Journal of Business Ethics (2007) 72:27–39
DOI 10.1007/s10551-006-9153-y

Psychometric Properties of the Reidenbach–Robin Multidimensional Ethics Scale

Joan M. McMahon
Robert J. Harvey

ABSTRACT. The factor structure of the Multidimensional Ethics Scale (MES; Reidenbach and Robin: 1988, *Journal of Business Ethics* **7**, 871–879; 1990, *Journal of Business Ethics* **9**, 639–653) was examined for the 8-item short form ($N = 328$) and the original 30-item pool ($N = 260$). The objectives of the study were: to verify the dimensionality of the MES; to increase the amount of true cross-scenario variance through the use of 18 scenarios varying in moral intensity (Jones: 1991, *Academy of Management Review* **16**, 366–395); and, to examine the items for measurement precision using item-response theory (IRT) methods. Results of confirmatory and exploratory factor analysis failed to conclusively support the hypothesized 3- (short form) or 5-factor (long form) structure; both instruments were instead dominated by a general factor. Item response theory analyses using Samejima's (1969, *Psychometrika Monograph Supplement* **34**, (4, Pt. 2)) graded response model revealed that many items in the 30-item pool performed very well, and suggested that a different collection of items be used to form a short-form version of the MES. Our proposed 10-item instrument includes more discriminating items than the 8-item version, and has the added advantage of including two items from each of the five ethical philosophies represented in the original 30-item pool.

KEY WORDS: ethical decision-making, ethical judgment, factor analysis, multidimensional ethics scale

Joan M. McMahon is an Assistant Professor of Management in the School of Business at Christopher Newport University, teaching courses in Organizational Behavior, Leadership, and Human Resources. She has a B.A. in Speech from the State University of New York, College at Oneonta; an M.Ed. In Early Childhood Education from James Madison University; and an M.S. and Ph.D. in Industrial/Organizational Psychology from Virginia Polytechnic Institute and State University.

Robert J. Harvey is an Associate Professor of Psychology at Virginia Polytechnic Institute and State University. He has a B.A. in Psychology and an M.A. in Experimental Psychology from the University of Missouri at Kansas City, and a Ph.D. in Industrial/Organizational Psychology from Ohio State University. Dr. Harvey has authored a number of articles in the Journal of Applied Psychology, *the* Journal of Personality Assessment, Personnel Psychology, *and others. He is the author of the chapter on job analysis in the Handbook of Industrial and Organizational Psychology.*

Ethical judgment is included as a major element in many ethical decision-making models (e.g. Hunt and Vitell, 1986; Rest, 1986; Ferrell et al., 1989; Jones, 1991; Street et al., 2001). In light of the ethical failings involving individuals in high organizational positions that have received notoriety recently, it clearly remains a topic of importance for both researchers and practitioners. Unfortunately, many of the methods employed to *operationalize* ethical judgment remain troublesome, especially those involving single-item scales (e.g., Morris and McDonald, 1995; Davis et al., 1998; Frey, 2000a, b; Tsalikis et al., 2001).

Reidenbach and Robin (1988, 1990) developed a multi-item inventory designed to tap five domains relevant to ethical decisions: (a) *Deontology* is concerned with one's duty to follow ethical rules; (b) *Utilitarianism* involves acting in a manner that will provide the greatest good for the greatest number; (c) *Relativism* is based on the idea that no universal ethical rules exist; (d) *Egoism* is concerned with promoting an individual's long-term self-interests; and (e) *Justice* is based on the Aristotelian notion that equals should be treated equally. A 30-item pool (denoted here as the MES-30) was developed to assess these five domains (see Table I); based on exploratory factor analyses (EFA) of it, an 8-item short form (denoted MES-8)

TABLE I

Reidenbach and Robin's items for full and short-form (bold and italicized) MES versions

Item	
Justice	
0	***Just/Unjust (Moral Equity)***
23	***Fair/Unfair (Moral Equity)***
8	Results/Does not result in an equal distribution of good and bad
Relativism	
13	***Culturally acceptable/Unacceptable (Relativism)***
25	Individually acceptable/Unacceptable
11	Acceptable/Unacceptable to people I most admire
3	***Traditionally acceptable/Unacceptable (Relativism)***
27	***Acceptable/Unacceptable to my family (Moral Equity)***
Egoism	
21	Self promoting/Not self promoting
9	Selfish/Not selfish
29	Self sacrificing/Not self sacrificing
10	Prudent/not prudent
19	Under no moral obligation/Morally obligated to act otherwise
4	Personally satisfying/Not personally satisfying
5	In the best interests of the company/Not in the best interests of the company
Utilitarianism	
6	Efficient/Inefficient
16	OK/Not OK if actions can be justified by their consequences
12	Compromises/Does not compromise an important rule by which I live
18	On balance, tends to be good/Bad
7	Produces the greatest/Least utility
17	Maximizes/Minimizes benefits while minimizes/maximizes harm
14	Leads to the greatest/Least good for the greatest number
26	Results in a positive/Negative cost-benefit ratio
28	Maximizes/Minimizes pleasure
Deontology	
24	***Violates/Does not violate an unwritten contract (Contractualism)***
2	Violates/Does not violate my ideas of fairness
20	***Morally right/Not morally right (Moral Equity)***
22	Obligated/Not obligated to act this way
1	***Violates/Does not violate an unspoken promise (Contractualism)***
30	Duty bound to act this way/Not duty bound to act this way

was also produced. Their factor analysis of the MES-8 yielded a three-factor structure (*Moral Equity, Relativism*, and *Contractualism*).

The moral equity dimension includes four items: fair/unfair, just/unjust, acceptable/unacceptable to my family, and morally/not morally right. Reidenbach and Robin (1990) suggest that this dimension represents the notion of good and bad, and taps into concepts of right and wrong learned in childhood. The relativism dimension includes two items: traditionally acceptble/unacceptable, and culturally acceptable/unacceptable. This dimension represents social, rather than individual, considerations and taps into concepts learned through experiences with cultural norms. The contractualism dimension includes two items: violates/does not violate an unspoken promise, and violates/does not violate an unwritten contract. This dimension represents the notion of obligation and social contract.

Psychometric research on the MES

Numerous studies have subsequently examined the psychometric properties of the MES-8 (e.g., Jones, 1991; Reidenbach et al., 1991; Flory et al., 1992; Tansey et al., 1992; Humphreys et al., 1993; LaTour and Henthorne, 1994; Tansey et al., 1994; Henthorne and LaTour, 1995; Clark and Dawson, 1996; LaFleur et al., 1996; Robin et al., 1996; Robin et al., 1997; Simpson et al., 1998; Snipes et al., 1999; Cruz et al., 2000; Cohen et al., 2001; Ellis and Griffith, 2001; Razzaque and Hwee, 2002; Loo, 2004). Many involved performing EFA on the MES-8, with several reporting 3-factor solutions consistent with the Reidenbach and Robin results (e.g., Flory et al., 1992; Humphreys et al., 1993; Clark and Dawson, 1996; LaFleur et al., 1996; Robin et al., 1996, 1997; Simpson et al., 1998; Cruz et al., 2000; Cohen et al., 2001; Ellis and Griffith, 2001; Loo, 2004). For example, Robin et al. (1997) examined ratings of 10 different groups using 18 scenarios, concluding that ''with few exceptions, these statistics support the over-all fit of the ethics model across the varied situational contexts considered here'' (p. 571). However, a 2-factor view of the MES-8 that combines the *Moral Equity* and *Relativism* scales has emerged in some studies (e.g., Reidenbach et al., 1991; Tansey et al., 1992; LaTour and Henthorne, 1994; Henthorne and LaTour, 1995; Snipes et al., 1999; Razzaque and Hwee, 2002), and others have reported results suggesting a 1-factor model (e.g., Tansey et al., 1994). Reidenbach and Robin offer two possible explanations for these varying factor structures: high inter-item correlations that would be expected of items tapping a construct (ethical judgment) that embodies overlapping ethical philosophies (1990); and, the possibility that the relativism and contractualism dimensions provide ''*valuative rules* by which the moral equity of a decision is accessed'' (Robin et al., 1996, p. 281).

Regarding the MES-30, less consistent results have been obtained. In a female sample, Tsalikis and Ortiz-Buonafina (1990) found five factors in three scenarios and six factors in a fourth; for males, five emerged for two scenarios and six emerged for the others. Hansen (1992) reported solutions from 5 to 8 factors, and a 4-factor solution from a subset of the MES-30. Cohen et al. (1993) developed a 20-item subset; after removing items that loaded on more than one factor, analysis of the remaining 15 items, using PC/varimax and the eigenvalues >1.0 rule for retaining factors, yielded solutions that differed appreciably across scenario. Cohen et al. concluded that Reidenbach and Robin's items ''may well provide the basis of multidimensional scales, but a scale must be constructed and validated for each application studied'' (p. 25). Similar results that varied across scenario were reported by Davis et al. (2001) and Kujala (2001).

Unfortunately, a number of concerns can be identified with respect to the above studies. First, Reidenbach and Robin (1988, 1990) designed the MES to be capable of measuring ethical judgments across a range of scenarios. However, during its development and in many subsequent factor analyses, relatively few scenarios were examined, and the factor structure was examined using a *within-scenario* approach (i.e., conducting a separate factor analysis for each scenario). When one's goal involves identifying the general dimensions underlying ethical judgments, such a strategy may well be problematic; arguably, one should instead seek to maximize variability across scenarios in the ratings (particularly, with respect to situational attributes that may influence ethical judgments). For the same reason that one would not attempt to identify the general dimensions of personality by factoring only ratings made of people who are highly introverted, neurotic, disagreeable, conscientious, and open to experience, factoring the MES on a scenario-by-scenario basis is unlikely to identify stable and comprehensive dimensions.

Second, the factor analytic methods that were used have received some criticism (e.g., see Linn, 1968; Tucker et al., 1969; Lee and Comrey, 1979; Ford et al., 1986; Fabrigar et al., 1999). For example, many studies used the principal components model, which assumes that *no measurement errors* or other construct-irrelevant sources of variance exist; such an assumption is quite tenuous for ratings collected using a 1–7 bipolar scale. Many studies relied on the eigenvalues >1.0 rule to determine the number of factors; this rule has repeatedly been shown to lead to erroneous results. Orthogonal rotations were the norm; unfortunately, if the true latent dimensions are non-orthogonal, such methods may do a poor job recovering that structure (even if the correct number of factors is retained). Thus,

although it is possible that the dimensionality of the MES varies across situations, the lack of consistency seen in past studies may simply be illusory, reflecting the use of arbitrary decision rules (particularly eigenvalues >1.0) that Monte Carlo studies have long shown to be potentially misleading (e.g., Tucker et al., 1969; Fabrigar et al., 1999).

The present study

Given these limitations, we concluded that additional psychometric research on the MES is essential. First, given that past research offers a firm foundation for specifying competing dimensional hypotheses (e.g., for MES-8, 1-, 2-, and 3-factor; for MES-30, 1- and 5-factor), we used *confirmatory* factor analysis (CFA) methods to examine the dimensionality of the MES. Based on past EFA results, we hypothesized that CFA would show good model-fit for the 3-factor model of the MES-8 and the 5-factor view of the MES-30. However, if a 1-factor model can provide levels of model-fit that rival the alternative models, such a finding would directly question the MES's success regarding a primary objective (i.e., providing a *multidimensional* view of ethical perceptions). Likewise, even if the 3- or 5-factor models fit better than the 1-factor model, if levels of factor correlation become excessive in the higher-dimensional models, such a finding would question the MES's success in meeting its primary goal.

Second, we wanted to increase the amount of true cross-scenario variance. This was accomplished by (a) using a larger number of scenarios (18) than has typically been seen in past studies, (b) attempting to vary the types of ethical decisions depicted via developing scenarios varying in *moral intensity* (see Jones, 1991), and (c) conducting factor analyses at the cross-scenario level. However, the question of whether scenario-specific MES variance exists remains an important one. That is, with the notable exception of Cohen et al. (1993), most researchers have tended to view the MES's goal as being to define a common profile of underlying ethical constructs whose dimensional structure is consistent across situations. In an attempt to determine the degree to which the dimensional structure of the MES is scenario-influenced, we compared the fit of CFA models incorporating only scenario-based factors (i.e., one factor per scenario on which all MES items load) against models that fit only ''trait'' (MES dimension) factors, and models with both. From the perspective of the MES's developers, one would hope to find that scenario factors do not provide better model-fit than MES factors; however, in light of past research on method-variance (e.g., Harvey et al., 1985), we hypothesized that even if good fit for the MES models is obtained, scenario-based factors would further improve fit.

Finally, we wanted to examine the performance of the item-analysis methods used to produce the MES-8. Specifically, to what extent did these methods produce a short-form that strikes a good balance between including the most discriminating items versus representing all of the constructs in the full pool? Based on our review of the item analysis procedures used in past studies, as well as our first hypothesis above, we hypothesized that significant differences would be seen between the MES-8 pool versus a short-form developed using item-response theory (IRT) methods. Since the MES uses 7-point ordered-category scales, we used Samejima's (1969) graded-response IRT model to calibrate the MES-30; items were selected for a short-form version based on both their IRT discrimination (*a*) parameters, as well as their a priori category in the Reidenbach and Robin taxonomy. However, even if the pools formed via principal component (PC)- versus IRT-based methods are similar, it is still important to use IRT to assess the degree of measurement precision provided by the MES. That is, although many studies (e.g., Loo, 2004) have examined the MES using classical test theory (CTT) reliability estimates, scales typically do not provide consistent measurement precision across the full range of scores (e.g., see Harvey and Murry, 1994; Embretson and Reise, 2000). Thus, we were interested in examining test information functions for the MES to assess both the amount of precision, and consistency of precision, across the scale.

Method

Participants and instruments

Two subject pools of undergraduates at a large southeastern university were used; the first

($N = 328$) was used strictly to examine the psychometric properties of the MES-8; the second ($N = 260$) was used to examine the psychometric properties of the MES-30 (which includes the 8 items from the MES-8). Extra credit toward psychology classes was awarded to all participants. Eighteen scenarios that varied in terms of moral intensity (Jones, 1991) were used (see McMahon and Harvey, 2006 for the actual scenarios used and details regarding scenario development). Table I summarizes the thirty MES items. A bipolar format was used to present each MES item, with seven rating points provided between the poles. Data collection for the MES-8 pool used a paper-and-pencil format; a web-based survey was used with the MES-30 pool, with unanchored radio buttons to record responses. In the MES-30 sample, an additional item was included to collect a global judgment (item 15); hence, MES-30 items are numbered from 0–14 and 16–30. Item responses were scaled such that a '1' represented a low ethical rating (that is, the participant judged the questionable action taken in the scenario as being ethical), and a '7' represented a high ethical rating. In the MES-30 sample, items 1, 2, 9, and 24 were reverse-scored in order to achieve that scaling.

Procedure

Participants were asked to read scenarios describing business situations with ethical overtones. Three versions (control, low, and high moral intensity; see Jones, 1991) of 18 scenarios were used; in the MES-8 sample, participants were randomly assigned to the control, low, or high intensity condition and saw the 18 scenarios appropriate for their condition in one of three different orders of presentation. For the MES-30 sample, to reduce the length of the rating task, participants were presented with low and high versions of six randomly selected scenarios (ordering of each pair was randomized); raters could not progress to the next scenario until they provided non-missing responses for all items. For both samples, in addition to the MES, 12 items measuring perceived moral intensity were also rated as a check on the intensity manipulation. In the MES-30 sample, one additional item was included as a measure of intention (i.e. ''I would have made the same decision'').

Analyses

MES-8

CFAs were performed via SAS/PROC CALIS with maximum-likelihood estimation using the covariance matrix; follow-up oblique EFAs were conducted using PROC FACTOR with the common-factor model (squared-multiple correlation communality estimates, oblique Harris–Kaiser rotation with $p = 0.5$). Three ''trait'' models were examined in the CFAs: (a) the 3-factor Reidenbach and Robin (1988, 1990) model of *Moral Equity, Relativism*, and *Contractualism* (see Table I); (b) a 2-factor model based on studies (e.g., Henthorne and LaTour, 1995) in which the *Moral Equity* and *Relativism* items loaded together; and (c) a general-factor model in which all MES items load on a single dimension. CFAs were conducted on aggregated data (i.e. by averaging ratings of each MES item across the 18 scenarios) and on disaggregated data (i.e., analyzing a 144-by-144 matrix in which each scenario was rated on each MES item). Factor metric was determined by fixing each variance to 1.0. To address the ''scenario factor'' issue, additional models included an 18-factor model in which the 8 MES items in each scenario loaded on a factor, as well as variants of the ''trait'' models that included the 18 scenario factors.

MES-30

Since each participant rated 44 items on both the high and low version of 6 scenarios, and because all raters did not rate all possible scenarios, it was not possible to test disaggregated CFA models of the MES-30. To ensure that the ratings were made with maximum rater attention and minimum potential fatigue – while maintaining diversity in rated scenarios – the ratings used in the factor analyses were taken from the first scenario presented to each MES-30 participant. Since these were randomly selected from the 36 possible high/low scenarios, diversity was assured. CFA models tested 1- and 5-factor models (the original 5 factors theorized by Reidenbach and Robin, see Table I) as well as a 3-factor model in which the item parameters for items in the two MES-30 domains that did not map onto the three MES-8 domains (see Table I) were estimated using only unique variances. This latter model was included to quantify the impact of

moving from a 5- to 3-dimensional view of ethics when the MES-8 was developed. EFA was conducted for 1-factor through 12-factor models. IRT analyses for MES-30 were conducted using MULTILOG 6.0.

Results and discussion

MES-8 factor analyses

Table II presents the CFA results for the short-form MES items. As hypothesized, in the aggregated data these results indicate that the 3-factor model provides the best model fit, and that the oblique version fits much better than the orthogonal (i.e., lower χ^2 values reflect smaller discrepancies between the actual versus reproduced variance/covariance matrices, and thus better fit; in contrast, values of GFI, AGFI, CFI and NNFI range from 0 to 1, with higher values reflecting better fit). However, the high factor correlations between these three supposedly distinct constructs (i.e., $r = 0.57$, 0.59, and 0.69 for *Moral Equity – Relativism*, *Moral Equity – Contractualism*, and *Relativism – Contractualism*, respectively) raise questions regarding discriminant validity. Consistent with the view that the MES-8 primarily measures a general ethical dimension, the 1-factor model provides fit levels that are relatively close to the 3-factor results, and the EFA results using data aggregated across the 18 scenarios show that the first factor accounts for 81.7% of the total common variance (6.5:1 ratio of first-to-second eigenvalues). Thus, although a scree plot indicates a break at three, it also confirms that a powerful first factor is present. Similar results are seen using the data from the MES-30 sample (which does not aggregate over scenario, but does reflect heterogeneity in rated scenario) when factoring only the eight items in common between the two (models 13–14): a virtually identical scree plot is present, and the first factor is even stronger (96.4% of total common variance, 10.4:1 ratio).

Regarding the question of scenario-based variance, the Table II results for disaggregated data (i.e., with items representing the 144 combinations of MES and scenario) show that models with only orthogonal and oblique scenario-factors (models 3, 4) provide much stronger fit (e.g., CFI = 0.58 and 0.60 versus 0.18 and 0.19) than the substantive 3-factor MES models (1, 2). Although improvements in fit are seen when the substantive factors are added to the scenario-factor models (models 5–7), the scenario factors explain considerably more variance in MES ratings than the substantive MES factors when rating targets are not homogeneous. As in the aggregated data, only relatively minor improvements in fit are achieved when comparing the 1- versus 3-factor models (5 versus 6–7). Thus, empirical support exists in both types of data regarding the superiority of the CFA 3-factor model over the 1- and 2-factor models; however, the EFA results indicate that a 1-factor view of the MES-8 is also plausible (and perhaps preferable, given the strong inter-factor correlations), and the disaggregated-data results indicate that the factor structure of the MES-8 is clearly not invariant across scenarios (even for the 1-factor model).

MES-30 factor analyses

Table III presents the CFA results for the full MES item pool; as with the results for the MES-8, the 1-factor model provides levels of model-fit that approach the levels seen for the 5-factor (deontology, utilitarianism, relativism, egoism, and justice) multidimensional model advanced by Reidenbach and Robin (even more closely in this case). As with the MES-8, the oblique Reidenbach and Robin model suffered from disturbingly high factor correlations (in the present case, factor correlations ranged from 0.90 to 1.0, with two estimates – involving *Justice-Egoism* and *Justice-Deontology* – exhibiting a Heywood case by attempting to exceed 1.0) that tend to give preference to the 1-factor model on the basis of parsimony and discriminant validity. Consistent with this view, the results for the 3-factor models (which included all items, but viewed items not classified into the three domains in the MES-8 factor structure as being 100% unique variance) show very poor model fit. Clearly, this 30-item pool is dominated by a general factor representing ethical perceptions.

The fit indices for follow-up EFA analyses on the full item pool are shown in Table IV; here, maximum-likelihood EFAs were conducted, and the model-fit indices were examined to shed light on the underlying dimensionality of the item pool (as with

TABLE II

Confirmatory factor analysis goodness-of-fit indices for 1-factor, 2-factor, 3-factor, and 18 scenario-factor models using MES-8 item pool

Model	χ^2	df	GFI	AGFI	CFI	NNFI
1. 3-factor orthogonal, disaggregated data	55394.8929***	10152	0.2219	0.1999	0.1834	0.1718
2. 3-factor oblique, disaggregated data	55084.3066***	10149	0.2218	0.1995	0.1889	0.1772
3. 18 scenario factors (orthogonal), disaggregated data	33696.9537***	10152	0.2764	0.2559	0.5750	0.5690
4. 18 scenario factors (oblique), disaggregated data	32290.6738***	9999	0.2861	0.2546	0.5976	0.5857
5. 18 scenario factors plus 1 general factor (oblique within scenarios), disaggregated data	27707.7758***	9855	0.4241	0.3899	0.6778	0.6633
6. 3-factor plus 18 scenario factors (all orthogonal), disaggregated data	26439.7645***	10008	0.4760	0.4533	0.7034	0.6949
7. 3-factor plus 18 scenario factors (oblique within cluster), disaggregated data	25004.9049***	9852	0.5053	0.4757	0.7265	0.7142
8. Aggregated data, 1-factor model	786.0004***	20	0.6831	0.4297	0.7261	0.6166
9. Aggregated data, 3-factor orthogonal model	450.4552***	20	0.7482	0.5467	0.8461	0.7845
10. Aggregated data, 3-factor oblique model	103.7641***	17	0.9270	0.8455	0.9690	0.9489
11. Aggregated data, 2-factor orthogonal	591.1287***	20	0.7445	0.5401	0.7958	0.7141
12. Aggregated data, 2-factor oblique	390.5119***	19	0.8004	0.6219	0.8672	0.8042
13. 3-factor using only 8 short-form items from MES-30, orthogonal	250.0701***	20	0.8090	0.6562	0.7872	0.7020
14. 3-factor using only 8 short-form items from MES-30, oblique	39.3912**	17	0.9628	0.9211	0.9793	0.9659

Note: GFI = goodness of fit index; AGFI = adjusted GFI; CFI = Bentler's comparative fit index; NNFI = Bentler and Bonnet non-normed fit index. Models 1–12 are based on data from the MES-8 sample; Models 13–14 are based on the MES-30 sample, using only the 8 items common to both.

***Denotes χ^2 values $p < 0.001$.

**Denotes χ^2 values $p < 0.01$.

TABLE III

Confirmatory factor analysis goodness-of-fit indices for 1-factor, 3-factor, and 5-factor models using MES-30 item pool

Model	χ^2	df	GFI	AGFI	CFI	NNFI
1. 1-factor	1202.5955***	405	0.7299	0.6898	0.8102	0.7962
2. 3-factor from short form, orthogonal	3779.3821***	427	0.2531	0.1867	0.2024	0.1875
3. 3-factor short form, oblique	3568.7032***	424	0.2864	0.2174	0.2518	0.2324
4. 5-factor a priori, orthogonal	2367.4798***	405	0.5900	0.5292	0.5331	0.4985
5. 5-factor a priori, oblique[a]	1095.7645***	397	0.7559	0.7140	0.8338	0.8178

Note. GFI = goodness-of-fit index; AGFI = adjusted GFI; CFI = Bentler's comparative fit index; NNFI = Bentler and Bonnet non-normed fit index. $N = 260$.

[a]Contains two out-of-bounds estimates of factor correlations that were fixed at 1.0.

***Denotes χ^2 values $p < 0.001$.

the CFA results, lower values of χ^2 reflect improved model fit; for the remaining fit indices, lower values of SBC and AIC – and higher values of TLRC, which ranges from 0 to 1 – reflect better fit and smaller residuals). Consistent with the CFA results, although models in higher dimensionalities improve somewhat over the strong results seen for the 1-factor model, the 30-item pool is clearly dominated by a general factor (e.g., 78% of the estimated total common variance is accounted for by the 1-factor model).

MES-30 IRT

IRT item parameters for the MES-30 items (Table V) were derived using the GRM. Given the dominant general factor seen above, all MES-30 items (with the exception of 10 – ''Prudent,'' which had a negative item-total correlation and was therefore excluded) were calibrated using IRT; the test standard-error functions are presented in Figure 1. As hypothesized, although the MES-8 includes many of the items with high *a* parameters, it

TABLE IV

Exploratory factor analysis fit indices for 1-factor through 12-factor models using MES-30 item pool

# Factors	χ^2	df	AIC	SBC	TLRC	% TCV
1	1202.5955***	405	392.5955	– 1049.4806	0.8006	0.7832
2	900.6334***	376	148.6334	– 1190.1829	0.8608	0.8659
3	728.4494***	348	32.4494	– 1206.6678	0.8926	0.9164
4	618.8046***	321	– 23.1954	– 1166.1742	0.9102	0.9515
5	507.1572***	295	– 82.8428	– 1133.2439	0.9322	0.9838
6	419.1350***	270	– 120.8650	– 1082.2491	0.9497	1.0074
7	349.2791***	246	– 142.7209	– 1018.6486	0.9637	1.0275
8	284.58932**	223	– 161.41068	– 955.44269	0.97868	1.0441
9	238.75442*	201	– 163.24558	– 878.94258	0.98800	1.0583
10	197.46316	180	– 162.53684	– 803.45953	0.99752	1.0706
11	158.78663	160	– 161.21337	– 730.92243	1.00838	1.0813
12	128.92154	141	– 153.07846	– 655.13457	1.01651	1.0877

Note. AIC = Akaike information criterion; SBC = Schwarz Bayesian criterion; TLRC = Tucker–Lewis reliability criterion; %TCV = percentage of estimated total common variance accounted for by factor solution.

***Denotes χ^2 values $p < 0.001$.

**Denotes χ^2 values $p < 0.01$.

*Denotes χ^2 values $p < 0.05$.

TABLE V

Item parameters for 30 MES items

	Item	Loading	Item-total	Mean	SD	a	b1	b2	b3	b4	b5	b6
27	***Acceptable to my family (Relativism)****	***0.843***	***0.791***	***5.45***	***1.65***	***2.75***	***-2.98***	***2.02***	***−1.36***	***−0.76***	***−0.27***	***0.57***
18	On balance tends to be good (Utilitarianism)*	0.843	0.817	4.94	1.68	2.43	−2.73	−1.85	−1.07	−0.50	0.33	1.30
20	***Morally right (Deontology)****	***0.840***	***0.787***	***5.58***	***1.42***	***2.62***	***−3.49***	***−2.36***	***−1.79***	***−1.14***	***−0.29***	***0.71***
23	***Fair (Justice)****	***0.837***	***0.780***	***5.45***	***1.65***	***2.44***	***−2.97***	***−1.90***	***−1.36***	***−0.96***	***−0.23***	***0.62***
25	Individually acceptable (Relativism)*	0.793	0.788	4.97	1.77	2.20	−2.66	−1.96	−1.03	−0.46	0.17	0.97
11	Acceptable to people I most admire (Relativism)	0.782	0.739	5.47	1.53	1.87	−4.40	−2.56	−1.78	−0.85	−0.25	0.74
0	***Just (Justice)****	***0.748***	***0.729***	***5.15***	***1.80***	***1.82***	***−2.75***	***−1.84***	***−1.16***	***−0.72***	***−0.13***	***1.00***
14	Leads to the greatest good for the greatest number (Utilitarianism)*	0.737	0.722	4.80	1.73	1.74	−2.78	−1.84	−1.00	−0.33	0.34	1.53
17	Maximizes benefits while minimizes harm (Utilitarianism)	0.702	0.681	4.69	1.70	1.57	−3.10	−2.15	−0.95	−0.02	0.65	1.52
9	Selfish (Egoism-R)*	0.696	0.634	5.69	1.66	1.59	−3.14	−2.33	−1.71	−1.28	−0.66	0.29
2	Violates my idea of fairness (Deontology-R)*	0.682	0.625	5.47	1.73	1.73	−2.85	−1.99	−1.53	−1.22	−0.56	0.68
19	Under no moral obligation to act otherwise (Egoism)*	0.662	0.628	5.02	1.70	1.74	−3.01	−2.29	−1.45	−0.62	0.06	1.22
16	OK if actions can be justified by consequences (Utilitarianism)	0.654	0.652	4.42	1.86	1.32	−2.38	−1.77	−0.83	0.04	0.76	1.83
30	Duty bound to act this way (Deontology)	0.649	0.648	5.15	1.60	1.36	−3.82	−2.67	−1.82	−0.53	0.14	1.08
8	Results in an equal distribution of good and bad (Justice)	0.648	0.628	4.94	1.72	1.43	−2.98	−2.09	−1.54	−0.41	0.21	1.40
22	Obligated to act this way (Deontology)	0.628	0.620	5.24	1.60	1.32	−3.84	−3.01	−1.55	−0.77	−0.03	1.10
3	***Traditionally acceptable (Relativism)***	***0.610***	***0.618***	***4.77***	***1.85***	***1.27***	***−3.29***	***−1.71***	***−1.06***	***−0.41***	***0.40***	***1.51***
6	Efficient (Utilitarianism)	0.602	0.612	4.25	1.97	1.09	−2.78	−1.45	−0.44	0.41	0.90	1.83
24	***Violates an unwritten contract (Deontology-R)***	***0.577***	***0.547***	***5.05***	***1.79***	***1.18***	***-3.42***	***−2.23***	***−1.56***	***−0.75***	***−0.04***	***1.25***
1	***Violates an unspoken promise (Deontology-R)***	***0.576***	***0.542***	***5.11***	***1.85***	***1.17***	***-2.87***	***−2.14***	***−1.62***	***−0.89***	***−0.27***	***1.26***
13	***Culturally acceptable (Relativism)***	***0.557***	***0.550***	***4.33***	***1.87***	***1.02***	***−3.29***	***−1.70***	***−0.59***	***−0.03***	***0.95***	***2.09***
7	Produces the greatest utility (Utilitarianism)	0.499	0.508	4.01	1.64	0.82	−3.76	−2.01	−0.90	1.22	2.05	3.21
5	In the best interests of the company (Egoism)	0.487	0.488	4.26	2.14	0.84	−2.99	−1.29	−0.39	0.25	0.67	1.80
26	Results in a positive cost-benefit ratio (Utilitarianism)	0.440	0.465	4.11	1.89	0.73	−3.53	−1.88	−0.84	0.55	1.40	2.74
12	Compromises an important rule by which I live (Utilitarianism-R)	0.426	0.395	4.59	1.95	0.89	−3.21	−2.13	−1.17	−0.39	0.33	1.89
4	Personally satisfying (Egoism)	0.409	0.402	4.70	1.97	0.81	−3.32	−2.33	−1.52	−0.35	0.32	1.69
28	Maximizes pleasure (Utilitarianism)	0.367	0.372	4.15	1.79	0.59	−4.62	−2.64	−1.17	0.92	1.85	3.36
29	Self sacrificing (Egoism)	0.248	0.231	5.04	1.82	0.50	−6.42	−4.36	−3.08	−1.11	−0.02	1.84
21	Self promoting (Egoism)	0.180	0.185	3.53	2.06	0.21	−7.42	−1.68	1.18	3.21	5.67	9.09
10	Prudent (Egoism)	−0.024	−0.036	3.85	1.66	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note. Loading = loading of each item on the unrotated first principal axis; item–total = corrected item–total correlation; *a* and *b1–b6* present item parameters from Samejima's (1969) GRM calibration. $N = 260$. Bold-italic items denote ones included in the 8-item short-form version of the MES.

*Denotes items that would be included in a 10-item short form based on magnitude of loading on general factor and content-domain balance.

also includes others that rank much lower in terms of discriminating power. Based on the IRT findings, we suggest a 10-item short form, rather than the 8-item MES.

To select the new 10-item subset, items were first ranked in terms of descending values of their loading on the first unrotated principal axis underlying the item correlation matrix (which produces a ranking quite similar to that produced by sorting on the *a* parameter; in both cases it is designed to identify items offering the highest levels of information regarding the underlying trait). We then sampled the top two items from each of the five *a priori* dimensions (for deontology: morally right/not morally right, and, violates/does not violate my idea of fairness; for utilitarianism: on balance tends to be good/bad, and, leads to the greatest/least good for the greatest number; for relativism: acceptable/unacceptable to my family, and, individually acceptable/unacceptable; for egoism: selfish/not selfish, and, under no moral obligation/morally obligated to act otherwise; and for justice: fair/unfair, and, just/unjust). Interestingly, the standard-error functions for the 29-item pool (the MES-30 excluding ''Prudent'') and our 10-item subset exhibit respectable measurement precision, with a flat function from approximately 1 z unit above the mean through -3 z units.

Although our proposed 10-item short form (''MES-10'', see Table VI) unavoidably provides less information, it is encouraging to note that a scale one-third the length of the MES-30 still produces consistent precision across a wide range. The results for the MES-8, in contrast, show clearly inferior measurement precision with respect to the MES-10, with standard errors over twice the size of those seen

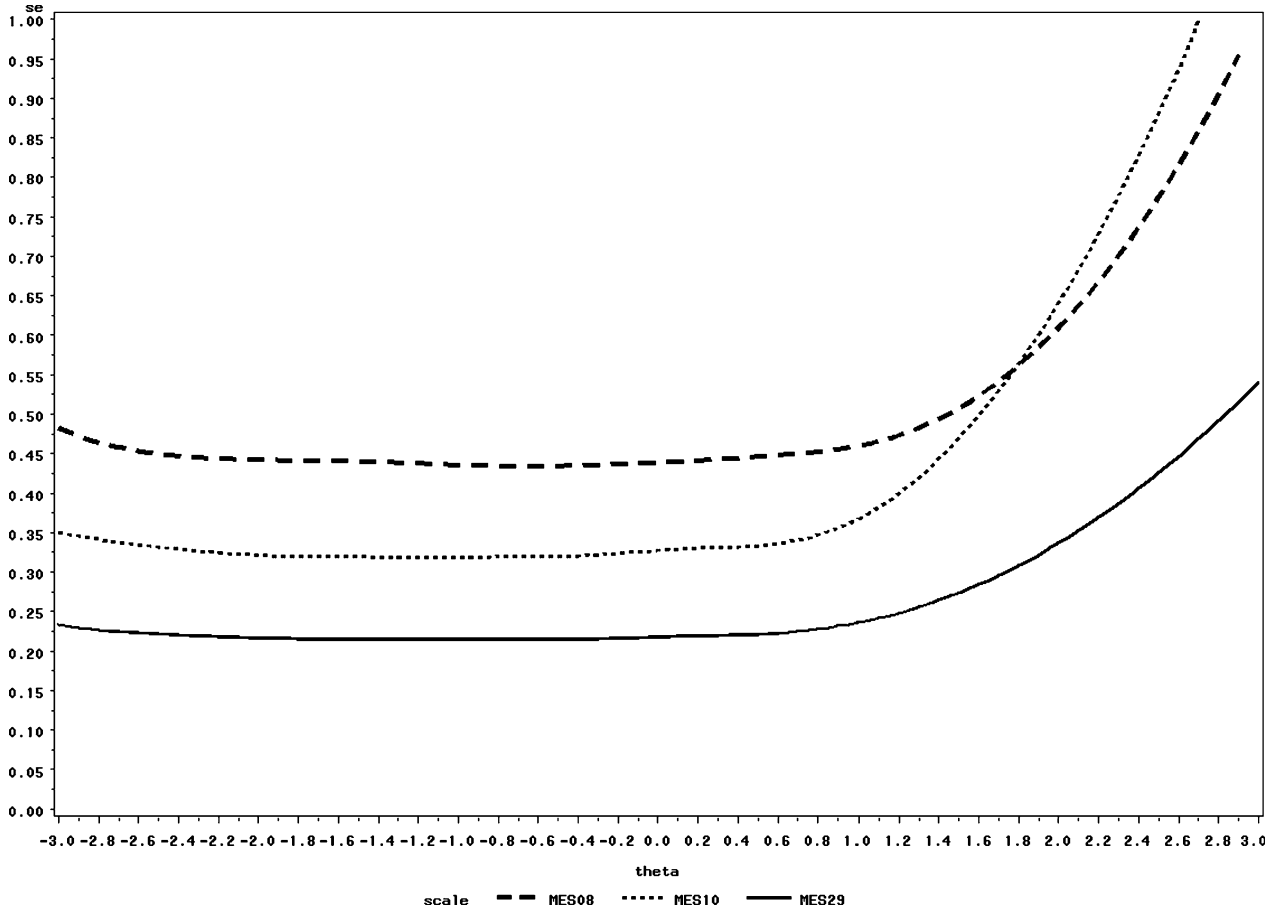

Figure 1. Test standard error functions for 29 MES item pool (solid line), 10-item short form (finely dashed line), and MES-8 pool (heavy dashed line) calibrated using the Samejima (1969) graded response model.

TABLE VI
Proposed "MES-10" items

Deontology
Morally right/Not morally right
Violates/Does not violate my idea of fairness
Utilitarianism
On balance tends to be good/Bad
Leads to the greatest/Least good for the greatest number
Relativism
Acceptable/Unacceptable to my family
Individually acceptable/Unacceptable
Egoism
Selfish/Not selfish
Under no moral obligation/Morally obligated to act otherwise
Justice
Fair/Unfair
Just/Unjust

for corresponding values produced by the full pool. An additional benefit of the MES-10 is that it retains 2 items from each of the five ethical domains used to inform Reidenbach and Robin (1988) in their initial development of items, thus achieving their original goal of providing a multidimensional instrument for ethical judgment based on strong philosophical underpinnings.

Conclusions

Our results provide a "good news and bad news" view of the MES. Let's look at the bad news first and get it out of the way. On the negative side, our results contradicted the view that the MES (either the MES-8 or the MES-30) provides a clearly multidimensional assessment of ethical perceptions; although interpretable higher-dimensionality solutions can be produced, their high cross-scale correlations raise serious concerns regarding discriminant validity. Additionally, even for the 1-factor view of the MES, our results indicate that scenario-based effects may play a significant role. Accordingly, if the MES is to be used as an indicator of one's general pattern of ethical decisions or perceptions, it would be advisable to collect measures across a wide range of scenarios and then aggregate across scenarios.

On the positive side, the long-form (MES-30) pool was shown, via IRT, to provide strong measurement precision across a wide range of scale values, and to provide its maximum precision in the range in which it is arguably most important (i.e., for average-and-below ethical perceptions). Our suggested MES-10 produces measurement precision consistent with the MES-30 and superior to the MES-8. In terms of utility, the length of both the MES-8 and the MES-10 make them more practical than the MES-30. However, with 2 items representing each of five ethical domains (deontology, utilitarianism, relativism, egoism, and justice), the MES-10 has the potential to be more theoretically informative than the MES-8. Further research on the MES-10 is encouraged.

References

Clark, J. W. and L. E. Dawson: 1996, 'Personal Religiousness and Ethical Judgements: An Empirical Analysis', *Journal of Business Ethics* **15**, 359–372.

Cohen, J., L. Pant and D. Sharp: 1993, 'A Validation and Extension of a Multidimensional Ethics Scale', *Journal of Business Ethics* **12**, 13–26.

Cohen, J. R., L. W. Pant and D. J. Sharp: 2001, 'An Examination of Differences in Ethical Decision-making

between Canadian Business Students and Accounting Professionals', *Journal of Business Ethics* **30**, 319–336.

Cruz, C. A., W. E. Shafer and J. R. Strawser: 2000, 'A Multidimensional Analysis of Tax Practitioner's Ethical Judgements', *Journal of Business Ethics* **24**, 223–244.

Davis, M. A., M. G. Andersen and M. B. Curtis: 2001, 'Measuring Ethical Ideology in Business Ethics: A Critical Analysis of the Ethics Position Questionnaire', *Journal of Business Ethics* **32**, 35–53.

Davis, M. A., N. B. Johnson and D. G. Ohmer: 1998, 'Issue-Contingent Effects on Ethical Decision Making: A Cross-Cultural Comparison', *Journal of Business Ethics* **17**, 373–389.

Ellis, T. S. and D. Griffith: 2001, 'The Evaluation of IT Ethical Scenarios Using a Multidimensional Scale', *The DATA BASE for Advances in Information Systems* **32**, 75–85.

Embretson, S. E. and S. P. Reise: 2000, *Item Response Theory for Psychologists* (Erlbaum, Mahwah, NJ).

Fabrigar, L. R., D. T. Wegener, R. C. MacCallum and E. J. Strahan: 1999, 'Evaluating the Use of Exploratory Factor Analysis in Psychological Research', *Psychological Methods* **4**(3), 272–299.

Ferrell, O. C., L. G. Gresham and J. Fraedrich: 1989, 'A Synthesis of Ethical Decision Models for Marketing', *Journal of Macromarketing* **11**, 55–64.

Flory, S. M., T. J. Phillips, Jr., R. E. Reidenbach and D. P. Robin: 1992, 'A Multidimensional Analysis of Selected Ethical Issues in Accounting', *The Accounting Review* **67**, 284–302.

Ford, J. K., R. C. MacCallum and M. Tait: 1986, 'The Application of Exploratory Factor Analysis in Applied Psychology: A Critical Review and Analysis', *Personnel Psychology* **39**(2), 291–314.

Frey, B. F.: 2000a, 'The Impact of Moral Intensity on Decision Making in a Business Context', *Journal of Business Ethics* **26**, 181–195.

Frey, B. F.: 2000b, 'Investigating Moral Intensity with the World-Wide Web: A Look at Participant Reactions and a Comparison of Methods', *Behavior Research Methods, Instruments, and Computers* **32**, 423–431.

Hansen, R. S.: 1992, 'A Multidimensional Scale for Measuring Business Ethics: A Purification and Refinement', *Journal of Business Ethics* **11**, 523–534.

Harvey, R. J., R. S. Billings and K. J. Nilan: 1985, 'Confirmatory Factor Analysis of the Job Diagnostic Survey: Good News and Bad News', *Journal of Applied Psychology* **70**, 461–468.

Harvey, R. J. and W. D. Murry: 1994, 'Scoring the Myers-Briggs Type Indicator: Empirical Comparison of Preference Score versus Latent-trait Methods', *Journal of Personality Assessment* **62**, 116–129.

Henthorne, T. L. and M. S. LaTour: 1995, 'A Model to Explore the Ethics of Erotic Stimuli in Print Advertising', *Journal of Business Ethics* **14**, 561–569.

Humphreys, N., D. P. Robin, R. E. Reidenbach and D. L. Moak: 1993, 'The Ethical Decision Making Process of Small Business Owner/Managers and Their Customers', *Journal of Small Business Management* **31**, 9–22.

Hunt, S. D. and S. Vitell: 1986, 'A General Theory of Marketing Ethics', *Journal of Macromarketing* **6**, 5–16.

Jones, T. M.: 1991, 'Ethical Decision Making by Individuals in Organizations: An Issue-Contingent Model', *Academy of Management Review* **16**, 366–395.

Kujala, J.: 2001, 'A Multidimensional Approach to Finnish Managers' Moral Decision-Making', *Journal of Business Ethics* **34**, 231–254.

LaFleur, E. K., R. E. Reidenbach, D. P. Robin and P. J. Forrest: 1996, 'An Exploration of Rule Configuration Effects on the Ethical Decision Processes of Advertising Professionals', *Journal of the Academy of Marketing Science* **24**, 66–76.

LaTour, M. S. and T. L. Henthorne: 1994, 'Ethical Judgments of Sexual Appeals in Print Advertising', *Journal of Advertising* **23**, 81–90.

Lee, H. B. and A. L. Comrey: 1979, 'Distortions in a Commonly Used Factor Analytic Procedure', *Multivariate Behavioral Research* **14**, 301–321.

Linn, R. L.: 1968, 'A Monte Carlo Approach to the Number of Factors Problem', *Psychometrika* **33**(1), 37–71.

Loo, R.: 2004, 'Support for Reidenbach and Robin's (1990) Eight-item Multidimensional Ethics Scale', *The Social Science Journal* **41**, 289–294.

McMahon, J. M. and R. J. Harvey: 2006, 'An Analysis of the Factor Structure of Jones' Moral Intensity Construct', *Journal of Business Ethics* **64**, 381–404.

Morris, S. A. and R. A. McDonald: 1995, 'The Role of Moral Intensity in Moral Judgments: An Empirical Investigation', *Journal of Business Ethics* **14**, 715–726.

Razzaque, M. A. and T. P. Hwee: 2002, 'Ethics and Purchasing Dilemma: A Singaporean View', *Journal of Business Ethics* **35**, 307–326.

Reidenbach, R. E. and D. P. Robin: 1988, 'Some Initial Steps Toward Improving the Measurement of Ethical Evaluations of Marketing Activities', *Journal of Business Ethics* **7**, 871–879.

Reidenbach, R. E. and D. P. Robin: 1990, 'Toward the Development of a Multidimensional Scale for Improving Evaluations of Business EthRoics', *Journal of Business Ethics* **9**, 639–653.

Reidenbach, R. E., D. P. Robin and L. Dawson: 1991, 'An Application and Extension of a Multidimensional

Ethics Scale to Selected Marketing Practices and Marketing Groups', *Journal of the Academy of Marketing Science* **19**(2), 83–92.

Rest, J. R.: 1986, *Moral Development: Advances in Research and Theory* (Praeger, New York).

Robin, D. P., E. W. King and R. E. Reidenbach: 1996, 'The Effect of Attorneys' Perceived Duty to Client on Their Ethical Decision Making Process', *American Business Law Journal* **34**, 277–299.

Robin, D. P., R. E. Reidenbach and B. J. Babin: 1997, 'The Nature, Measurement, and Stability of Ethical Judgments in the Workplace', *Psychological Reports* **80**, 563–580.

Samejima F. (1969). Estimation of Latent Ability using a Response Pattern of Graded Scores. Psychometrika Monograph Supplement **34** (4, Pt. 2).

Simpson, P. M., G. Brown and R. E. Widing, II: 1998, 'The Association of Ethical Judgment of Advertising and Selected Advertising Effectiveness Response Variables', *Journal of Business Ethics* **17**, 125–136.

Snipes, R. L., M. S. LaTour and S. J. Bliss: 1999, 'A Model of the Effects of Self-efficacy on the Perceived Ethicality and Performance of Fear Appeals in Advertising', *Journal of Business Ethics* **19**, 273–285.

Street, M. D., S. C. Douglas, S. W. Geiger and M. J. Martinko: 2001, 'The Impact of Cognitive Expenditure on the Ethical Decision-making Process: The Cognitive Elaboration Model', *Organizational Behavior and Human Decision Processes* **86**, 256–277.

Tansey, R., G. Brown, M. R. Hyman and L. E. Dawson, Jr.: 1994, 'Personal Moral Philosophies and the Moral Judgments of Salespeople', *Journal of Personal Selling and Sales Manager* **14**, 59–75.

Tansey, R., M. R. Hyman and G. Brown: 1992, 'Ethical Judgments About Wartime ads Depicting Combat', *Journal of Advertising* **21**, 57–74.

Tsalikis, J. and M. Ortiz-Buonafina: 1990, 'Ethical Beliefs' Differences of Males and Females', *Journal of Business Ethics* **9**, 509–517.

Tsalikis, J., B. Seaton and P. L. Shepherd: 2001, 'Relativism in Ethical Research: A Proposed Model and Mode of Inquiry', *Journal of Business Ethics* **32**, 231–246.

Tucker, L. R., R. F. Koopman and R. L. Linn: 1969, 'Evaluation of Factor Analytic Research Procedures by Means of Simulated Correlation Matrices', *Psychometrika* **34**(4), 421–459.

Joan M. McMahon
Management and Marketing,
Christopher Newport University,
1 University Place, Newport News,
VA, 23606,
U.S.A.
E-mail: jmcmahon@cnu.edu

Robert J. Harvey
Virginia Polytechnic Institute and State University,
121 Williams Hall, Blacksburg,
VA, 24061,
U.S.A.
E-mail: harveyrj@vt.edu