2014

Measuring students' perceptions of plagiarism: Modification and Rasch validation of a plagiarism attitude scale

Steven J. Howard
University of Wollongong, stevenh@uow.edu.au

John F. Ehrich
Monash University, jehrich@uow.edu.au

Russell Walton
University of Wollongong, rwalton@uow.edu.au

Publication Details

Howard, S. J., Ehrich, J. F. & Walton, R. (2014). Measuring students' perceptions of plagiarism: Modification and Rasch validation of a plagiarism attitude scale. Journal of Applied Measurement, 15 (4), 372-393.

Measuring students' perceptions of plagiarism: Modification and Rasch validation of a plagiarism attitude scale

Abstract
Plagiarism is a significant area of concern in higher education, given university students' high self-reported rates of plagiarism. However, research remains inconsistent in prevalence estimates and suggested precursors of plagiarism. This may be a function of the unclear psychometric properties of the measurement tools adopted. To investigate this, we modified an existing plagiarism scale (to broaden its scope), established its psychometric properties using traditional (EFA, Cronbach's alpha) and modern (Rasch analysis) survey evaluation approaches, and examined results of well-functioning items. Results indicated that traditional and modern psychometric approaches differed in their recommendations. Further, responses indicated that although most respondents acknowledged the seriousness of plagiarism, these attitudes were neither unanimous nor consistent across the range of issues assessed. This study thus provides rigorous psychometric testing of a plagiarism attitude scale and baseline data from which to begin a discussion of contextual, personal, and external factors that influence students' plagiarism attitudes.

Keywords
modification, rasch, validation, plagiarism, attitude, perceptions, scale, students, measuring

Disciplines
Education | Social and Behavioral Sciences

Publication Details
Howard, S. J., Ehrich, J. F. & Walton, R. (2014). Measuring students' perceptions of plagiarism: Modification and Rasch validation of a plagiarism attitude scale. Journal of Applied Measurement, 15 (4), 372-393.

Measuring Students' Perceptions of Plagiarism: Modification and Rasch Validation of a Plagiarism Attitude Scale

Steven J. Howard[1], John F. Ehrich[2], and Russell Walton[1]

[1]School of Education, University of Wollongong, Wollongong, NSW, Australia, 2522

[2]Faculty of Education, Monash University, Melbourne, VIC, Australia, 3800

email addresses: stevenh@uow.edu.au, john.ehrich@monash.edu, rwalton@uow.edu.au

Corresponding author: Steven Howard, School of Education, University of Wollongong, Wollongong, NSW 2522, Australia; email stevenh@uow.edu.au; Phone +61 2 4221 5165.

Keywords: Rasch analysis; factor analysis; plagiarism; attitudes; university; education

Abstract

Plagiarism is a significant area of concern in higher education, given university students' high self-reported rates of plagiarism. However, research remains inconsistent in prevalence estimates and suggested precursors of plagiarism. This may be a function of the unclear psychometric properties of the measurement tools adopted. To investigate this, we modified an existing plagiarism scale (to broaden its scope), established its psychometric properties using traditional (EFA, Cronbach's alpha) and modern (Rasch analysis) survey evaluation approaches, and examined results of well-functioning items. Results indicated that traditional and modern psychometric approaches differed in their recommendations. Further, responses indicated that although most respondents acknowledged the seriousness of plagiarism, these attitudes were neither unanimous nor consistent across the range of issues assessed. This study thus provides rigorous psychometric testing of a plagiarism attitude scale and baseline data from which to begin a discussion of contextual, personal, and external factors that influence students' plagiarism attitudes.

Measuring Students' Perceptions of Plagiarism: Modification and Rasch Validation of a Plagiarism Attitude Scale

The ability to demonstrate an understanding of course content by communicating core concepts is an important skill expected of all university graduates. Mastery of this skill requires that students be able to extract and synthesize ideas from the existing literature and, in doing so, distinguish synthesis from personal evaluation in order to properly attribute ideas to their source. However, there is evidence that a growing proportion of university students are engaging in forms of academic misconduct that circumvent the acquisition and demonstration of these skills (e.g., misappropriating ideas without proper acknowledgment of the source, cheating on an exam). Specifically, estimates suggest that as many as three-quarters of university students have engaged in at least some form of academic misconduct in their academic career (Brimble and Stevenson-Clarke, 2005; McCabe and Bowers, 1994).

One form of academic misconduct that has received significant attention is plagiarism. Although there is no universally accepted definition of plagiarism, it is generally agreed that it entails using the ideas, words, or works of another without appropriate acknowledgment of their source (Gibaldi, 2003). This can include misappropriating others' ideas (paraphrasing others' ideas without proper acknowledgment of their source), 'cut and paste' plagiarism (importing unquoted, unattributed, and nearly/completely word-for-word excerpts from another source into an academic assessment), or misappropriating entire works (submitting work completed entirely by another, such as ghost-written essays). Definitions of plagiarism are also increasingly acknowledging self-plagiarism as a form of academic misconduct, in an effort to deter resubmission of students' previous works (Bretag and Mahmud, 2009; Walker, 1998).

Plagiarism has remained a significant area of concern for researchers and educators since an early study of university students' self-reported academic misconduct (Bowers, 1964)

found that three-quarters of respondents had engaged in academically dishonest practices. Subsequent replications and extensions of these findings have provided similar results, with some forms of academic misconduct actually on the rise (Brimble and Stevenson-Clarke, 2005; McCabe and Bowers, 1994). Although these studies capture forms of misconduct that extend beyond the misappropriation of source material, plagiarism prevalence estimates show similarly worrying trends. For instance, self-reports of plagiarism indicate that anywhere from 30% to 60% of students have engaged in some form of plagiarism as least once in their academic career, with more than 10% admitting to engaging in substantial amounts of plagiarism (Breen and Maassen, 2005; Hughes and McCabe, 2006; McCabe, 2005; McCabe and Trevino, 1993, 1995, 1996; McCabe, Trevino, and Butterfield, 2001, 2004; McGowan, 2005; Selwyn, 2008; Underwood and Szabo, 2003). It has been further suggested that these estimates may, in fact, underestimate the true prevalence of plagiarism, given that they are derived from students' self-reported transgressions (Selwyn, 2008).

Although there is little doubt about the troubling prevalence of plagiarism at the tertiary level (despite the variability in specific prevalence estimates), the precursors of plagiarism are less clear. For instance, it has been suggested that the factors exacerbating plagiarism are both individual and institutional in nature. Specifically, individual factors found to relate to plagiarist behaviors include age, sex, personal/professional pressures, grade point average, language skills, cultural background, competitiveness, and self-efficacy (Ashworth, Bannister, and Thorne, 1997; Breen and Maassen, 2005; Devlin and Gray, 2007; Handa and Power, 2005; Harris, 1995; Martin, Rao, and Sloan, 2011; Maxwell, Curtis, and Vardanega, 2008; Perry, Kane, Bernesser, and Spicker, 1990; Song-Turner, 2008; Underwood and Szabo, 2003). Institutional factors that have been suggested as potential precursors of plagiarism include mild or non-existent punishments for plagiarism, the absence of academic integrity policies, insufficient staggering of academic assessments, and a lack of explicit instruction in

the conventions of academic writing (Bamford and Sergiou, 2005; Bannister and Ashworth, 1998; Ellery, 2008; Gullifer and Tyson, 2010; Harris, 2001; Marshall and Garry, 2006; McCabe and Trevino, 1993; Park, 2003; Ryan, Bonanno, Krass, Scouller, and Smith, 2009; Sutherland-Smith, 2008; Zimitat, 2008). Other factors that have been suggested to exacerbate plagiarist behaviors include 'electronic opportunism' (capitalizing on the increasing electronic availability of scholarly works; Rocco and Warglien, 1995), students' declining perceptions of plagiarism as a serious offense (Baruchson-Arbib and Yaari, 2004), and an increasing perception that plagiarizing is a necessary 'equalizer', in the context of faculty members often turning a blind eye to students' plagiarism practices (McCabe, Trevino, and Butterfield, 2001). The relative primacy or redundancy of these various factors for explaining plagiarist behaviors remains unclear.

Are Inconsistent Findings a Function of the Measurement Instrument?

It may be the case that the lack of clarity around the prevalence of plagiarism (with estimates ranging from less than 30% to more than 60%) and precursors of plagiarism (encompassing a myriad of personal, institutional, and other external factors) is at least partly a function of the measurement tools adopted. That is, researchers investigating university students' attitudes toward plagiarism often adopt surveys as their main data collection instrument (Mavrinac, Brumini, Bilic-Zulle, and Petrovecki, 2010). In such studies, plagiarism is measured in terms of students' attitudes toward plagiarism and the perceived stressors that exacerbate plagiarist behaviors. However, students' differing definitions of plagiarism (which often deviate from formal academic definitions; Brimble and Stevenson-Clarke, 2005; Crisp, 2007) and a lack of precision in measurement may result in survey items measuring very different constructs. In the simplest case, a student with an accurate conception of what constitutes plagiarism (e.g., the various forms of misappropriation described earlier) may have very different opinions of the seriousness of plagiarism relative

to a student with misconceptions of plagiarism (e.g., believing that ‘cut and paste’ plagiarism is appropriate if there is an in-text reference, changing a few words in a sentence/paragraph is sufficient). To illustrate the point, in explaining plagiarism findings that run counter to public perceptions, McCabe (2005) remarked “a partial explanation may be that there is some confusion in the minds of students, and faculty, as to exactly what each question is seeking” (p. 6).

Although this is an issue that pervades all forms of survey research, there are steps that can be taken to evaluate the function of these measures. Specifically, psychometric analyses allow researchers to evaluate the accuracy (validity) and consistency (reliability) with which the survey instrument measures the underlying construct of interest (in this case, plagiarism). However, plagiarism research rarely includes psychometric analyses of the survey instrument (Ehrich, Howard, Tognolini, and Bokosmaty, 2013). In fact, Gururajan and Roberts (2005) argue that, given the absence of validity and reliability evidence of plagiarism surveys, there remains no concrete evidence of university students’ attitudes toward plagiarism in the Australian higher education context. Nevertheless, tertiary institutions continue to enact change on the basis of these potentially problematic findings (Scanlon, 2003).

While the majority of survey-based plagiarism research does not report the psychometric properties of their measurement instruments (e.g., Harris, 2001), there have been some recent efforts to rectify this situation (Gururajan and Roberts, 2005; Mavrinac et al., 2010). These studies adopt traditional test theory approaches to survey evaluation (e.g., exploratory factor analysis, Cronbach’s alpha), which evaluates the clustering of items into cohesive subscales and the reliability with which these subscales measure the underlying construct. They do not, however, evaluate individual item fit relative to this construct. Further, these approaches aim to describe the structure of the underlying data (modeling the data), rather than evaluating the structure the data should exhibit (fitting data to a theoretical model of optimal measurement;

Andrich, 1988, 2004). Traditional test theory approaches, therefore, can result in the derivation of scales that may appear psychometrically valid and reliable, yet contain items with significant misfit to the underlying construct. This raises issues for interpretation and subsequent actions on the basis of these results.

In contrast, Rasch modeling (a specific instance of item response theory; Tennant and Conaghan, 2007) evaluates the structure of responses relative to the specifications of the Rasch model. As this model is based upon the guiding principle of measurement invariance, measures found to demonstrate good fit to this theoretical model should function similarly well for all respondents. This is critical if the instrument is to be used to measure a construct in a valid and reliable manner across a range of samples (Hagquist, Bruce, and Gustavsson, 2009). Further, Rasch analysis identifies anomalies in the survey-generated data in order to refine the method of assessment (Ehrich et al., 2013). This permits identification of individual items that are problematic in their function (e.g., items that are poor indices of the latent construct under measure, items with problematic response categories, items that display bias amongst demographic subsamples, etc.). Rasch modeling thus governs the *construction* of measurement instruments (rather than *describing* the data collected), providing analytical data as to whether the items comprising the scale function as intended. Despite these distinct advantages over traditional test evaluation approaches, only one plagiarism study employing Rasch analysis could be found (Deckert, 1993), which investigated students' ability to recognize plagiarism. To date, there appears to be no measure of students' attitudes toward plagiarism that has been subjected to Rasch modeling, despite the increased confidence in the measurement instrument that such analyses can provide.

The Current Study

The current study aimed to evaluate the psychometric properties of a plagiarism attitudes scale (a modified version of the Attitudes Toward Plagiarism questionnaire; Mavrinac et al.

2010). Although the original scale displayed good psychometric properties among Croatian university students using traditional test theory methods, it remains unclear whether all items of the scale function in accordance with the Rasch model (that is, appropriately measure the constructs they are purported to measure). To evaluate this, the results of traditional (i.e., exploratory factor analysis, Cronbach's alpha) and modern scale evaluation techniques (i.e., Rasch analysis) were contrasted. This study thus sought to broaden the applicability of this scale, establish its psychometric properties using a modern test theory approach (i.e., Rasch modeling), and evaluate the results of administration of the scale. On the basis of these results, recommendations are made for future administrations of the instrument.

Method

Participants

Participants were 300 undergraduate Education students in attendance at pre-identified lectures in an Australian university. Most students were enrolled in their first (n = 117) or second year of study (n = 172), with a fewer number of third (n = 10) and fourth year students (n = 1). Three students failed to respond to the demographic questionnaire. Available demographic data indicated that there were significantly more females (n = 244) than males (n = 53), the ratio of which is representative of student enrolments in this faculty. The majority of students were young adults (females: M_{age} = 21.30, SD = 4.40; males: M_{age} = 23.00, SD = 6.10) ranging from 18 to 50 years in age.

Instrument

The original Attitudes Toward Plagiarism questionnaire (Mavrinac et al., 2010) consists of 29 statements designed to measure plagiarism attitudes and precursors among university students in science-based faculties. To respond, participants indicate their (dis)agreement with each item using a five-point Likert scale (i.e., 1 = *strongly disagree*, 2 = *disagree*, 3 = *neither agree nor disagree*, 4 = *agree*, 5 = *strongly agree*). Previous research (Mavrinac et

al., 2010) administering this scale to Croatian medical, psychological, and engineering students identified three subscales (i.e., positive attitudes toward plagiarism, negative attitudes toward plagiarism, and subjective plagiarism norms), each with acceptable psychometric properties (Cronbach's alpha for all subscales > .70). However, a range of contemporary plagiarism issues do not appear to be captured by this scale (e.g., facilitating plagiarism, intentionality) and a number of the scale's items appear to be more relevant to students in science-based faculties (e.g., "a plagiarized paper does no harm to science"; Mavrinac et al., 2010, p. 198). As such, a number of items were modified or removed to render the scale more applicable to the entire range of university students. In addition, a number of items were added to broaden the range of plagiarism issues assessed. To ensure participants' adoption of a common definition of plagiarism, the following definition preceded the plagiarism items: "using another person's ideas designs, words or works without appropriate acknowledgement. This includes taking a full sentence of text without properly quoting (i.e., using "quotation marks" and an in-text citation), insufficient paraphrasing (e.g., changing only a few words of a sentence), and self-plagiarism (i.e., copying from your own previous work)."

Statistical Analyses

Given the uncertainty around whether all items of this scale accurately and appropriately assess the subscale construct being measured, analyses sought to contrast the results of traditional and modern scale evaluation approaches. Specifically, exploratory factor analysis was first conducted to examine the factor structure of the underlying data. Cronbach's alpha was then used to investigate the reliability of the identified subscales. A series of Rasch analyses were then conducted to evaluate the scale using a modern test theory approach. This analysis was conducted using the polytomous Rasch model with partial credit parameterization (for a complete discussion of Rasch modeling, see Hagquist et al., 2009;

Masters, 1982; Tennant and Conaghan, 2007). While there were some grounds for adopting an alternative approach (e.g., limited sample size; Andrich, 1978), partial credit parameterization was chosen because no data have yet been collected on the functioning of the response categories of this scale. Finally, descriptive statistics for well-functioning items were examined to gain insight into students' attitudes toward plagiarism.

Results

Exploratory Factor Analysis (EFA)

As in the original instrument, eight items were reverse scored (see Table 1), so that for all items higher scores reflected softer (less serious) attitudes toward plagiarism. A subsequent exploratory factor analysis (EFA) using maximum likelihood estimation and an oblique (direct oblimin) factor rotation was conducted on the 36 items of the Revised Attitude Toward Plagiarism scale. The Kaiser-Meyer-Olkin (KMO) statistic indicated sufficient sampling, KMO = .892 ('good' according to common rules of thumb; Field, 2009), and KMO values for all individual variables were at acceptable levels (all >.5; Field, 2009). Bartlett's test of sphericity was significant, $X^2(630) = 3790.00$, $p < .001$, indicating that inter-item correlations were sufficiently large to justify EFA analysis.

EFA results identified nine eigenvalues greater than 1 (explaining 58.74% of the variance), however, factor loadings and the scree plot suggested a three-factor solution was optimal (explaining 37.85% of the variance). Further, only three factors explained at least 5% of the variance. The three-factor solution can be interpreted as: Factor 1 (*Factors that Exacerbate Plagiarism*) with 14 items (items 9, 12, 18, 21, 22, 23, 26, 28, 29, 30, 31, 33, 34, 36); Factor 2 (*Justification for Plagiarism*) with 6 items (items 1, 2, 3, 6, 11, 27); and Factor 3 (*Severity and Penalty*) with 9 items (items 4, 5, 7, 8, 15, 16, 17, 19, 25). All other items failed to load significantly (standardized factor loadings < .30) on any of these three factors

(see Table 1). Reliability analysis (Cronbach's alpha) suggested acceptable reliability for all subscales (Factor 1 = .89; Factor 2 = .72; Factor 3 = .79).

Rasch Analysis of the Full Scale and Subscales

A series of Rasch analyses subsequently were conducted, first on the 36 items of the full scale and then on each of the subscales identified by EFA. The polytomous Rasch model with partial credit parameterization was adopted, to provide detailed information on the functioning of the response categories of the scale. Analyses were run using Rasch Unidimensional Measurement Modeling (RUMM) 2030 software (Andrich, Sheridan, and Luo, 2010).

Working of the response categories. To examine the response categories of the full 36-item scale (to determine their concordance with the Rasch model), an examination of the ordering of thresholds (the points between response categories) was undertaken. Ordered thresholds reflect logical response choices relative to the person's level of the latent trait being measured (e.g., severity of attitude toward plagiarism), whereas disordered thresholds indicate misfit of the data to the Rasch model. Results revealed disorder in 22 of the 36 items. For all disordered thresholds, response categories were re-scored (i.e., collapsed into smaller categories) to improve fit to the Rasch model (Tennant and Conaghan, 2007).

Rescoring of items with disordered thresholds. Category characteristic curves (CCCs) were examined for all items with disordered thresholds. Disordering of thresholds between *Strongly Disagree* and *Disagree* was evident for 13 items (for an example CCC, see Figure 1), the collapsing of which resulted in the ordering of thresholds for each of these items. The remaining items with disordered thresholds were resolved by: collapsing *Strongly Agree* and *Agree* for five items; collapsing *Strongly Agree* and *Agree*, as well as *Disagree* and *Strongly Disagree* for two items; and dichotomising three items. This resulted in the ordering of the thresholds of all items (see Figure 2).

Overall fit to the model. An analysis of item fit statistics, which evaluates how well items function according to the expectations of the Rasch model, suggested poor overall fit. Specifically, the item–trait interaction (a measure of overall fit including unidimensionality assumptions) was statistically significant, $X^2(144) = 428.20$, $p < .001$, indicating poor overall fit of the data to the model. Further, the overall (unweighted) item fit residual ($M = 0.30$, $SD = 2.00$) indicated that the data did not provide good fit to the model (with good fit being indicated by a mean close to 0 and standard deviation close to 1). In contrast, the Person Separation Index (PSI) indicated good reliability (PSI = 0.89) and the overall person fit residual indicated good fit between responses and those expected by the Rasch model ($M = –0.21$, $SD = 1.24$).

Individual item fit. Individual item fit for all 36 questions is given in Table 2. Misfit was found in nine items, on the basis of mean fit residuals exceeding the acceptable range between –2.50 and +2.50 (all items), significant chi-square and *F* statistics indicating significant misfit (one item), and ICCs with a lack of congruence between the theoretical s-shaped curve and actual values (see Figure 3).

Targeting and sample size. The person–item location distributions for the full scale and all subscales are depicted in Figure 4. These distributions indicated that participants' responses were not evenly spread, with the majority of respondents having more severe attitudes toward plagiarism. The items, however, were sufficiently diverse to elicit a range of attitudes toward plagiarism. Although the sample size was somewhat small ($N = 300$) relative to the number of items, it was sufficient for subscale analysis.

Assumptions of local independence. To ensure that there was no discernable pattern in the standardized residuals, response dependency and multidimensionality were examined (Tennant and Conaghan, 2007). Response dependency occurs when items are connected to each other in that the response to one item affects the response of another (or others). A

principal components analysis (PCA) of the standardized residuals and examination of the residual correlation matrix for highly correlated items allowed investigation of response dependency (items that are highly correlated are deemed response dependent). Examination of the residual correlation matrix indicated that a number of items were highly correlated, providing evidence of response dependency and the presence of subscales. A subtest analysis further indicated that the reliability of the scale was artificially inflated due to the presence of multiple response-dependent items (as evidenced by a decline in reliability in this analysis; Marais and Andrich, 2008).

The presence of functional subscales (multidimensionality) was further evidenced by: a significant item–trait interaction, $X^2(144) = 428.20$, $p < .001$; a PCA of the residuals generating a dominant factor with an eigenvalue of 3.60 (< 2 suggests unidimensionality; Linacre, 2011) accounting for only 10.1% of the variance (>20% suggests unidimensionality; Reckase, 1979); and *t*-tests (Smith, 2002) indicating that 70 of 300 (23.3%) participants had significantly different locations (5–7% suggests unidimensionality; Hagquist et al., 2009). Taken together, these results provide evidence that the 36-item scale was not unidimensional in nature.

Differential Item Functioning (DIF). To investigate the possibility of item bias, DIF was conducted to determine whether different groups (i.e., males, females) responded differently on items, despite having the same level of the latent trait being measured (Tennant and Conaghan, 2007). In RUMM2030, DIF is determined by ICCs and the relationship between the slopes of the observed means for the class intervals of the comparison group (Andrich et al., 2010). This is evaluated statistically by analyses of variance on the standardized residuals for persons and items. Specifically, a two-way ANOVA is run on the class intervals of the comparison group, yielding a main effect of class interval, a main effect of the comparison group, as well as any interaction between these effects. Significant results

indicate the presence of DIF. This and all subsequent DIF analyses followed this approach. At a Bonferroni-adjusted alpha of .0004, no evidence of DIF was detected.

Summary. Initial Rasch analysis of the 36-item scale revealed that, in its full form, it is problematic. Results clearly indicated the presence of multiple dimensions and reliability of the omnibus scale was poor. Further, many of the questions' response categories were not functioning according to the Rasch model. That is, the threshold estimates failed to increase monotonically (e.g., participants with higher levels of the trait respond in a manner that is reflective of these higher levels). While this problem was resolved by collapsing categories, it should be noted that this method is considered controversial (Andrich, 2012). Finally, nine questions did not function as intended according to the Rasch model. Overall, this indicates that the full scale is not suitable to be used as a total score measure to quantify the severity of students' attitudes toward plagiarism. However, the three subscales identified by EFA were shown to be more reliable as independent measures. Subsequent Rasch analyses on these subscales followed the same method and procedures as above, except where specified. The results are summarised as follows.

Subscale 1 (*Factors that Exacerbate Plagiarism*)

This first 14-item subscale (*Factors that Exacerbate Plagiarism*) displayed improved reliability relative to the full scale, with clear evidence of unidimensionality. However, there remained problems arising from misfitting items (see Table 3). Items considered for deletion (to obtain a subscale with improved psychometric properties) were defined as items with fit residuals more extreme than ± 2.5, significant chi-square and *F* statistics (*p*s < .001 after Bonferroni adjustment), and/or items with problematic ICCs. This procedure entailed removing misfitting items and re-running the Rasch analyses to examine fluctuations in PSIs and item–trait interactions. An item meeting these criteria was ultimately retained if the PSI was reduced and the item–trait interaction became 'more significant' after the item was

removed (an item was deleted if the reverse). This procedure was repeated until: (1) no additional misfitting items were apparent; and (2) improved psychometric properties were attained.

Results of these analyses revealed a reliable eight-item subscale (see Table 3). The improved psychometric properties of this subscale was evidenced by: a non-significant item–trait interaction, $X^2(32) = 40.90$, $p = .140$, indicating that overall fit was good; a slightly low PSI (0.67) and Cronbach's alpha coefficient (0.72), suggesting marginally acceptable reliability; individual item fit ($M = 0.29$, $SD = 1.07$) and person fit residuals ($M = –0.36$, $SD = 0.73$), indicating good overall fit of the data to the model; and no individual item misfit evident upon examination of ICCs. This eight-item subscale further demonstrated similar targeting as the full scale (see Figure 4), with most participants having harsher attitudes toward plagiarism, with item locations again providing evidence of a good range of items. This subscale also met assumptions of local independence, according to the expectations of the Rasch model (PCA of standardized residuals indicated no response dependency, with all correlations < .17). Unidimensionality assumptions were also met (PCA analysis indicated a dominant factor accounting for 20.7% of the variance, with an eigenvalue of 1.65). A *t*-test of person locations (Smith, 2002), using two subsets of positively and negatively loaded items, indicated that 9 out of 300 (3%) participants had significantly different locations at the 5% level. Taken together, these results indicate that the eight-item subscale was unidimensional. Finally a DIF analysis of males and females was run on all eight items with Bonferroni corrections (alpha = .002). No evidence of DIF for sex was found.

Subscale 2 (*Justification for Plagiarism*)

The second six-item subscale (*Justification for Plagiarism*) displayed similarly improved reliability compared to the full scale. However, there remained issues surrounding the fit and reliability of this subscale. Rasch analysis revealed a significant item–trait interaction, $X^2(24)$

= 37.00, p = .040, indicating that overall fit was slightly below acceptable levels. The PSI (0.59) and Cronbach's alpha coefficient (0.64) was substantially lower than what is considered acceptable, indicating poor reliability. However, overall individual item fit (M = 0.08, SD = 1.04) and person fit residuals (M = –0.26, SD = 0.88) indicated that the data fit the model well. Examination of the ICCs for individual items indicated that all were functioning as expected (hence no items were deleted; see Table 3).

The targeting of the scale was consistent with the prior analysis, indicating a substantially greater number of participants with harsher attitudes toward plagiarism and, again, the spread of items covered a good range (see Figure 4). The subscale also met assumptions of local independence (PCA of residuals revealed no response dependency; all correlations < .07). Unidimensionality assumptions were also met, with the first dominant factor accounting for 29.1% of the variance and an eigenvalue of 1.75. A *t*-test of the person locations (Smith, 2002) also provided evidence of unidimensionality (only 4.3% of participants had significantly different locations at the 5% level). No evidence of DIF on the latent trait variable 'sex' (alpha = .003) was detected. Despite generally good psychometric properties, the poor reliability (perhaps a function of the limited number of items in the subscale; John and Benet-Martinez, 2000) suggests that this subscale should be used with caution.

Subscale 3 (*Severity and Penalty*)

This nine-item subscale (*Severity and Penalty*) displayed good psychometric properties after deletion of a single misfitting item (based on fluctuations in PSI and item–trait interactions; see Table 3). Rasch analysis of the resulting eight-item subscale revealed a non-significant item–trait interaction, $X^2(36) = 59.70$, p = .060, indicating acceptable overall fit of the data to the model. Although the PSI was slightly low (0.67), the Cronbach's alpha coefficient (0.70) indicated an acceptable level of reliability. Overall individual item fit (M =

0.28, *SD* = 1.00) indicated good fit to the model, as did overall person fit residuals (*M* = –0.30, *SD* = 1.10). Item misfit was detected in the fit residuals for a single item, X^2 = 18.00, *p* < .002 and *F* = 5.83, *p* < .001, however, this item was ultimately retained because its removal reduced the PSI and Cronbach's alpha of the scale.

The targeting of this subscale was again consistent with prior analyses indicating a greater number of persons with harsher attitudes toward plagiarism, while the spread of items was even across levels of severity of attitudes toward plagiarism (see Figure 4). Response dependency and unidimensionality assumptions were also met. PCA analysis revealed no highly correlated item loadings, with a first dominant factor accounting for 21.7% of the variance and an eigenvalue of 1.70. A *t*-test of the person locations comprising positively and negatively loaded items indicated that 22 out of 300 (7.3%) of persons had significantly different locations at the 5% level. All results suggest a unidimensional subscale. There was no evidence of DIF on 'sex' after Bonferroni corrections (alpha = .003).

Descriptive Statistics of the Revised Subscales

Descriptive statistics for all revised subscales are provided in Table 4. Examination of this data for the psychometrically sound subscales (subscales 1 and 3) indicated that, although respondents tended to take a strong stance against plagiarism, these attitudes were not unanimous. For instance, while the majority of students believed that 'plagiarism is as bad as stealing an exam' (51.7%) and 'plagiarism undermines independent thought' (62.0%), there were a number of students that disagreed with these sentiments (21.3% and 6.7%, respectively). Further, there was a minority of respondents who indicated that plagiarism 'is not a big deal' (7.3%), 'should not be considered a serious offence' (9.3%), and is sometimes 'necessary' (3.7%). The situation most strongly perceived as tempting students to plagiarize was when peers make their previous works available (7.7%), followed by external time pressures (5.7%) and the perception that punishment for plagiarism would be light (1.7%).

There was also a relatively large contingent that believed the majority of students have engaged in plagiarist behaviours (22.7%) and that this 'does no harm to the value of a university degree' (7.3%). Although only a small minority expressed lenient attitudes toward plagiarism, the majority of respondents indicated that they did not view self-plagiarism in the same light. That is, 64.0% of respondents indicated that 'self-plagiarism should not be punishable in the same way as plagiarism is'.

Discussion

The current study involved modification of an existing plagiarism attitudes scale (to broaden its scope and applicability) and evaluation of its psychometric properties (using both traditional and modern psychometric approaches), before evaluating results of administration of the scale. Supplementing traditional approaches to scale evaluation (i.e., EFA, Cronbach's alpha) with modern psychometric approaches provided information on measurement issues that are not fully addressed by traditional psychometric approaches. Specifically, Rasch analysis provided additional evidence of unidimensionality (required in order to generate sum scores from items in a subscale), appropriateness of response orderings (to determine whether the ordering of response categories was appropriate relative to patterns of responding), measurement invariance (to evaluate whether items remained constant across participants with different levels of the latent trait), and differential item functioning (to examine whether items were biased across different sample subgroups; Tennant and Conaghan, 2007). Another important advantage of Rasch analysis was its utility for evaluating the scale on an item-by-item basis, rather than merely evaluating the functioning of the overall (sub)scale. Despite growing recognition of the advantages of these modern evaluation methods, few studies have adopted this approach to survey-based investigations of plagiarism and none appear to have done so for investigating students' attitudes toward plagiarism. As a result, it can be argued

that there remains no standardized measure of, or reliable evidence about, Australian university students' attitudes toward plagiarism (e.g., Gururujan and Roberts, 2005).

Given that traditional test theory approaches aim to *summarize* the existing data, whereas modern approaches seek to *construct* measures, vast differences in conclusions can arise from these approaches (Tatum, 2000; Wright, 1996). It was therefore expected that our analyses would differ in their advocacy for the inclusion and exclusion of individual items. In line with these expectations, the results of our Rasch analysis provided a stark contrast to those of our traditional test theory analyses. Whereas traditional test theory approaches suggested the presence of three subscales with good reliability, subsequent Rasch analyses highlighted problems with each of the subscales. Specifically, the 29 items that clustered together into three subscales during EFA did not all index their underlying constructs according to Rasch analyses. Instead, Rasch analyses indicated that the survey instrument functioned best (i.e., provided optimal validity and reliability metrics) as two independent eight-item subscales that capture students' attitudes toward plagiarism and perceived precursors of plagiarist behaviors.

These results contrast Mavrinac et al.'s (2010) psychometric evaluation of their scale, which revealed a well-functioning 29-item scale with three psychometrically valid (construct validity determined by EFA) and reliable (Cronbach's alpha >.70) subscales. The results of our true score theory analyses paralleled those of Mavrinac et al. (2010), insofar as both advocated a 29-item scale comprised of three subscales (albeit with a somewhat different composition of items), each with similarly good reliability. However, subsequent Rasch analyses indicated that the second subscale, as currently constituted, could not be interpreted as reliably measuring its underlying construct (justification for plagiarism). Although this six-item subscale (*Justification for Plagiarism*) was found to have low reliability, the inclusion of additional items measuring this latent construct may serve to improve its reliability. As such,

we advocate modification of the scale rather than its complete removal, because justification for plagiarism is an important factor in understanding why students commit acts of plagiarism (Bamford and Sergiou, 2005; Bannister and Ashworth, 1998).

In addition, a further 10 items from Mavrinac et al.'s (2010) original scale were ultimately removed from our final scale due to low EFA factor loadings in the current study (five items) or misfit identified by Rasch analyses (five items). Although this discord may have been influenced in part by the modification of a number of items from the original scale, many unaltered items also displayed poor fit to the Rasch model. This is consistent with the fact that the result of any psychometric evaluation of a scale is a function of the items (evidenced by misfitting items), participants' characteristics (evidenced by differences in EFA results in the current and previous studies), and the aims of the psychometric analyses (summarizing data versus measurement construction; Alwin, 1991). As such, this divergence in results is likely the result of both context (e.g., different countries, different faculties, new and modified items) and analyses (e.g., Rasch analyses identifying misfitting items that were deemed appropriate by traditional analyses).

Rasch analyses thus provide an important advantage to survey-based research, because significant problems arise when interpreting questionnaire results in the absence of sufficient validity and reliability evidence. As stated by Alwin (1991), "the quality of the conclusions drawn from the survey depends heavily on the quality of measurement" (p. 5). For instance, students who believe that plagiarism only occurs when substantial amounts of text are taken verbatim without acknowledgement may agree that 'plagiarists do not belong at university' (Item 13, which did not load significantly on any of the subscales). Others who recognize that plagiarism may involve more minor infractions, by those who genuinely do not understand academic writing conventions, may take the opposite view. In such cases, it is unclear what position students are actually taking in their responses. Similarly, students with different

conceptions of plagiarism may have different opinions of whether ‘everyone else is doing it’ (Item 22, which loaded significantly on subscale 1 but misfit the Rasch model). According to Eley (1994), such questions pose a problem for the validity of a scale, in that “it is only reasonable to include questions to which students are qualified to respond” (p. 6). In the current study, despite previous psychometric evidence for the construct validity and reliability of the original 29-item scale (Mavrinac et al., 2010), 11 of those items (five from subscale 1 and all of subscale 2) appear problematic for subsequent interpretation. Further, only three of the seven items added to the scale (items 30 to 36) functioned well according to the Rasch model. As such, attempts to implement educational strategies on the basis of those invalid or unreliable survey items is, to say the least, problematic.

Examination of the data derived from well-functioning items, as determined by Rasch analyses, indicated generally strong attitudes against plagiarism. Nevertheless, there was a concerning minority that felt plagiarism was ‘not a big deal’, was sometimes necessary, and, at least in some cases, should be ignored. The strongest temptation appeared to stem from peers sharing their previous assignments, followed by external time pressures and perceptions of light punishments for plagiarism. That these potential precursors were cited as temptations for many who did not see plagiarism as a necessity suggests that convenience may also be an important consideration in plagiarist behaviors (even if it is neither ‘necessary’ nor ‘right’). Although only a minority of respondents held these concerning attitudes, the majority of students did not hold self-plagiarism in the same regard. In fact, responses suggested that self-plagiarism has not received the same notoriety as other more-traditional forms of plagiarism. These results indicated that, although the majority of students acknowledge the seriousness of plagiarism, these attitudes were neither unanimous nor consistent across the range of issues assessed.

Despite extensive evidence of the good psychometric properties of this revised scale, these descriptive results need to be interpreted in the context of the limits of survey research. That is, survey research is restricted by: (i) the ease with which participants can determine the most 'socially acceptable' response (Schwartz, 1999); (ii) the attempt to assess complex issues using often oversimplified questions (Kalton and Schuman, 1982); and (iii) the potential disconnect between respondents' perceptions and actions (Schuman and Johnson, 1976). Although none of these issues can conclusively be remedied in the conduct of survey research, measures were put in place to mitigate their impact. For instance, to maximize accuracy of responding, participants' responses remained anonymous throughout the conduct of this study. This was particularly important for items assessing whether students had engaged in plagiarism. Further, to ensure appropriate questioning, items from the original scale (Mavrinac et al., 2010) were reviewed and, if necessary modified, to ensure their applicability to the entire range of university students. Despite these measures, further research is required to determine the extent to which students' stated perceptions on this scale correspond with their actions (e.g., whether instances of plagiarism increase during times of compressed deadlines).

In addition, the restricted sample of male respondents may have prevented detection of differential item functioning for males and females. Although the large contingent of female respondents reflects the broader demographic makeup of the faculty, DIF analyses would, nevertheless, have been better served with a larger proportion of male respondents. Because a rigorous psychometric approach should always include DIF analysis on these latent trait variables (Tennant and Conaghan, 2007), it is recommended that future administrations of this instrument attain a more gender-balanced sample to enable a more meaningful DIF analysis.

The current study provides a number of important advances to the field of plagiarism research. First, whereas traditional psychometric analyses remain the norm in survey research (Alwin, 1991), supplementing these methods with Rasch analyses allowed us to construct a measure that permits confident quantification and measurement of its underlying constructs. This Rasch-derived 16-item scale differed substantially from that derived by traditional test theory analyses. In highlighting the importance of these methods, Wright (1996) states "only Rasch analysis constructs the kind of objective linear variables that social scientists need to quantify their constructs, map their fields of study, test their hypotheses, and measure the values of their social programs" (p. 24). Second, this study provides extensive evidence of the validity and reliability of this plagiarism measure. That this instrument demonstrated good fit to the Rasch model suggests its utility for measurement in a valid and reliable manner across a range of samples (Hagquist et al., 2009). Lastly, the current study provides initial baseline data from administration of the scale. Although often omitted from validation papers, baseline data is important to facilitate future discussions of differences as a function of contextual (e.g., cultural), individual (e.g., intrapersonal factors), and external (e.g., academic pressures) factors. This study represents a first step in facilitating these discussions, and determining appropriate actions, on the basis of measures that have undergone rigorous psychometric testing.

References

Alwin, D. F. (1991). Research on survey quality. *Sociological Methods & Research, 20*(1), 3–29.

Andrich, D. (1978). Rating formulation for ordered response categories. *Psychometrika, 43,* 561–573.

Andrich, D. (1988). *Rasch models for measurement*. Beverly Hills: SAGE Publications.

Andrich, D. (2004). Controversy and the Rasch model: A characteristic of incompatible paradigms? *Medical Care, 42*, 7–16.

Andrich, D. (2013). An expanded derivation of the threshold structure of the polytomous Rasch model that dispels any 'threshold disorder controversy'. *Educational and Psychological Measurement, 73,* 78–124.

Andrich, D., Sheridan, B., and Luo, G. (2010). RUMM2030: A Windows program for the Rasch Unidimensional Measurement Model [Computer software]. Perth: RUMM Laboratory.

Ashworth, P., Bannister, P., and Thorne, P. (1997). Guilty in whose eyes? University students' perceptions of cheating and plagiarism in academic work and assessment. *Studies in Higher Education, 22,* 187–203.

Bamford, J., and Sergiou, K. (2005). International students and plagiarism: An analysis of the reasons for plagiarism among international foundation students. *Investigations in University Teaching and Learning, 2,* 16–22.

Bannister, P., and Ashworth, P. (1998). Four good reasons for cheating and plagiarism. In C. Rust (Ed.), *Improving student learning symposium* (pp. 233–241). Oxford: Oxford Centre for Staff Development, Oxford Brookes University.

Baruchson-Arbib, S., and Yaari, E. (2004). Printed versus Internet plagiarism: A study of students' perception. *International Journal of Information Ethics, 1*, 1–7.

Bowers, W. J. (1964). *Student dishonesty and its control in college*. New York: Bureau of Applied Social Research.

Breen, L., and Maassen, M. (2005). Reducing the incidence of plagiarism in an undergraduate course: The role of education. *Issues in Educational Research, 15,* 1–16.

Bretag, T., and Mahmud, S. (2009). Self-plagiarism or appropriate textual re-use? *Journal of Academic Ethics, 7,* 193–205.

Brimble, M., and Stevenson-Clarke, P. (2005). Perceptions of the prevalence and seriousness of academic dishonesty in Australian universities. *Australian Educational Researcher, 32,* 19–44.

Crisp, G. T. (2007). Staff attitudes to dealing with plagiarism issues: Perspectives from one Australian university. *International Journal of Educational Integrity, 3,* 3–15.

Deckert, G. D. (1993). Perspectives on plagiarism from ESL students in Hong Kong. *Journal of Second Language Writing, 2,* 131–148.

Devlin, M., and Gray, K. (2007). In their own words: A qualitative study of the reasons Australian university students plagiarize. *Higher Education Research and Development, 26,* 181–198.

Ehrich, J. F., Howard, S. J., Tognolini, J., and Bokosmaty, S. (2013). Measuring attitudes toward plagiarism: Issues and psychometric solutions. Manuscript submitted for publication.

Eley, M. G. (1994). Using students' ratings in the evaluation of teaching: Monash University's MonQueST project (1993 John Smyth Memorial Lecture). *VIER Bulletin, 72,* 3–28.

Ellery, K. (2008). Undergraduate plagiarism: A pedagogical perspective. *Assessment & Evaluation in Higher Education, 33,* 507–516.

Field, A. (2009). *Discovering statistics using SPSS* (3rd Ed.). London: SAGE Publications.

Gibaldi, J. (2003). *MLA handbook for writers of research papers*. New York: Modern Language Society of America.

Gullifer, J., and Tyson, G. A. (2010). Exploring university students' perceptions of plagiarism: A focus group study. *Studies in Higher Education, 35,* 463–481.

Gururajan, R., and Roberts, D. (2005, July). *Attitude towards plagiarism in information systems in Australian universities.* Paper presented at the Asia Pacific Conference on Information Systems, Bangkok, Thailand.

Hagquist, C., Bruce, M., and Gustavsson, P. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies, 46,* 380–393.

Handa, N., and Power, C. (2005). Land and discover! A case study investigating the cultural context of plagiarism. *Journal of University Teaching and Learning Practice, 2,* 64–84.

Harris, R. A. (1995). Overseas students in the United Kingdom university system. *Higher Education, 29,* 77–92.

Harris, R. A. (2001). *The plagiarism handbook: Strategies for preventing, detecting, and dealing with plagiarism*. Los Angeles: Pyrczak.

Hughes, J. M. C., and McCabe, D. L. (2006). Understanding academic misconduct. *Canadian Journal of Higher Education, 36,* 49–63.

John, O. P., and Benet-Martinez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis and C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339–369). New York: Cambridge University Press.

Kalton, G., and Schuman, H. (1982). The effect of the question on survey responses: A review. *Journal of the Royal Statistical Society, 145,* 42–73.

Linacre, J. M. (2011). *A users guide to Winsteps/Ministep Rasch-model computer programs*. Retrieved from http://www.winsteps.com/winsteps.htm

Marais, I., and Andrich, D. (2008). Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *Journal of Applied Measurement, 9*, 200–215.

Marshall, S., and Garry, M. (2006). NESB and ESB students' attitudes and perceptions and plagiarism. *International Journal for Educational Integrity, 2*, 26–37. Retrieved from http://www.ojs.unisa.edu.au/index.php/IJEI/issue/view/4

Martin, D. E., Rao, A., and Sloan, L. R. (2011). Ethnicity, acculturation, and plagiarism: A criterion study of unethical academic conduct. *Human Organization, 70*, 88–96.

Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149–174.

Mavrinac, M., Brumini, G., Bilic-Zulle, L., and Petrovecki, M. (2010). Construction and validation of attitudes toward plagiarism questionnaire. *Croatian Medical Journal, 51*, 195–201.

Maxwell, A., Curtis, G. J., and Vardanega, L. (2008). Does culture influence understanding and perceived seriousness of plagiarism? *International Journal for Educational Integrity, 4*(2), 25–40. Retrieved from http://ojs.ml.unisa.edu.au/index.php/IJEI/issue/view/75

McCabe, D. L. (2005). Cheating among college and university students: A North American perspective. *International Journal for Educational Integrity, 1*(1). Retrieved from http://ojs.ml.unisa.edu.au/index.php/IJEI/issue/view/3

McCabe, D. L., and Bowers, W. J. (1994) Academic dishonesty among males in college: A thirty year perspective. *Journal of College Student Development, 35*, 5–10.

McCabe, D. L., and Trevino, L. K. (1993). Academic dishonesty: Honor codes and other contextual influences. *Journal of Higher Education, 64*, 522–538.

McCabe, D. L., and Trevino, L. K. (1995). Cheating among business students: A challenge for business leaders and educators. *Journal of Management Education, 19,* 205–218.

McCabe, D. L., and Trevino, L. K. (1996). What we know about cheating in college: Longitudinal trends and recent developments. *Change*, *28*(1), 28–33.

McCabe, D. L., Trevino, L. K., and Butterfield, K. D. (2001). Cheating in academic institutions: A decade of research. *Ethics and Behavior, 11,* 219–232.

McCabe, D. L., Trevino, L. K., and Butterfield, K. D. (2004). Academic integrity: How widespread are cheating and plagiarism? In D. R. Karp and T. Allena (Eds.), *Restorative justice on the college campus: Promoting student growth, and responsibility, and reawakening the spirits of campus community* (pp. 130–141). Springfield, IL: Charles D. Thomas.

McGowan, U. (2005). Does educational integrity mean teaching students NOT to use their own words? *International Journal for Educational Integrity*, *1*(1). Retrieved from http://www.ojs.unisa.edu.au/index.php/IJEI/issue/view/3

Park, C. (2003). In other (people's) words: Plagiarism by university students – literature and lessons. *Assessment & Evaluation in Higher Education, 28,* 471–488.

Perry, A. R., Kane, K. M., Bernesser, K. J., and Spicker, P. T. (1990). Type A behavior, competitive achievement-striving, and cheating among college students. *Psychological Reports, 66,* 459–465.

Reckase, M. D. (1979). Unifactor latent trait models applied to multi-factor tests: Results and implications. *Journal of Educational Statistics, 4,* 207–230.

Rocco, E., and Warglien, M. (1995). *Computer mediated communication and the emergence of electronic opportunism*. Provo, UT: Association for Consumer Research. Retrieved from http://eprints.biblio.unitn.it/archive/00000034/01/CEEL96_01.pdf

Ryan, G., Bonanno, H., Krass, I., Scouller, K., and Smith, L. (2009). Undergraduate and postgraduate pharmacy students' perceptions of plagiarism and academic honesty. *American Journal of Pharmaceutical Education*, *73*(6), 1–8.

Scanlon, P. M. (2003). Student online plagiarism: How do we respond? *College Teaching*, *51*, 161–165.

Schuman, H., and Johnson, M. P. (1976). Attitudes and behaviors. *Annual Review of Sociology*, *2*, 161–207.

Schwartz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93–105.

Selwyn, N. (2008). 'Not necessarily a bad thing'. A study of online plagiarism amongst undergraduate students. *Assessment & Evaluation in Higher Education*, *33*, 465–479.

Smith, Jr., E. V. (2002). Understanding Rasch measurement: Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, *3*, 205–231.

Song-Turner, H. (2008). Plagiarism: Academic dishonesty or 'blind spot' of multicultural education? *Australian University Review*, *50*(2), 39–51.

Sutherland-Smith, W. (2008). *Plagiarism, the Internet and student learning: Improving academic integrity.* Abingdon, UK: Routledge.

Tatum, D. S. (2000). Rasch analysis: An introduction to objective measurement. *Laboratory Medicine*, *31*(5), 272–274.

Tennant, A., and 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 paper? *Arthritis Care & Research*, *57*, 1358–1362.

Underwood, J., and Szabo, A. (2003). Academic offences and e-learning: Individual propensities in cheating. *British Journal of Educational Technology*, *34*, 467–477.

Walker, J. (1998). Student plagiarism in universities: What are we doing about it? *Higher Education Research & Development*, *17*, 89–106.

Wright, B. D. (1996). Comparing Rasch measurement and factor analysis. *Structural Equation Modeling*, *3*, 3–24.

Zimitat, C. (2008). A student perspective of plagiarism. In T. S. Roberts (Ed.), *Student plagiarism in an online world: Problems and solutions* (pp. 10–22). Hershey, PA: Information Science Reference.

Table 1a

EFA Factor Loadings for Subscale 1 (Factors that Exacerbate Plagiarism)

#	Item	*Factor Loading*
9	Short deadlines or a heavy workload give me the right to plagiarise a bit.	.44
12	If another student gives me permission to copy from his/her paper, I'm not doing anything bad because I have his/her permission.	.38
18	A plagiarised paper does no harm to the value of a university degree.	.35
21	Those who say they have never plagiarised are lying.	.38
22	Sometimes I'm tempted to plagiarise, because everyone else is doing it (e.g., students, researchers, academic staff).	.75
23	I keep plagiarising because I haven't been caught yet.	.61
26	Sometimes I copy a sentence or two just to become inspired for further writing.	.60
28	Plagiarism can be justified if I currently have more important obligations or tasks to do.	.63
29	Sometimes, it is necessary to plagiarise.	.59
30	I am tempted to plagiarise if I have permission from a friend to copy his or her work.	.73
31	Plagiarism is against my ethical values.	.37
33	I am tempted to plagiarise if I currently have more important obligations or tasks to do.	.81
34	Plagiarism is only a big deal if a substantial portion of the paper has been plagiarised.	.34
36	I am tempted to plagiarise because, even if caught, the punishment will be light (the reward outweighs the risk).	.54

Note. **Bolded** item numbers identify items that were reverse scored for EFA and Rasch analyses. Only factor loadings > .30 (statistically significant) are presented here.

Table 1b

EFA Factor Loadings for Subscale 2 (Justification for Plagiarism)

#	Item	*Factor Loadings*
1	Sometimes you cannot avoid using other people's words, because there are only so many ways to describe something.	.48
2	It is justified to use previous descriptions of a concept or theory, because they remain the same.	.40
3	Self-plagiarism is not punishable because it is not harmful (you cannot steal from yourself).	.75
6	Undergraduate students, because they are just learning the ropes, should receive milder punishment for plagiarism.	.43
11	It is justified to use your own previous work, without providing citation, in order to complete the current work.	.67
27	Since plagiarism is taking other people's words rather than tangible assets, it should not be considered a serious offence.	.38

Note. **Bolded** item numbers identify items that were reverse scored for EFA and Rasch analyses. Only factor loadings > .30 (statistically significant) are presented here.

Table 1c

EFA Factor Loadings for Subscale 3 (Severity and Penalty)

#	Item		***Factor Loadings***
4	Plagiarised parts of a student's paper should be ignored if the paper is otherwise of high quality.		.50
5	Self-plagiarism should not be punishable in the same way as plagiarism is.		.65
7	If you cannot write well because of unfamiliarity with the topic area, it is justified to copy parts of a paper already published in that area in order to accurately represent those ideas.		.32
8	I could not write a good academic paper without plagiarising.		.48
15	Given a commonly perceived decline in moral and ethical standards, it is important to discuss issues like plagiarism and self-plagiarism.		.61
16	Plagiarism is as bad as stealing an exam.		.55
17	Plagiarism undermines independent thought.		.67
19	Since plagiarism is taking other people's words rather than tangible assets, it should not be considered a serious offence.		.34
25	Plagiarism is not a big deal.	.32	.44

Note. **Bolded** item numbers identify items that were reverse scored for EFA and Rasch analyses. Only factor loadings > .30 (statistically significant) are presented here.

Table 2

Fit of the 36 Items of the Revised Plagiarism Attitude Scale

Item No.	Location	*SE*	Fit Resid	χ^2	*p*	*F*	*p*
1	–0.295	0.081	1.097	1.837	.766	0.428	.788
2	–1.351	0.067	0.812	4.158	.385	1.006	.405
3	0.396	0.075	0.227	3.032	.552	0.652	.626
4	–1.429	0.067	0.807	7.473	.113	1.920	.107
5	–1.120	0.060	2.024	8.303	.081	1.986	.097
6	0.427	0.086	0.473	2.720	.606	0.631	.641
7	0.663	0.079	–0.714	6.514	.164	1.820	.125
8	1.014	0.091	**–2.845**	18.852	**<.001**	6.986	**<.001**
9	1.406	0.199	–0.294	1.973	.741	0.579	.678
10	–0.586	0.069	0.054	2.450	.654	0.601	.662
11	0.817	0.085	–1.287	11.020	.026	3.135	.015
12	–0.737	0.056	**4.594**	24.307	**<.001**	5.085	<**.001**
13	–1.711	0.058	**2.656**	28.738	**<.001**	6.448	**<.001**
14	0.074	0.071	1.378	4.872	.301	0.996	.410
15	–0.477	0.058	0.281	4.816	.307	1.171	.324
16	0.011	0.074	–1.016	14.848	.005	4.180	.003
17	0.181	0.073	–0.236	7.068	.132	1.715	.146
18	–0.179	0.069	0.815	11.393	.022	2.486	.044
19	–0.180	0.090	1.608	10.777	.029	2.403	.050
20	–0.967	0.091	0.370	1.451	.835	0.323	.862
21	0.779	0.095	–0.707	9.535	.049	2.562	.039
22	2.820	0.360	–1.775	8.500	.075	3.372	.010
23	–1.130	0.066	**3.656**	25.923	**<.001**	5.334	**<.001**
24	–0.271	0.072	**–2.667**	20.116	**<.001**	6.888	**<.001**
25	–0.209	0.062	–0.687	5.176	.270	1.358	.248
26	–1.008	0.063	2.326	14.186	.007	2.973	.020
27	2.010	0.087	–1.893	19.965	**.001**	6.198	**<.001**
28	0.703	0.075	**–3.262**	31.062	**<.001**	12.005	**<.001**
29	0.658	0.095	–0.355	6.775	.148	1.749	.139
30	–0.656	0.066	–0.566	2.268	.687	0.826	.509
31	–1.133	0.059	**5.656**	52.108	**<.001**	10.310	**<.001**
32	1.036	0.105	–0.865	16.396	.003	4.719	.001
33	–0.254	0.060	0.159	3.118	.538	0.897	.466
34	–0.737	0.088	**3.125**	21.380	**<.001**	4.694	**<.001**
35	0.527	0.083	–0.431	5.652	.227	1.497	.203
36	0.908	0.168	–1.702	9.406	.052	2.904	.022

Note. Items with fit residuals (Fit Resid) <-2.5 and > 2.5 (in **bold**) are considered misfitting. *p* = Bonferroni adjusted *p* values (**bolded** are significant at $p < .001$; .05/36).

Table 3

Item Fit by Subscale (after EFA and Rasch Deletions)

Item Number	Location	*SE*	Fit Resid	χ^2	*p*	*F*	*p*
Subscale 1: Factors that Exacerbate Plagiarism							
9	2.032	0.364	0.421	3.314	.507	0.826	.509
18	–1.141	0.077	0.984	5.496	.240	1.406	.232
21	–0.077	0.100	–0.688	5.279	.260	1.007	.404
28	–0.196	0.081	–2.162	9.989	.041	3.111	.016
29	–0.118	0.098	–1.437	6.041	.196	1.490	.205
30	–1.673	0.075	–0.019	1.758	.780	0.537	.709
33	–1.207	0.067	0.055	2.351	.671	0.666	.616
36	2.380	0.426	0.515	6.641	.156	1.005	.405
Subscale 2: Justification for Plagiarism							
1	–0.663	0.085	1.361	2.429	.657	0.613	.654
2	–1.779	0.072	1.203	4.973	.290	1.073	.370
3	0.081	0.077	–0.412	7.826	.098	1.496	.203
6	0.136	0.089	–0.572	8.503	.075	1.750	.139
11	0.548	0.084	–1.281	7.857	.097	1.979	.098
27	1.677	0.089	0.237	5.396	.249	0.991	.413
Subscale 3: Severity and Penalty							
4	–1.093	0.067	0.782	3.057	.548	0.804	.523
5	–0.770	0.060	0.833	2.234	.693	0.608	.657
7	0.944	0.077	–1.090	3.687	.450	0.902	.463
15	–0.145	0.058	–0.496	2.182	.702	0.700	.592
16	0.337	0.074	–0.934	17.995	**.001**	5.830	**<.001**
17	0.458	0.072	0.481	4.215	.378	0.987	.415
19	0.155	0.089	1.730	11.801	.019	2.776	.027
25	0.114	0.061	0.924	0.444	.979	0.104	.981

Note. Items with fit residuals (Fit Resid) <-2.5 and > 2.5 (in **bold**) are considered misfitting. *p* = Bonferroni adjusted *p* values (**bolded** are significant).

Table 4

Descriptive Statistics by Subscale

	Subscale 1: Factors that Exacerbate Plagiarism (8 items)	***M***	***SD***	***%Agree***
9	Short deadlines or a heavy workload give me the right to plagiarise a bit.	1.45	0.73	2.3
18	A plagiarised paper does no harm to the value of a university degree.	2.04	0.92	7.3
21	Those who say they have never plagiarised are lying.	2.95	1.07	22.7
28	Plagiarism can be justified if I currently have more important obligations or tasks to do.	1.63	0.74	1.0
29	Sometimes, it is necessary to plagiarise.	1.78	0.90	3.6
30	I am tempted to plagiarise if I have permission from a friend to copy his or her work.	1.85	1.04	7.7
33	I am tempted to plagiarise if I currently have more important obligations or tasks to do.	1.85	0.98	5.7
36	I am tempted to plagiarise because, even if caught, the punishment will be light (the reward outweighs the risk).	1.56	0.81	1.7
	Subscale 2: Justification for Plagiarism (6 items)	***M***	***SD***	***%Agree***
1	Sometimes you cannot avoid using other people's words, because there are only so many ways to describe something.	3.20	1.14	44.0
2	It is justified to use previous descriptions of a concept or theory, because they remain the same.	2.99	0.95	27.3
3	Self-plagiarism is not punishable because it is not harmful (you cannot steal from yourself).	3.70	1.09	58.7
6	Undergraduate students, because they are just learning the ropes, should receive milder punishment for plagiarism.	3.14	1.12	36.0
11	It is justified to use your own previous work, without providing citation, in order to complete the current work.	3.02	1.20	34.4
27	Since plagiarism is taking other people's words rather than tangible assets, it should not be considered a serious offence.	3.41	1.23	49.0
	Subscale 3: Severity and Penalty (8 items)	***M***	***SD***	***%Agree***
4	Plagiarised parts of a student's paper should be ignored if the paper is otherwise of high quality.	1.86	0.88	4.4
5	Self-plagiarism should not be punishable in the same way as plagiarism is.	3.79	1.10	64.0
7	If you cannot write well because of unfamiliarity with the topic area, it is justified to copy parts of a paper already published in that area in order to accurately represent those ideas.	1.59	0.74	1.6
15	Given a commonly perceived decline in moral and ethical standards, it is important to discuss issues like plagiarism and self-plagiarism.	3.92	0.91	69.0
16	Plagiarism is as bad as stealing an exam.	3.52	1.22	51.7
17	Plagiarism undermines independent thought.	3.77	0.90	62.0
19	Since plagiarism is taking other people's words rather than tangible assets, it should not be considered a serious offence.	2.27	0.97	9.3
25	Plagiarism is not a big deal.	2.01	1.00	7.3
	Misfitting Items	***M***	***SD***	***%Agree***
8	I could not write a good academic paper without plagiarising.	1.55	0.84	4.0
10	When I do not know what to write, I borrow from a paper published in an unrelated field.	1.44	0.76	2.4
12	If another student gives me permission to copy from his/her paper, I'm not doing anything bad because I have his/her permission.	1.52	0.77	2.6
13	Plagiarists do not belong at university.	3.32	1.23	46.3
14	The names of students who plagiarise should be disclosed to all academic staff in the faculty.	2.13	1.13	12.3
20	Authors say they do not plagiarise, when in fact they do.	2.87	0.87	15.7
22	Sometimes I'm tempted to plagiarise, because everyone else is doing it (e.g., students, researchers, academic staff).	2.12	1.03	8.0
23	I keep plagiarising because I haven't been caught yet.	1.52	0.82	3.0
24	I study in a plagiarism-free environment.	3.24	1.14	39.3
26	Sometimes I copy a sentence or two just to become inspired for further	2.41	1.13	18.7

	writing.			
31	Plagiarism is against my ethical values.	3.73	1.15	60.6
32	If I lend a paper to another student to look at and then that student turns it in as his/her own and is caught, I should not be punished also.	3.58	1.31	57.3
34	Plagiarism is only a big deal if a substantial portion of the paper has been plagiarised.	2.28	1.17	17.0
35	Uploading a paper anonymously to the Internet facilitates plagiarism and is therefore wrong.	3.20	1.09	31.6

Note. %Agree = proportion of respondents that indicated agreement or strong agreement with the statement. Misfitting items are those that failed to load well on any of the subscales (determined by EFA) or provided cause for deletion during Rasch analyses. They are included here to illustrate that while interpretation of these results is easy, it is problematic to assume that they measure the underlying construct of interest. In addition, all items in subscale 2 should be interpreted with caution due to poor reliability.

Figure 1. Category Characteristic Curve for Question 1. At no point in the graph does the curve for score = 1 appear more likely to be chosen over the curve for score = 0, thus showing disorder of response.

Figure 2. Threshold map showing ordered thresholds of all 36 items of the Revised Attitudes toward Plagiarism scale.

Figure 3. ICCs for Question 8 (top) and Question 24 (bottom). In these graphs the observed values (represented by dots) are slightly steeper than expected values (the s-shaped curve), indicating slight over-discrimination.

Figure 4. Person–item location distribution for: (a) the full 36-item scale; (b) the eight-item subscale '*Issues that may Exacerbate Plagiarism*'; (c) the six-item subscale '*Justification for Plagiarism*'; and (d) the eight-item subscale '*Severity and Penalty*'. In all cases, the top graph represents the spread of persons and the bottom graph shows the spread of items.