive review of established instruments found concerns with content appropriateness and group administrability (Lim & Chapman, 2021b).

Based on critical stages recommended by the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association & National Council on Measurement in Education [AERA, APA & NCME], 2014), the MRQ was preliminarily validated in part on a classic test theory (CTT) factor analytic approach to establish its factorial structure, and the analyses found both quantitative and qualitative support for validity evidence and the reliability of the MRQ (Lim & Chapman, 2021c); quantitative support was established via the factor analytic approach (i.e., exploratory factor analysis followed by confirmatory factor analysis) and parallel analysis, while qualitative support was established via evidence from: (1) content appropriateness where an expert panel critique the initial item pool, and (2) response processes, where the items were validated through engaging five students, within the age range by which the MRQ was intended, in cognitive interviews (Willis, 2017).

**Rasch Measurement Theory**

Despite its widespread application in the field of educational measurement since the 1920s (Kohli et al., 2015), the CTT-based factor analytic approach is subject to several limitations, owing in part to its circular dependency in terms of conceptualisation. In CTT, person statistics (i.e., observed scores) are inherently item-sample dependent and often assumed to be normally distributed, while item statistics (i.e., difficulty and discrimination) are person-sample dependent (Boone, 2016; Ewing et al., 2005; Kohli et al., 2015). This restricts the applicability of CTT in various important measurement situations (e.g., situations in which different tests must be equated). Further, in situations that involve rating scale data (e.g., ordinal scales such as the Likert scale), approaches grounded in CTT do not provide a basis for exploring the additivity of scores, a critical attribute of any valid measure (Wu & Leung, 2017). These call for an additional validation of the MRQ, and Rasch Measurement Theory (RMT), posited as an elaboration of CTT (Andrich & Marais, 2019), fits the purpose of this study. The view that Rasch analysis is designed for dichotomous or polytomous response data, makes no distributional assumptions and enables rating scales to be modified through identifying and dealing with misfitting items so that an instrument fittingly measures a latent trait supports RMT as a suitable method for extending the validation of the MRQ (Hendriks et al., 2012; Pallant & Tennant, 2007).

Developed by Georg Rasch in 1957, an important feature of RMT is a table of expected response probabilities reflecting Rasch’s view that a person with a greater proficiency than another should have a higher probability of solving or endorsing an item; conversely, an item that is more difficult or more difficult to endorse than another implies that for any person, the probability of endorsing or solving the other item is higher (Andrich, 1997; Andrich & Marais, 2019; Bond & Fox, 2015; Rasch, 1960). By first locating and ordering persons and item difficulty on a log-linear scale reflecting degrees of the latent trait (e.g., easy to difficult; most to least endorsable), incorporating the ordering features of Guttman scaling, and the focus on probabilistic distributions of examinees’ performance at the item level, rather than on test-level information, RMT offers an alternative basis for constructing measurements within and beyond education (Pallant & Tennant, 2007; Rasch, 1960).

Following the ordering of person and items on the log-linear scale, psychometric properties of an instrument can be determined; these include the dimensionality of the instrument, fit to the Rasch model (i.e., person and item fit), threshold ordering, differential item functioning (i.e., item bias), local independence and person-separation index (PSI) (i.e., internal reliability). An important advantage of the Rasch approach derives from the fact that the

| Level                  | Stage                                      | MOE Stipulated grade attainment |
|------------------------|--------------------------------------------|---------------------------------|
| Pre-Conventional       | 1. Obedience and punishment orientation    | K2-12                           |
|                        | 2. Self-interest orientation               | K2-12                           |
| Conventional           | 3. Interpersonal accord and conformity     | K4-12                           |
|                        | 4. Authority and social-order maintaining  | K7-12                           |
| Post-Conventional      | 5. Social contract orientation             | K10-12                          |
|                        | 6. Universal ethical principles            | K10-12                          |

**Figure 1.** Kohlberg’s stage-based theory of moral development.
item parameters obtained do not depend on the characteristics of the persons taking the test, and that the person parameters do not depend on the specific items chosen for a given test (Andrich & Marais, 2019; Bond & Fox, 2015). The parameters produced through Rasch analysis are thus independent of specific sample characteristics; the interpretation of person measures are taken with reference to the items defining the purported latent trait, as opposed to CTT, where person measures are interpreted with reference to the sample mean (Ewing et al., 2005). In view of this, the Rasch approach, taken as an elaboration of CTT, addresses issues with the CTT-based factor analytic approach.

Further, Rasch measurement models are tenable as a basis for examining scores obtained through rating scales, because these models provide a means by which the hierarchical structure, unidimensionality and additivity of the scores can be evaluated.

It is noteworthy that while there have been various terms involving Rasch analysis and its application to dichotomous or polytomous scoring structures, Andrich et al. (2018) suggested that there remains only one model (i.e., the unidimensional Rasch model for ordered categories) involved with different types of items; terms such as Dichotomous Model, Partial Credit Model, Rating Model are unnecessary and have misled some to assume that there are different Rasch models.

**Moral Reasoning Questionnaire**

The 26-item MRQ was developed by the recommendations of Lim & Chapman (2021b) in their review of existing moral reasoning measures whilst considering the intended audience (see Appendix A). Based upon Kohlberg’s stage-based theory of moral development (Kohlberg, 1984), the MRQ was initially developed with 30 items, of which four were discarded during the preliminary validation. The MRQ is intended to be delivered online and the items within are of a two-tier response format. In tier one, respondents would select one of two ‘action’ options after reading a moral dilemma vignette. Tier two, an ordering response format, would then be presented based on respondents’ selection in tier one; tier two requires respondents to rank the options in order of importance to themselves, and each of these options correspond to a level based on Kohlberg’s stage-based theory of moral development. Figure 2 presents an example of an item within the MRQ.

Responses to the MRQ are scored based on the scoring matrix presented in Table 1. This scoring matrix had been applied during the preliminary validation and there was no evidence to suggest, based on the CTT factor analytic approach, that it was inappropriate (Lim & Chapman, 2021c) (see Appendix B for results).


Table 1. Scoring matrix of two-tier items.

| Rank order (levels of moral judgment) | Score |
|---------------------------------------|-------|
| (1) Post-conventional                  | 5     |
| (2) Conventional                      |       |
| (3) Pre-conventional                  |       |
| (1) Conventional                       | 4     |
| (2) Post-conventional                  |       |
| (3) Pre-conventional                  |       |
| (1) Post-conventional                  | 3     |
| (2) Pre-conventional                   |       |
| (3) Conventional                      |       |
| (1) Conventional                       | 2     |
| (2) Pre-conventional                   |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 1     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 0     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |

Table 2. Scoring matrix of two-tier items with revised categories.

| Rank order (levels of moral judgment) | Score |
|---------------------------------------|-------|
| (1) Post-conventional                  | 2     |
| (2) Conventional                      |       |
| (3) Pre-conventional                  |       |
| (1) Conventional                       | 2     |
| (2) Post-conventional                  |       |
| (3) Pre-conventional                  |       |
| (1) Post-conventional                  | 1     |
| (2) Pre-conventional                   |       |
| (3) Conventional                      |       |
| (1) Conventional                       | 1     |
| (2) Pre-conventional                   |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 1     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 0     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |

Participants

Data for this study were drawn from the responses of participants who took part in the preliminary validation that was approved by the authors’ respective institutional review board (IRB). As required by the IRBs, each participant received a consent form and a participant information sheet that specified that: (1) participant involvement in the research was voluntary (2) participants were free to withdraw at any stage without prejudice in any way, with no reason required for withdrawal and (3) all data would be anonymised and each participant would not be identifiable. As all the participants were considered minors, parental consent was sought before each of the participants took part in the study. The participants were from three secondary schools (grades 8 to 12, aged between 12 and 18) in Singapore that agreed to support the study following the access permission that was granted by the MOE in 2015. Of the 670 participants whose parents/guardians agreed to let them participate in this study, 497 were female. The age range of the participants was 12 to 18 years (M = 14.24, SD = 1.30 years). As to schools, 36.7% (n = 246) of the participants were from school M (a mixed-sex government secondary and secular school), 44.2% (n = 296) from school P (a single-sex autonomous government-aided secondary and mission school) and 19.1% (n = 128) from school X (a mixed-sex autonomous government secondary and secular school). All school types follow the national syllabus with the following distinctions: (1) instituted by various community and religious organisations, government-aided schools include mission schools and serve the educational needs of specific communities; (2) autonomous schools offer a wider range of programmes for students.

The participants from the three participating schools were considered diverse as they represented different educational levels and streams. At the point of data collection, 17.6% (n = 118) of the participants were from secondary one, 27.8% (n = 186) were from secondary two; 28.4% (n = 190) were from secondary three, and 26.3% (n = 176) were from secondary four. As to streams, 79.4% (n = 532) were from the express stream, 16.3% (n = 109) from the normal-academic stream and 4.3% (n = 29) from the normal-technical stream; the majority of students in Singapore secondary schools are from the express stream. There were no missing values or incomplete responses when the data were processed after collection.

Rasch analysis

The Rasch analysis was conducted using the Rasch Unidimensional Measurement Model (RUMM2030) software version 5.4 (RUMM Laboratory Pty Ltd, Perth, Australia). The analysis was performed primarily to assess the data fit to an unrestricted Rasch model, without assuming a uniform distance between response thresholds. The MRQ and its items were evaluated, using parametric statistical tests, for: (1) threshold ordering and reliability, (2) the overall model fit, (3) individual item and person fit (4), the item characteristic curves, (5) local dependency and differential item functioning and (6) dimensionality.

Threshold ordering and reliability

Initial results obtained from the Rasch analysis indicated the presence of disordered thresholds across a number of items in the MRQ, and a significant chi-square statistic, \( \chi^2 \) (270, N = 669) = 517.27, \( p < .001 \) (see Appendix C for illustration of disordered thresholds). The reliability of the MRQ, however, appeared to be good, with a PSI of 0.88. This PSI value indicated: (1) a good spread of item estimates given the presence of multiple thresholds (six response categories) for each item and (2) a high estimated true variance in respondents’ moral reasoning levels (i.e., only 12% of the variance attributable to error variance).

The presence of disordered thresholds suggested that respondents might not have been able to distinguish between the six response categories presented within the MRQ based on the Rasch model. Given this result, a

---

**Table 1. Scoring matrix of two-tier items.**

| Rank order (levels of moral judgment) | Score |
|---------------------------------------|-------|
| (1) Post-conventional                  | 5     |
| (2) Conventional                      |       |
| (3) Pre-conventional                  |       |
| (1) Conventional                       | 4     |
| (2) Post-conventional                  |       |
| (3) Pre-conventional                  |       |
| (1) Post-conventional                  | 3     |
| (2) Pre-conventional                   |       |
| (3) Conventional                      |       |
| (1) Conventional                       | 2     |
| (2) Pre-conventional                   |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 1     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 0     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |

**Table 2. Scoring matrix of two-tier items with revised categories.**

| Rank order (levels of moral judgment) | Score |
|---------------------------------------|-------|
| (1) Post-conventional                  | 2     |
| (2) Conventional                      |       |
| (3) Pre-conventional                  |       |
| (1) Conventional                       | 2     |
| (2) Post-conventional                  |       |
| (3) Pre-conventional                  |       |
| (1) Post-conventional                  | 1     |
| (2) Pre-conventional                   |       |
| (3) Conventional                      |       |
| (1) Conventional                       | 1     |
| (2) Pre-conventional                   |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 1     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |
| (1) Pre-conventional                   | 0     |
| (2) Conventional                      |       |
| (3) Post-conventional                  |       |
remediation was undertaken to have the categories revised into a smaller number, as suggested by Andrich and Marais (2019), and the data was re-scored based on this revised scoring matrix and subjected to a second round of Rasch analysis (results reported from here on). The revised categories, presented in Table 2, were premised on the following assumptions: (1) that a respondent would score 2 if she or he identified the pre-conventional level as the lowest level of moral judgment; (2) a respondent would score 0 if she or he identified the levels of moral judgment opposite to that of Kohlberg’s stage-based theory of moral development; and (3) a respondent would score 1 for all other rank order permutations.

In the Rasch analysis performed on the MRQ using the revised categories, items A5, A11, A13 and A20 were removed as disordered thresholds (i.e., persons at a higher moral reasoning stage demonstrated a probability of endorsing a more endorsable item lower compared to that of persons at a lower moral reasoning stage and vice versa) remained evident. All other items did not have disordered thresholds using the revised categories. This is in agreement with the preliminary validation that also identified these four items as items that likely caused a model misfit based on the factor analytic approach. Figure 3 presents the threshold map for the remaining 26 items and Table 3 shows the uncentralised item thresholds that indicate all category responses have been used as expected consistently.

Despite the removal of items A5, A11, A13 and A20, the Rasch analysis continued to find adequate reliability estimates (i.e., a PSI of .84 and Cronbach’s $\alpha$ of .89) based on the minimum PSI recommended by Tennant and Conaghan (2007) as .7. These outcomes suggest that the MRQ scale should be able to differentiate between at least two groups of respondents, given that the PSI reflects whether the scale is adequately robust and provides for differentiation between subgroups or individuals within the data set (Andrich, 1982; Streiner et al., 2015).

**Overall model fit**

In terms of the overall fit of the model, the $\chi^2$ test for the overall fit to the Rasch model remained significant at $\chi^2$.
(234, n = 647) = 333.74, p < .001 based on the Rasch analysis with revised categories. While a p-value less than .01 indicates that the data would likely not fit the Rasch model, it should be noted that the \( \chi^2 \) statistic increases with sample size, and that other sources of evidence such as the item characteristic curve should also be reviewed (Andrich & Marais, 2019). In view of the sensitivity of the \( \chi^2 \) statistic to the sample size, an adjusted sample was used, and the overall fit test indicated \( \chi^2 (234, n = 527) = 271.84, p = .05 \); the adjusted sample is an algebraic adjustment option in RUMM2030 involving the same data set (Andrich & Marais, 2019). This non-significant item–trait interaction \( \chi^2 \) statistic suggested: (1) a good overall model fit; and (2) that the items collectively measure a common latent trait. The good overall model fit was also supported by the item fit residual (\( M = -.14, SD = 1.53 \)) and person fit residual (\( M = -.16, SD = .95 \)), which had means close to zero and standard deviations close to one, as suggested by Andrich and Marais (2019), and Tennant and Conaghan (2007).

### Individual item and person fit

Examining the individual item and person fit outputs, fit residuals should lie within the range of ± 2.5 for an item to be considered fitting to the Rasch model (Tennant & Conaghan, 2007; Tennant & Pallant, 2006) though Andrich and Marais (2019) stated that ‘there are no absolute criteria for interpreting fit statistics’ (p. 196). In this case, all items (with Bonferroni adjustment) except for A9 (2.71) and 26 (−2.65) were within the threshold and \( p \)-values were >.01 (Table 4). Based on this evidence, the vast majority of items within the MRQ meet the criteria for adequate fit.

### Item characteristic curves

Item characteristic curves (ICC) are reviewed as part of any Rasch analysis (e.g., when the \( p \)-value of the \( \chi^2 \) statistic is less than .01) (Andrich & Marais, 2019). In this Rasch analysis, reviewing all the 26 ICCs did not reveal major concerns, with the exception of items A9 and A26 that showed modest non-systematic misfits. The ICC of item A9 (Figure 4) suggests that respondents at the post-conventional level of moral reasoning found it difficult to endorse this item, and this likely influenced the item fit. The ICC for item A26 (Figure 5) suggests that the item misfit observed could have been influenced by over-discrimination where respondents at the post-conventional level found the item slightly too endorsable while those at

### Table 3. Uncentralised item thresholds.

| Item label in RUMM2030 | Item number in MRQ | Endorsability | Mean   | Threshold 1 | Threshold 2 |
|------------------------|--------------------|---------------|--------|-------------|-------------|
| 0001                   | A1                 | .01           | .01    | −.60        | .63         |
| 0002                   | A2                 | .04           | .04    | −.37        | .45         |
| 0003                   | A3                 | −.01          | −.01   | −.46        | .44         |
| 0004                   | A4                 | −.08          | −.08   | −.86        | .69         |
| 0005                   | A6                 | .17           | .17    | −.59        | .94         |
| 0006                   | A7                 | 1.00          | 1.00   | −.41        | 2.41        |
| 0007                   | A8                 | .33           | .33    | −1.08       | 1.74        |
| 0008                   | A9                 | −.45          | −.45   | −2.86       | 1.95        |
| 0009                   | A10                | .11           | .11    | −.70        | .92         |
| 0010                   | A12                | −.06          | −.06   | −1.01       | .88         |
| 0011                   | A14                | .09           | .09    | −.54        | .72         |
| 0012                   | A15                | −.78          | −.78   | −1.21       | −.35        |
| 0013                   | A16                | −.51          | −.51   | −1.47       | .45         |
| 0014                   | A17                | −.24          | −.24   | −1.21       | .73         |
| 0015                   | A18                | −.00          | −.00   | −1.20       | 1.20        |
| 0016                   | A19                | −.66          | −.66   | −1.09       | −.22        |
| 0017                   | A21                | −.33          | −.33   | −.47        | 1.12        |
| 0018                   | A22                | −.46          | −.46   | −1.53       | .62         |
| 0019                   | A23                | .93           | .93    | −.05        | 1.91        |
| 0020                   | A24                | −.17          | −.17   | −.68        | .33         |
| 0021                   | A25                | .19           | .19    | −.38        | .77         |
| 0022                   | A26                | .18           | .18    | −.52        | .88         |
| 0023                   | A27                | .59           | .59    | −.34        | 1.51        |
| 0024                   | A28                | .16           | .16    | −.77        | 1.08        |
| 0025                   | A29                | −.48          | −.48   | −1.05       | .09         |
| 0026                   | A30                | −.21          | −.21   | −.85        | .43         |
the pre-conventional level found it slightly more difficult to endorse.

Though the fit residuals of items A9 and A26 were slightly beyond the ±2.5 range, their $\chi^2$ probability with Bonferroni adjustment was not less than .01. In view of this and the modest item misfit, it was concluded that all 26 items measure a common underlying construct (i.e., moral reasoning). With regard to person fit, only 10 respondents excluding extreme cases had a fit residual, out of the ±2.5 range, between −3.69 and 3.11. This could indicate anomalies in the score patterns of these respondents, which may have reflected fatigue. As there was no data entry error and given the good overall model fit, these respondents were not removed.

**Local dependency and differential item functioning**

Violations of local dependency were investigated based on the residual correlation matrix (Pallant & Tennant, 2007). From the Rasch analysis, the maximum inter-item residual correlation ($r = .18$) was between items A2 and A3, less than the .2 threshold proffered by Andrich et al. (2018). This indicates minimal local dependency and how a respondent performs for an item would have little or no bearing on other items.

Differential item functioning (DIF) occurs ‘when items do not function in the same way for different groups of people, who otherwise have the same value on the trait’ (Andrich & Marais, 2019, p. 199). The DIF analyses performed in this study afford the determination of evidence of any item bias within the MRQ, although these analyses were somewhat limited by the data that could be obtained in the study. As the MRQ is designed to assess moral reasoning across students at secondary one to four across gender, schools and streams, DIF was assessed for school, stream, gender and level. A Bonferroni correction (i.e., dividing the probability value of significance by the number of tests of fit) was used in these analyses. Table 5 presents the $p$-values of the ANOVA (main effect) used to determine the existence of DIF by level, gender, school and educational stream. From the results, it was concluded that no DIF was evident for school, educational stream, gender and level, with the exception of item A10 that exhibited item bias (i.e., uniform DIF) across different levels, $F(3, 607) = 12.70, p < .001$; there was no evidence of non-

---

**Table 4.** Fit statistics for MRQ with revised categories.

| Item label in RUMM2030 | Item number in MRQ | Fit residual | $\chi^2$ probability |
|------------------------|-------------------|-------------|----------------------|
| I0001                  | A1                | 1.11        | 0.22                 |
| I0002                  | A2                | 0.73        | 0.60                 |
| I0003                  | A3                | −0.47       | 0.78                 |
| I0004                  | A4                | −0.19       | 0.84                 |
| I0005                  | A6                | 1.68        | 0.10                 |
| I0006                  | A7                | 0.32        | 0.14                 |
| I0007                  | A8                | 1.27        | 0.31                 |
| I0008                  | A9                | 2.71        | 0.01                 |
| I0009                  | A10               | 0.12        | 0.11                 |
| I0010                  | A12               | 1.7         | 0.01                 |
| I0011                  | A14               | −2.35       | 0.06                 |
| I0012                  | A15               | −0.99       | 0.19                 |
| I0013                  | A16               | −2.02       | 0.03                 |
| I0014                  | A17               | 0.85        | 0.88                 |
| I0015                  | A18               | −2.25       | 0.02                 |
| I0016                  | A19               | −1.41       | 0.36                 |
| I0017                  | A21               | −2.23       | 0.07                 |
| I0018                  | A22               | 0.63        | 0.77                 |
| I0019                  | A23               | −0.02       | 0.52                 |
| I0020                  | A24               | −1.39       | 0.29                 |
| I0021                  | A25               | 1.16        | 0.20                 |
| I0022                  | A26               | −2.65       | 0.01                 |
| I0023                  | A27               | −2.08       | 0.06                 |
| I0024                  | A28               | 1.32        | 0.18                 |
| I0025                  | A29               | 1.24        | 0.29                 |
| I0026                  | A30               | −0.46       | 0.60                 |

**Figure 4.** ICC of item A9.
uniform DIF for item A10 based on the ANOVA (interaction effects), $F(27, 607) = 0.74, p = .83$.

To appreciate the extent of the item bias, the ICC with level plots for item A10 of the MRQ (Figure 6) was generated and reviewed. Visually, there is item bias between the secondary one (S1) and secondary four (S4) levels at 0 to 0.5 logits and above 2.5 logits. The uniform DIF could be significant due to the large sample size.
given that a small difference in how an item functions across subgroups would result in a significant statistical test (Dogan et al., 2018; Teresi et al., 2021). Further, upon reviewing the wording and structure of the vignette of item A10 (Figure 2), the DIF was deemed to be benign rather than adverse (Douglas et al., 1996). Despite the lack of evidence of non-uniform DIF that would warrant a removal of the offending item (Pallant & Tennant, 2007), item A10 was removed to ascertain if model fit would be impacted. The Rasch analysis found that model fit was not impacted even with the removal of item A10. In view of these, and the evidence supporting the adequacy

Figure 6. ICC with level plots for item A10.

Figure 7. PCA t-test of ± loaded items on first principal component.
of how the data fits the Rasch model, item A10 was eventually retained.

**Dimensionality**

Further to the non-significant item-trait interaction $\chi^2$ statistic, evidence from the principal components analysis (PCA) supported the unidimensionality of the MRQ. All the items loaded on one principal component, which also supported the assumption of item local independence. The PCA revealed two patterns of items that loaded (positively and negatively) onto the first principal component. An independent $t$-test was then done on these two sets of items to make separate person estimates. For unidimensionality, no more than 5% of the $t$-test results should be significant ($p < .05$) (Andrich & Marais, 2019; Smith, 2002). As only 4.33% of respondents showed a significant difference between the person locations based on the two sets of items, the MRQ was concluded to be unidimensional (Figure 7).

The findings presented so far point to a good overall model fit based on the Rasch model, and support the unidimensionality of the MRQ. The person-item threshold distribution that places student (person) and item location estimates on the same logit scale (Figure 8) shows that the items and thresholds spanned almost the range of person scores except for some who scored very high on the MRQ. This could be explained by the MOE’s
expectation that more secondary school students fall within
the conventional to post-conventional levels of moral
development.

Further analyses suggested that inferences drawn from
the measure would not be confounded by students’ demo-
graphic attributes (i.e., gender, school or educational
stream). Females did have slightly higher moral reasoning
scores, but did not differ significantly from males \( F(1, 668) = \) 2.91, \( p = .09 \) (Figure 9). Moral reasoning scores
were also not influenced by where the school participants
were from \( F(2, 667) = .09, p = .92 \) (Figure 10) or by the
educational stream \( F(1, 668) = .31, p = .58 \) (Figure 11)
that the participants were in. Nonetheless, for the levels of
study, there was a statistically significant difference in
moral reasoning scores \( F(3, 666) = 2.87, p = .04 \)
(Figure 12). While this difference is to be expected, as
moral reasoning should be developmental, there was no
definitive trend in the analysis that students at higher
levels of study scored higher on the MRQ.

**Discussion**

Based on the analyses presented in six areas (i.e., threshold
ordering and reliability, overall model fit, individual item
and person fit, item characteristic curves, local dependency
and differential item functioning, and dimensionality), there
was adequate evidence to affirm the unidimensionality and hence the intended purpose of the MRQ. The MRQ items also proved to be functioning as anticipated, based on how the data fit the Rasch model.

The Rasch analysis presented item A7 as the least endorsable item on the log-linear scale (i.e., item location = .998), with adjacent category thresholds −.4 and 2.4 logits (Figure 13). Based on the polytomous Rasch model as expressed by equation (1) (Andrich & Marais, 2019), the probabilities of a student with a proficiency at the zero logit scoring 0, 1 and 2 are .38, .57 and .05, respectively; the probabilities of an average student with a mean of the order of 1.624 (Figure 8) scoring 0, 1 and 2 are .23, .65 and .11, respectively. In the same vein, the most endorsable item on the log-linear scale, item A15 (i.e., item location = −.779) had adjacent category thresholds of −1.2 and −.3 (Figure 14). Based on equation (1), the probability of a student with proficiency at the zero logit scoring 0, 1 and 2 are .11, .37 and .52, respectively, and that of a student with a mean proficiency of the order of 1.624 are .00, .00 and .94, respectively.

\[
P(x_{ni} = x) = \frac{e^{-\tau_{1i} - \tau_{2i} - \ldots - \tau_{xi} + x(\theta_n - \delta_i)}}{\sum_{x'=0}^{m} e^{-\tau_{1i} - \tau_{2i} - \ldots - \tau_{x'i} + x'(\theta_n - \delta_i)}}
\]  

(1)
where $P(x_{ni} = x)$ is the probability that person $n$ selects in category 'x' respectively on item $i$, $\beta$ is the person location parameter, $\tau$ are the response probability thresholds, and $\delta$ is the mean of these thresholds.

Considering these and that the MOE expects most secondary school students to fall within the conventional to post-conventional levels of Kohlberg’s stage-based model, it appears that the targeting of the MRQ (Figure 8) could be further refined and subsequent versions of the MRQ could include more ‘difficult to endorse’ items to assess the conventional to post-conventional levels of moral reasoning.

More could also be done to establish the MRQ as a measure of moral reasoning independent of cognition and intelligence, given that various previous studies have reported correlations in the .20 to .50 range between moral judgments and measures of intelligence, aptitude and achievement (Rest, 1979; Thoma & Dong, 2014). The data used in this study were from students from mainstream secondary schools and hence the MRQ appears to be fit for purpose in Singapore mainstream secondary school students. To ascertain that there is no DIF across a wider range of cognitive levels, and hence provide assurance that students of higher scholastic aptitude do not necessarily exhibit higher levels of moral reasoning, the MRQ could be administered to students who have comparable literacy levels but are not from mainstream secondary schools.

Though the Rasch analysis suggested that inferences drawn from the measure would not be confounded by students’ demographic attributes (i.e., gender, school or educational stream), the disproportionate sample by gender and educational stream, owing to availability sampling, could optically suggest otherwise. Hence, a more representative sample could be invited to participate in subsequent studies. With more data, measurement invariance by educational stream and gender could be affirmed.

Disordered thresholds that were identified through this study called for the MRQ scoring matrix to be revised. Hence, the revised scoring matrix should be used moving forward so that a log-linear person-measure of the MRQ can be established for the meaningful comparison of respondents’ moral reasoning.

Based on the triangulation of evidence across all of these analyses, the present study together with the preliminary validation present the MRQ as an instrument that can be used in Singapore secondary schools to monitor students’ development in the area of moral reasoning. As an accessible instrument with sound psychometric properties founded upon both CTT and RMT, the MRQ would be suitable for use on a large-scale basis. A further advantage of this instrument is that minimal training is required for teachers to administer and score the test. This adds further support to the notion that the MRQ can provide a practical means by which students’ development in moral reasoning can be monitored, hence addressing a major gap identified in this context (Lim & Chapman, 2021b).

**Conclusion**

This paper detailed how the RMT approach was used to validate the moral reasoning scale based on the MRQ, how the analyses were interpreted and how identified issues were resolved. The RMT approach undertaken in this study served as an elaboration of the CTT-based factor analytic approach used within the preliminary validation of the MRQ.
The Rasch analysis found evidence to support, amongst the reported psychometric properties, the unidimensionality and intended purpose of the MRQ, though issues related to disordered thresholds were identified. This led to a revised scoring matrix upon which further analyses found that invariant comparisons of persons and items could be drawn. Hence, it appears that the MRQ presents a viable scale free of DIF for measuring moral reasoning in students within the context of Singapore secondary education.

By its nature, the ‘validation process never ends, as there is always additional information that can be gathered to more fully understand a test and the inferences that can be drawn from it’ (AERA, APA & NCME, 2014, p. 21). While this study is an extension of the preliminary validation and presents the MRQ as an instrument holding considerable promise for use within the Singapore context, further research might be needed to support adoption on a widespread basis. As a concluding example, other Rasch analysis software could be applied to ascertain the findings of this study.

Consent to participate
Consent to participate was obtained from all participants either in hard copy (in the case of qualitative interviews) or online (in the case of online instrument delivery) format.

Consent for publication
Consent for publication was obtained in the same format as for consent to participate.

Ethics approval
Approvals to conduct this research were obtained from the Human Research Ethics Committee of the University of Western Australia (RA/4/1/7813) and from the Ministry of Education, Singapore (RQ105-15(09)).

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD
Lyndon Lim https://orcid.org/0000-0002-8199-5761

References
American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association.
Andrich, D. (1982). An index of person separation in latent trait theory, the traditional KR-20 index, and the Guttman scale response pattern. Education Research and Perspectives, 9(1), 95–104. https://www.rasch.org/erp7.htm
Andrich, D. (1997). Georg Rasch in his own words. Rasch Measurement Transactions, 11(1), 542–543. https://www.rasch.org/rmt111i1.htm
Andrich, D., Lyne, A., Sheren, R., & Luo, G. (2018). Interpreting RUMM2030 part II polytomous data. RUMM Laboratory Pty Ltd.
Andrich, D., & Marais, I. (2019). A course in Rasch measurement theory: Measuring in the educational, social and health sciences. Springer.
Bond, T. G., & Fox, C. M. (2015). Applying the Rasch model: Fundamental measurement in the human sciences (3rd ed.). Routledge.
Boone, W. J. (2016). Rasch analysis for instrument development: Why, when, and how? CBE Life Sciences Education, 15(4), 1–7. https://doi.org/10.1187/cbe.16-04-0148
Dogan, N., Hambleton, R. K., Yurtcu, M., & Yavuz, S. (2018). The comparison of differential item functioning predicted through experts and statistical techniques. Cypriot Journal of Educational Sciences, 13(2), 137–148. https://doi.org/10.18844/cjes.v13i2.2427
Douglas, J., Roussos, L., & Stout, W. (1996). Item-bundle DIF hypothesis testing: Identifying suspect bundles and assessing their differential functioning. Journal of Educational Measurement, 33(4), 465–484. http://www.jstor.org/stable/1435335 https://doi.org/10.1111/j.1745-3984.1996.tb00502.x
Ewing, M. T., Salzberger, T., & Sinkovics, R. R. (2005). An alternate approach to assessing cross-cultural measurement equivalence in advertising research. Journal of Advertising, 34(1), 17–36. 10.1080/00913367.2005.10639181
Hendriks, J., Fye, S., Styles, I., Skinner, S. R., & Merriman, G. (2012). Scale construction utilising the Rasch unidimensional measurement model: A measurement of adolescent attitudes towards abortion. Australasian Medical Journal, 5(5), 251–261. https://doi.org/10.4066%2F2AMJ.2012.952
Kohlberg, L. (1984). The psychology of moral development: The nature and validity of moral stages (essays on moral development, volume 2). Harper & Row.
Kohli, N., Koran, J., & Henn, L. (2015). Relationships among classical test theory and item response theory frameworks via factor analytic models. Educational and Psychological Measurement, 75(3), 389–405. https://doi.org/10.1177/0013164414559071
Lim, L., & Chapman, E. (2021c). Development and preliminary validation of the Moral Reasoning Questionnaire for secondary school students. SAGE Open. https://doi.org/10.1177/21582440221085271
Lim, L., & Chapman, E. (2021b). Moral reasoning assessment for Singapore secondary schools: A review. Issues in Educational Research, 31(4), 1121–1137. http://www.iier.org.au/iier31/lim.pdf
Lim, L., & Chapman, E. (2021a). Moral reasoning in secondary education curriculum: An operational definition. International Journal of Ethics Education. https://doi.org/10.1007/s40889-021-00129-z.
Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the hospital anxiety and depression scale (HADS). British Journal of Clinical Psychology, 46(1), 1–18. https://doi.org/10.1348/014466506 X96931
Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Danmarks Paedagogiske Institut.

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

Singapore Ministry of Education. (2012). 2014 syllabus, character and citizenship education (secondary). https://www.moe.gov.sg/-/media/files/programmes/2014-character-citizenship-education-secondary.pdf?la=en&hash=F4B55203C3F0B328B45A6E703E7529AC2C0FA0B

Singapore Ministry of Education. (2016). Character and citizenship syllabus (pre-university). https://www.moe.gov.sg/-media/files/secondary/syllabuses/2016-character-and-citizenship-education-syllabus-secondary.pdf

Smith, E. V. (2002). Detecting and evaluating the impact of multi-dimensionality using item fit statistics and principal component analysis of residuals. Journal of Applied Measurement, 3(2), 205–231.

Steiner, D. L., Norman, G. R., & Cairney, J. (2015). Health measurement scales: A practical guide to their development and use (5th ed.). Oxford University Press. https://doi.org/10.1093/med/9780199685219.001.0001

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Care and Research, 57(8), 1358–1362. https://doi.org/10.1002/art.23108

Tennant, A., & Pallant, J. (2006). Unidimensionality matters! (A tale of two Smiths?) Rasch Measurement Transactions, 20(1), 1048–1051. https://www.rasch.org/rmt/rtm201c.htm

Teresi, J. A., Wang, C., Kleinman, M., Jones, R. N., & Weiss, D. J. (2021). Differential item functioning analyses of the patient-reported outcomes measures information system (PROMIS®) measures: Methods, challenges, advances, and future directions. Psychometrika. https://doi.org/10.1007/s11336-021-09775-0

Thoma, S. J., & Dong, Y. (2014). The defining issues test of moral judgment development. Behavioral Development Bulletin, 19(3), 55–61. https://doi.org/10.1037/h0100590

Willis, G. (2017). Cognitive interviewing in survey design: State of the science and future directions. In D. Vannette & J. Krosnick (eds.), The Palgrave handbook of survey research. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-54395_6_14

Wu, H., & Leung, S. (2017). Can Likert scales be treated as interval scales? A simulation study. Journal of Social Service Research, 43(4), 527–532. https://doi.org/10.1080/01488376.2017.1329775

---

**Appendix A: Questionnaire**

**1A.** You are running for president of the Students’ Council, of which you are currently treasurer. The current president is struggling with her school work but is also running for presidency. At the elections, you see one of your friends whom you know hasn’t been attending council meetings regularly and is not supposed to vote. After the elections, this friend tells you he voted for you. The next day, the results are announced and you win by one vote. You recall that the friend who was not supposed to vote voted for you. You would…

- report the truth.
- not report the truth.

**1B1.** Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the truth. You would …

- _____ not report the truth as you might be disqualified for the presidency if you were found for hiding the truth.
- _____ report the truth as it is your duty to do so as the current treasurer, and it is against the Students’ Council rule to hide the truth.
- _____ report the truth as it is important that the elections are conducted fairly.

**1B2.** Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the truth. You would …

- _____ not report the truth as you would benefit more being a president than a treasurer. You believe you can be a better president than the current one.
- _____ not report the truth as you think that the current president will need more time to focus on her school work next year.
- _____ not report the truth as everyone has the right to pursue his/her goal.

**2A.** You are a prefect. During a test, you saw your best friend cheating. This is not the first time it has happened. You know that the parents of this friend have high expectations and your friend will be caned for poor results or if they knew about the cheating. Your best friend would definitely hate you for reporting the cheating. You would…

- report the cheating.
- not report the cheating.

**2B1.** Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the cheating. You would …

- _____ report the cheating as you might be removed from the prefectural board if you were found for not reporting it.
report the cheating as you want your friend to learn what is right.
report the cheating as it is important to be fair to everyone who took the test. (3)

2B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the cheating. You would …

not report the cheating as you would not want your best friend to hate you.
not report the cheating as you don’t want your best friend to be caned.
not report the cheating as everyone has a right to do what is best for himself.

3A. Your friend and you are in the same class. Your friend is very weak in Biology but better than you in other subjects. You just found out that she copied a Biology essay online and submitted it for an assignment. She asked you to keep quiet about it. This assignment is important and if she fails it, she will be retained for a year and she would be sad. You would …

report that your friend copied the essay.
not report that your friend copied the essay.

3B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the copying. You would …

report the matter to the teacher as she might give you a zero for the essay if she found out that you ignored the act of copying.
report the matter as copying is not allowed according to the school rules.
report the matter as it is important that everyone in the class is assessed fairly.

3B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the copying. You would …

not report the matter so that your friend won’t hate you.
not report the matter as you should be sincere in the friendship and your friend asked you to keep it a secret.
not report the matter as it is everyone’s right to strive to be promoted.

4A. You are in the school badminton team. Your friend is as good as you in badminton and has won many matches for the school. Both of you are competing for the position of the badminton team captain. In the promotional exam, you saw your friend cheat. After asking him, you found that if he failed the exam, his parents would stop him from playing badminton. He asks you to keep it to yourself. You would …

• report the cheating.
• not report the cheating.

4B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the cheating. You would …

report the matter to the teacher. Hiding the truth from the teacher might cost you the position of team captain.
report the cheating to the teacher as school rules state that cheating is a serious offence.
report the cheating as it is more important that everyone is treated fairly for the exam.

4B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the cheating. You would …

not report the matter to the teacher as your teammates would be unhappy with you if they knew you did.
not report the matter to the teacher as your friend asked you to keep it to yourself.
not report the matter as your friend has a right to pursue his interest in badminton even if it meant cheating for the promotional exam.

5A. In the class online chat, you noticed some of your friends talking bad about a classmate and they seem to be enjoying it. You would …

report this matter.
not report this matter.

5B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report this matter. You would …

report it to the teacher so that you might not face detention class if the online chat was found out.
report this issue to the teacher as the school encourages us not to badmouth others.
report this if your friends continue talking bad about the classmate. It is more important that your friends learn to respect others.
5B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report this matter. You would …

______ not report the remarks as they may not be true and more importantly, it wouldn’t benefit you if your classmates found out you reported it.
______ not report this as everyone seems to be enjoying the chat. After all, the chat topic will soon die off.
______ not report this as everyone has the right to discuss about their opinions of others.

6A. During recess, you realised you forgot to do your Mathematics assignment. With only 15 min left to the next lesson, your best friend offers you his assignment to copy as he doesn’t want you to be scolded. Your Mathematics teacher is strict and will definitely inform your parents if you admitted you forgot to do the assignment as this is not your first time. You would …

• copy the assignment.
• not copy the assignment.

6B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to copy the assignment. You would …

______ copy your best friend’s assignment and submit it so that you won’t be scolded.
______ copy your friend’s assignment as the teacher does not allow any late submission and school rules state the assignments should be submitted on time.
______ copy your best friend’s assignment and submit it on time so that your parents and teacher would not be disappointed with you.

6B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to copy the assignment. You would …

______ not copy your best friend’s assignment as you might be punished if you were found out.
______ not copy your best friend’s assignment as it is against the school rule to copy assignments.
______ not copy but do what you can within 15 min and submit the assignment as it is more important to be honest.

7A. You are in a debate team and your class is expecting your team to win. Your friend is on an opposing team and the debate will take place tomorrow. You sit beside this friend and can see his debate notes. His notes could help your team win. You would…

• copy his notes.
• not copy his notes.

7B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to copy his notes. You would …

______ copy his notes but use it for yourself instead of sharing it entirely with your team so that you might have a chance to be the best speaker.
______ copy and share his notes with your team to have a better chance of winning.
______ copy his notes and share with your team as both teams have been given equal opportunities to prepare for the debate whatever it takes. It is a fair fight.

7B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to copy his notes. You would …

______ not copy his notes as you might be punished and disqualified if you were found out.
______ not copy his notes as the school does not allow copying without acknowledging the source.
______ not copy his notes nor take a look as both teams should have equal chances to win the debate.

8A. You are a prefect standing at the school gate to record the names of latecomers. The gate is just behind the assembly area. You best friend John arrives at the gate just as you shut the gate. You know he could be late as his parents’ quarrel affected him as he told you over the phone last night. You would …

• allow John in.
• not allow John in.

8B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to allow John in. You would …

______ allow John in by pretending to look away and not shutting the gate so that you won’t be reprimanded by the Prefect Master for letting a latecomer in.
______ allow John in. If your classmates see you allowing John in, they wouldn’t dislike you.
______ allow John in as he wasn’t late intentionally. It wasn’t really his fault.

8B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to allow John in. You would …
not allow John in as you might be punished by the teacher if she saw it.
not allow John in as you are duty bound as a prefect.
not allow John in as it is important to be fair to every latecomer.

9A. The class beside yours took the same Science test 2 h ago. Over recess, you heard your friend Jane asking what the questions were. Even though this is just a class test, Jane says that she needs whatever help to boost her failing results. Jane normally sits beside you during a test. You would …

• report Jane to the teacher.
• not report Jane to the teacher.

9B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report Jane to the teacher. You would …

_____ report Jane to the teacher as you might be given a zero if your teacher found out that you knew but didn’t report the matter.
_____ report Jane to the teacher as the school does not allow cheating.
_____ report Jane to the teacher as it is more important to ensure fairness in the test for everyone.

9B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report Jane to the teacher. You would …

_____ not report but ask Jane what the questions were and discuss the answers with her so that both she and you would benefit and she would value you as a friend more.
_____ not report but ask Jane to share the questions with the class so that everyone in the class can do well.
_____ not report Jane as this could be a morale booster for her to work harder in future tests.

10A. You are the class monitor. Your teacher tasked you to divide the class equally and randomly into groups for project work. You know there are some classmates who are notorious and will not contribute to a group. You would …

• divide the class randomly.
• divide the class randomly with conditions.

10B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to divide the class randomly. You would …

_____ divide the class randomly for fear of being accused and later punished by the teacher for being biased.
_____ divide the class randomly as the teacher instructed you to do so.
_____ divide the class randomly as it is important to be fair to everyone.

10B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to divide the class randomly with conditions. You would …

_____ divide the class randomly but for your own group you would ensure that members would only be those you can work with.
_____ divide the class randomly based on your classmates’ preferences so that they would be happy with you.
_____ divide the class randomly but ensure that each group has a classmate known to be notorious. Everyone should learn how to work with difficult people.

11A. Your teacher asked you to connect her laptop to the projector and then left the class for a while. While setting up the laptop, you saw your final exam marks on it. You noticed that your mark was 1 short of an A1 grade. An A1 grade could make a difference on your streaming for the following year and your parents will scold you if you don’t make it to the top stream. You would …

• amend the marks on the laptop.
• not amend the marks on the laptop.

11B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to amend the marks on the laptop. You would …

_____ change the mark on the laptop so that you won’t be scolded by your parents.
_____ not touch the laptop in case your classmates disapproved of your action.
_____ add 1 mark to everyone in the class so that everyone gets 1 mark more. This would be fairer to everyone.

11B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to change the marks on the laptop. You would …
12A. You are a prefect. You see your close friend who always helps you in your homework smoking in the toilet but he asks you to keep it to yourself. This is not his first time. You know that this classmate has abusive parents and if you report him, he would be suspended and his parents will hit him. You would …

• report him to the teacher.
• not report him to the teacher.

12B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report him to the teacher. You would …

• report him. If you don’t report him and the teacher finds out, you might be reprimanded and lose the position of prefect.
• report him to the teacher as you are duty bound as a prefect and the school rules do not allow smoking in school.
• report him as it is more important that he leads a healthier lifestyle.

12B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report him to the teacher. You would …

• not report him so that you won’t lose him as a close friend.
• not report him so that his parents would not abuse him or feel hurt. He would also not feel worried about suspension.
• not report him but ask him to admit to the teacher that he smoked in the toilet and ask the teacher for help to kick the habit.

13A. Your class has been asked to give feedback about your Science teacher using a feedback form. You are required to write your name on the feedback form for identification. The school requires that you be truthful in this feedback. Many classmates and you feel that this teacher is biased and impatient but most of your classmates intend to give nice and false feedback. You would …

• give true feedback.
• not to give true feedback.

13B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to give true feedback. You would …

• give true feedback so that you won’t be criticised by anyone for being dishonest.
• give true feedback as the school values honesty.
• give true feedback as it is more important for the teacher to know the truth and improve.

13B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to give true feedback. You would …

• give false and nice feedback so that you would not be marked by the teacher.
• give false and nice feedback so that you will not feel different from your peers.
• give false feedback as this teacher might only be hurt unnecessarily if everyone gives harsh and true feedback.

14A. You are the Chairman of a uniform group. The Vice-Chairman just left the uniform group as he transferred to another school. Your teacher is asking you for a name to fill the position of Vice-Chairman. Your close friend Samy is qualified and he told you he wants to assume the position. Yet, there is another schoolmate who is more qualified than Samy. Your teacher will likely accept whichever name you suggest as she is new to the school. You would …

• suggest Samy’s name.
• not suggest Samy’s name.

14B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to suggest Samy’s name. You would …

• suggest Samy’s name as the Vice-Chairman. If Samy becomes Vice-Chairman, you won’t have any trouble working with him as you know him well.
• suggest Samy’s name as the Vice-Chairman as Samy is your good friend and he told you he wants to be Vice-Chairman.
• suggest Samy’s name as the Vice-Chairman. You know that Samy can co-lead the uniform group well with you and bring it to greater heights.

14B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to suggest Samy’s name. You would …
15A. You just received your report book and noticed you got an A1 for Mathematics, different from the A2 grade you got for the actual paper. Apparently, your teacher keyed A1 into the report book by mistake. Some friends sitting around you saw the A1 grade and said it was definitely an error; they said you should just let it be in case their good grades were also errors made by the teacher. An A1 grade for Mathematics will help you get into the Science stream and your parents will be delighted. You would …

- report the error.
- not report the error.

15B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the error. You would …

- not report the error to the teacher. If she found out later that you knew but didn’t report the error, you would be scolded for being dishonest.
- report the error to the teacher as the school rules state that one must be honest.
- report the error to the teacher as it is important that the grading is accurate and fair to everyone.

15B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the error. You would …

- not report the error as it wasn’t your fault and you wouldn’t be blamed for the error.
- not report the error since your friends told you to let it be. You wouldn’t want to be blamed for any change in grade for your friends’ grades if the teacher found other mistakes in the class’s results.
- not report the error. You feel that there’s more benefit to leave the Mathematics grade as A1; you can get into the Science stream and your parents will be happy.

16A. You work part-time in a fairly expensive restaurant with your classmates. Your job is to wash dishes. After working in the restaurant for just a day, you are thoroughly disgusted. You have seen that the kitchen is very dirty and has an infestation of cockroaches. You mention this to the restaurant manager who just smiles and walks away. You would …

- report the matter to the authorities.
- not report the matter to the authorities.

16B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the matter. You would …

- not report the matter to the relevant authorities quickly so that you won’t be criminalised if the dirty kitchen was found out by authorities.
- report the matter to the relevant authorities as the law does not permit such unhygienic practices.
- report the matter as it is important that all diners consume clean food.

16B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the matter. You would …

- not report the matter to the relevant authorities so that you won’t be scolded by the restaurant manager.
- not report the matter otherwise your classmates may dislike you for causing them a loss of income should the restaurant be forced to close.
- not report it but do your best to clean the kitchen as it is more important that the kitchen be properly cleaned than authorities clamping down on the restaurant business.

17A. You are unprepared for a test today. To buy time, you consider feigning illness so that you can take it another day when you have more confidence. You know the teacher will not let you take the test another day unless you have a valid reason. You would …

- feign illness to skip the test.
- not feign illness to skip the test.

17B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to feign illness and skip the test. You would …

- feign illness to skip the test, have more time to prepare for it and score better than the rest.
- feign illness to skip the test. This way, you can have more preparation time and your parents or teacher won’t be disappointed with you for not doing well for the exam.
- feign illness to skip the test and have more time to prepare for it as it doesn’t hurt anyone to do so.
17B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to feign illness and skip the test. You would …

- ______ not feign illness to skip the test as you might be punished if you were found out.
- ______ not feign illness as it is against the school rule to lie and skip the test without a valid reason.
- ______ not feign illness to skip the test as it is important to be fair to everyone.

18A. Some of your friends consume alcohol on their 16th birthday parties. You attend many of these parties. Your parents do not allow you to drink but you tell them it is difficult to reject your friends at so many parties. Though your parents ask you to tell your friends you can’t drink due to medical reasons, you feel the real reason is that they don’t want to be seen as over-protective. You would…

- • take alcoholic drinks.
- • not take alcoholic drinks.

18B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to take alcoholic drinks. You would …

- ______ drink and enjoy the company but keep it from your parents. If your parents found out, they might ground you for a few days.
- ______ drink and enjoy the company at selected birthday parties and explain your intention to your parents. This way, both your friends and parents won’t get totally upset.
- ______ drink as you don’t see anything wrong with drinking as long as you do not overdo it.

18B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to take alcoholic drinks. You would …

- ______ not drink no matter how to avoid scolding and punishment by your parents.
- ______ tell your friends you won’t drink as the legal age for alcohol consumption is 18.
- ______ tell your friends you can’t drink as your parents are concerned with your health.

19A. You failed your test. Your parents are angry and order you to stay home. This evening, they are away for a wedding dinner and you know they won’t be back till late night. Your friends ask you out for dinner and really hope you will join them. You know that your parents will likely not find out if you went out and you feel that being grounded because you failed a test is unfair. You would …

- • go for dinner with your friends.
- • not go for dinner with your friends.

19B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to go for dinner with your friends. You would …

- ______ go for dinner with your friends and come home way before your parents are home so that you won’t get into trouble.
- ______ go for dinner with your friends so that your friends won’t be unhappy.
- ______ go for dinner with your friends because it is not right to ground you because you failed a test.

19B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to go for dinner with your friends. You would …

- ______ not go for the dinner in case your parents find out and punish you.
- ______ not go for dinner as you have to abide by house rules set by your parents.
- ______ not go for dinner because you can stay home and study and not disappoint your parents.

20A. Your brother spent all his money on gaming. Just as he asks you for money to buy textbooks as his exams are next week, you find a pouch containing $500. There is a name inside the pouch but you do not know the owner. You would …

- • report the lost pouch to the police.
- • not report the lost pouch to the police.

20B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the lost pouch to the police. You would …

- ______ report the lost pouch to the police in case someone saw you taking it and you could get into trouble.
- ______ report the lost pouch to the police as taking without permission is similar to stealing which is against the law.
- ______ report the lost pouch to the police as the owner may need the money urgently and source for alternative funding for your brother.
20B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the lost pouch to the police. You would …

- keep the $500 for your brother as he would be grateful to you.
- keep the $500 for your brother as he needs it urgently for his exams.
- take $250 for your brother and report the remaining amount and lost pouch to the police. At least, you would be happy that you were able to help your brother.

21A. You just bought a new handphone (e.g. iPhone 6+) using your parents’ money. As you were walking out of the shop, you accidentally dropped it and the screen cracked. No one saw you drop the phone. You would …

- return the handphone and tell the shop you found a crack on it. Ask for a one-for-one exchange.
- not return the handphone. Just live with the crack.

21B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to return the handphone. You would …

- return the handphone, say you found a crack on it and ask for a new one so that your parents won’t scold you for being clumsy.
- return the handphone, say you found a crack on it and ask for a new one so that you wouldn’t be wasting your parents’ hard-earned money.
- return the handphone, say you found a crack on it and ask for a new one as everyone has the right to make his/her hard-earned money worth and well-spent.

21B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to return the handphone. You would …

- not return the handphone so that your parents won’t scold you for being dishonest.
- not return the handphone as insisting that that you found a crack on it and asking for a new one is unlawful and can be considered as cheating.
- not return the handphone as it is more important to be fair to the store that sold you the handphone.

22A. Your parents do not earn much. You just bought a textbook from the school bookstore and noticed that the cashier returned you more change than she was supposed to. You would …

- return the excess change.
- not return the excess change.

22B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to return the excess change. You would …

- return the excess change as you do not want to be caught and sent for detention class for being dishonest.
- return the excess change as it is against the school rules to be dishonest.
- return the excess change as it is more important to be honest than to save some money.

22B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to return the excess change. You would …

- keep the excess change since nobody would find out you took it and you won’t get into any trouble.
- use the excess change to buy another textbook as you can help your parents save money and they would be happy.
- use the excess change to make a donation to the needy.

23A. Your friends and you join a science quiz. Depending on how many questions are answered correctly, the winning team will be awarded up to $10000 and they can give it a charitable organisation of their choice. 2 days before the quiz, you find a list of questions for the quiz on the table at the venue. Your team has been struggling to learn the science concepts included in the quiz. You would …

- note the list of questions.
- ignore the list of questions.

23B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to note the list of questions. You would …

- note the list of questions but not tell your team about it so that it doesn’t reflect poorly on you. During the quiz, you would try to answer as many questions as possible so that your team can win most of the $10,000.
- tell your team the questions so that the team won’t feel so stressed about preparing for the quiz.
- note the list of questions and get your team to win most of the $10,000 for the sake of charity. It is more important to win more than less for charity.
23B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to ignore the list of questions. You would …

______ ignore the list of questions as you might be caught and disqualified for cheating even if it is for charity.
______ ignore the list of questions as this is against competition rules.
______ ignore the list of questions as it is important to be fair to the other team and the prize sponsor.

24A. You are at a hawker centre. You see an old man selling curry puffs by the stairway. He isn’t licensed to sell but still does it for a living. His curry puffs could be dirty and cause food poisoning. You would …

• report the old man to the authorities.
• not report the old man to the authorities.

24B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to report the old man to the authorities. You would …

______ report the old man to the authorities as you wouldn’t want to be criticised for not reporting such unlicensed practices to authorities.
______ report the old man to the authorities as it is not right for people to sell food without a license by law.
______ report the old man as you would be helping other potential customers avoid eating the dirty curry puffs.

24B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to report the old man. You would …

______ not report the old man as you normally don’t eat his curry puffs and you don’t want to be in any unnecessary trouble.
______ not report the old man so that you won’t be seen by others as a heartless person.
______ not report the old man but speak to him and offer help in other ways like referring him for financial aid.

25A. You are 5 min to being late for school. Just when you are about to enter the school gate, you see an old man fall down. There is no one at the school gate to help him except you. The nearest person is about 100 metres away at the bus stop. You have been late for school before and you know that if you are late for school this time, you will be punished. You would …

• help the old man.
• not help the old man.

25B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to help the old man. You would …

______ help the old man as you do not want the public at the bus stop to criticise you for being unhelpful.
______ help the old man as you are wearing the school uniform and what you do reflects on the school.
______ help the old man even if you were late because his safety is more important.

25B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to help the old man. You would …

______ not help the old man as you don’t want to be punished for being late. Besides, others will see and come and help him.
______ not help the old man as the school rules require students to be punctual.
______ not help the old man as you do not want to disappoint your parents and teachers in being late again.

26A. You often eat with your family at a hawker centre. Each time you eat there, a few people will approach you and sell packets of tissue paper even though you have tissue packets. Some of these people are handicapped or old and others look young. You know that these people are not licensed to sell tissue paper. You would …

• buy tissue paper from anyone who approaches you.
• not buy tissue paper from anyone who approaches you.

26B1. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding to buy tissue paper from anyone who approaches you. You would …

______ buy tissue from anyone who approaches as you do not want people around to criticise you for being unhelpful.
______ buy tissue from anyone who approaches as they will be happy that someone is buying from them.
______ buy tissue from anyone who approaches you as they are just earning a living.

26B2. Rank the following options in order of importance (1 = most important and 3 = least important) when deciding not to buy tissue paper from anyone who approaches you. You would …

______ not buy tissue from anyone who approaches you as you might get into trouble if authorities see you supporting unlicensed sellers.
_____ not buy tissue from anyone who approaches you as the law doesn’t allow unlicensed sellers. _____ not buy tissue from anyone who approaches you so that licensed sellers such as shops around the hawker centre have a chance to earn a living.

Appendix B: Parallel analysis and exploratory and confirmatory factor analyses results of preliminary validation

Figure B1. Parallel analysis graphical output suggesting the unidimensionality of the MRQ.

| Item      | Loading on factor 1 | Item      | Loading on factor 1 |
|-----------|---------------------|-----------|---------------------|
| A27_score | 0.65828             | A10_score | 0.47641             |
| A14_score | 0.61441             | A4_score  | 0.47540             |
| A21_score | 0.57614             | A25_score | 0.45869             |
| A26_score | 0.56450             | A8_score  | 0.45352             |
| A24_score | 0.55848             | A28_score | 0.44563             |
| A18_score | 0.55351             | A3_score  | 0.44528             |
| A16_score | 0.52406             | A17_score | 0.42846             |
| A30_score | 0.52399             | A7_score  | 0.41866             |
| A12_score | 0.51746             | A2_score  | 0.40766             |
| A23_score | 0.50437             | A22_score | 0.39791             |
| A19_score | 0.49210             | A9_score  | 0.39659             |
| A15_score | 0.47729             | A6_score  | 0.38961             |
| A29_score | 0.47644             | A1_score  | 0.34928             |

Figure B2. Results of exploratory factor analysis of items performed in preliminary validation.
Appendix C Initial Rasch analyses – Disordered thresholds

| Estimation method | $\chi^2$  | df  | $\chi^2$/df | CFI | RMSEA | SRMR |
|-------------------|-----------|-----|-------------|-----|--------|------|
| Maximum likelihood with Satorra-Bentler scaled chi-square statistic | 459.62*   | 299 | 1.53        | 0.93 | 0.04   | 0.05 |

*Note: $\chi^2$ = Satorra-Bentler scaled chi-square statistic; df = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square

*p < .001

Figure B3. Results of confirmatory factor analysis of items performed in preliminary validation.

Figure C1. Threshold map from initial Rasch analysis of 30 items.
Figure C2. Item 8, with ordered thresholds.

Figure C3. Item 14, with disordered thresholds. The category characteristic curves presented in Figure C2 suggest that respondents with higher moral reasoning dispositions (i.e., logits on the log-linear scale) in relation to Kohlberg’s stage-base theory of moral development tended to have a higher probability of scoring higher for item 8. On the contrary, Figure C3 suggests that respondents with higher moral reasoning dispositions did not necessarily express a higher probability of scoring higher for item 14 that was found with disordered thresholds.