A MEASURE OF ADHERENCE TO CONVENTIONAL MORALITY

WILLIAM S. TOOKE
State University of New York—Plattsburgh

WILLIAM ICKES
University of Texas at Arlington

The authors describe the construction and validation of an empirical measure of individual differences in the degree of adherence to standards of "conventional morality." The items for this measure, which is administered to subjects as a Lifestyle Preference Questionnaire, were based on traditional codes of Western ethical conduct (such as the Ten Commandments and the "seven deadly sins"). The refined, 60-item measure of adherence to these codes was found to be reliable in terms of both its internal consistency (.91) and its stability over time (.94). The pattern of correlations with real-ideal self-image measures and other self-report measures relevant to either conventional moral behavior or conventional moral ideology provided considerable support for the discriminant-convergent validity of the Conventional Morality Scale (CMS). In addition, the validity of the CMS as a predictor of morally relevant behavior was supported by data concerning the relative "no-show" rates of high- versus low-CM subjects who had previously committed themselves to participate in a scheduled experimental session.

The evaluation of individual differences in moral beliefs and behavior has had a long and arduous history that predates by centuries the advent of modern empirical methods. This prehistory is rooted in the various philosophical and theological traditions that have attempted to reveal the "true nature" of morality and the processes involved in making appropriate ethical judgments. At one extreme, the Greek philosopher Epicurus is perhaps the best-remembered proponent of a purely hedonistic system of ethics advocating situational standards of morality in the service of self-gratification and self-interest. At the other extreme, the eighteenth-century philosopher Immanuel Kant (1724–1804) defined certain moral
The development of these individual differences in morality and ethics has been a second area of major interest to theorists and researchers. Piaget (1932) and Kohlberg (1964) have proposed influential theories suggesting that all individuals pass through specific stages of moral development that are ordered along a continuum from what can be described as amoral ignorance at one end to a refined ethical sensibility at the other. Both theories contain the assumption that moral development can be arrested at any particular stage, thus resulting in a wide range of individual differences.

Building on much of Piaget's theoretical work (e.g., 1932) regarding moral development, Kohlberg (1964) developed a measure designed to evaluate the moral values that presumably mediate the process of making ethical judgments. Kohlberg’s work is representative of initial attempts to empirically measure individual differences in the development of moral thought. Unfortunately, his measure consists entirely of hypothetical scenarios to which subjects are asked to respond by indicating what they believe would be the most ethical behavior on the part of the character(s) in each scenario and their rationales for choosing this behavior. Equally unfortunate from the standpoint of more empirically minded researchers, the subjects’ responses are then scored by panels of judges who are instructed to rate the responses of each subject according to the degree of moral attainment that each response represents in terms of Kohlberg’s predetermined stages of moral development. Whether one agrees or disagrees with Kohlberg’s conceptualization of these stages (compare Kurtines & Greif, 1974), it is clear that because this same conceptualization is inherent in the judges’ evaluations, their ratings are hopelessly confounded with Kohlberg’s a priori definition of moral judgment and development.

A methodology similar to Kohlberg’s is employed in the work of Hogan (1967, 1970), who suggests that the development of moral character can be described in terms of five dimensions: moral knowledge, socialization, empathy, autonomy, and moral judgment. In order to assess these dimensions, Hogan and Dickstein (1972) developed a measure of moral values that contains 15 statements to which subjects are asked to respond in terms of their moral judgments regarding each statement. As in Kohlberg’s methodology, the subjects’ responses are rated by judges in terms of the “maturity of moral judgment” displayed. Hogan’s measure is briefer and more readily scored than is Kohlberg’s, and his multidimensional approach to the study of mature moral values and judgments may prove to have some heuristic value. Once again, however, the emphasis is on the nature and etiology of individuals’ moral judgments as assessed by judges’ ratings rather than on individuals’ actual ethical behaviors.
CONVENTIONAL MORALITY

In all of these research traditions, the prediction of actual morally relevant behavior has typically been de-emphasized. In fact, apart from the often-controversial findings of a few early studies (see Hartshorne & May, 1928), the prediction of moral behavior based on assessments of individual differences in ethical standards has received relatively little attention (for some contemporary examples, see Schwartz, 1970; and Schwartz et al., 1969).

Similarly, the comparatively few attempts to address the relationship between moral judgment and moral behavior have yielded results that are nonconfirmatory, inconsistent, and/or difficult to interpret. For example, when Haan, Smith, and Block (1968) used the subjects' Moral Judgment Scale classification as a predictor of participation in a free speech movement "sit-in" at Berkeley, they found that fully 60% of the stage two subjects (the second highest percentage in the sample) participated in the sit-in. This finding is at odds with Kohlberg's theory regarding the moral judgments and behaviors characteristic of subjects assessed at the lower stages of his classification strategy. In a more recent study, Haan (1979) found no correspondence between level of moral judgment and student protest activity. Moreover, when Forsythe and Berger (1979) attempted to use the Ethics Position Questionnaire (EPQ) to predict behavior in a cheating situation, they found that this measure had virtually no utility in predicting individual differences in ethical (i.e., cheating) behavior.

We are not the first to note the inconsistent ties to moral behavior in the literatures on moral development and moral judgment. In a recent review, Blasi (1980) found only 11 published and unpublished studies that attempted to relate moral judgment to cheating (a prototypical, morally relevant behavior). Only six of these studies (four of which Blasi criticized on methodological grounds) yielded support for the expected positive relationships between moral judgment and resistance to temptation. Blasi's (1980) general conclusion is one with which we agree: Contradictory findings regarding the relationship between moral judgment and moral behavior, together with additional psychometric deficiencies inherent in this approach, suggest serious problems for theory and research in the moral judgment tradition.

The purpose of the present paper is to suggest a new approach to the assessment of individual differences in ethical and moral values—one that would ideally permit the prediction of relevant behavioral differences as well. Specifically, this paper concerns the construction, psychometric refinement, and construct validation of a paper-and-pencil measure that reflects degree of adherence to traditional, Western ethical rules of conduct such as the Ten Commandments. This measure, the Conventional Morality Scale (CMS), is composed of items that directly assess the degree to which an individual's self-reported behaviors reflect a high versus low level of "conventional morality." The evidence obtained to date suggests that this measure reflects a general orientation to morality that emphasizes the acceptance of traditional moral absolutes at the expense of purely hedonistic conduct.

This paper reviews the development of the Conventional Morality Scale in terms of its psychometric refinement, its construct validity, and its ability to predict individual differences in conventional ethical behavior. Finally, the application of the CMS to the study of individual differences and interpersonal behavior dynamics is discussed, and specific suggestions for its use in this regard are proposed.

CONSTRUCTION OF THE CONVENTIONAL MORALITY SCALE (CMS)

A preliminary item set of 80 self-descriptive statements was administered as a Lifestyle Preference Questionnaire to 267 undergraduate students at the University of Texas at Arlington. Nearly all of the items in this set were constructed using the "seven deadly sins" and the Ten Commandments as heuristic guidelines for generating item content. Items representative of the seven deadly sins were designed to assess: (1) lust ("I like to read erotic books or magazines"), (2) anger ("I am not the kind of person to hold a grudge"), (3) envy ("I envy people who have more than I do"), (4) sloth ("If I had enough money, I wouldn't work another day"), (5) greed ("I have a strong need to acquire things"), (6) pride ("If I accomplish something, I like to let as many people as possible know about it"), and (7) gluttony ("I indulge myself by eating large meals").

Items based on the Ten Commandments were intended to emphasize the "secular commandments" (those without overt religious connotations), such as: (1) stealing ("I have taken things I wanted without paying for them or returning them later"), and (2) respect for parents ("I tune out most of what my parents have to say to me"). Because the CMS was not intended to be a measure of religious propensities, items were not written to represent commandments of a strictly theological nature. However, one item did concern church attendance ("I attend church services at least once a week") because this type of behavior seems intuitively linked with conventional morality. This item is representative of the commandment to "keep holy the Sabbath day" and therefore does not deviate from the basic model.

In addition to the items noted above, the initial pool included a few items designed to assess more general ideological notions regarding
and relative concern or unconcern for the moral implications of one's behavior. (For example, "Morality and ethics don't really concern me"); "For me, what is right or wrong depends entirely on the situation; there are no absolute rules.")

For each of the items tested, five response alternatives were available: (1) extremely uncharacteristic of me, (2) somewhat uncharacteristic of me, (3) neither characteristic nor uncharacteristic of me, (4) somewhat characteristic of me, and (5) extremely characteristic of me. To avoid response set bias, some items were written and keyed in the direction of high conventional morality, whereas other items were written and keyed in the direction of low conventional morality.

SCALE REFINEMENT AND RELIABILITY

One serious and potentially damaging criticism of the CMS is that any such measure is likely to be hopelessly confounded with social desirability. That is, rather than accurately reporting the degree of their own behavioral endorsement of conventional moral values, respondents may answer the items in a way that they perceive to be most desirable by society at large. In order to correct for this potential confound, the original sample of 267 undergraduate students was asked to complete the Crowne-Marlowe Social Desirability Scale (M-C SDS; Crowne & Marlowe, 1964) in addition to the 80-item, unrefined version of the CMS. Items on the unrefined CMS that yielded both significant correlations with social desirability and low item-to-total correlations with the CMS were then eliminated from the CMS item pool. Through this procedure, a set of CMS items was retained that shared substantially less variance with social desirability.

Despite these precautions, we expected that significant overlap in the CMS and the SDS might still remain because, by definition, behavior that is conventional also tends to be socially desirable. To further "purify" the CMS as a measure of conventional morality, we simply parialed out any remaining variance in the CMS responses that was attributable to social desirability (as measured by the M-C SDS). We return to this issue in the discussion of convergent and divergent validity of the CMS.

Reliability assessment and additional scale refinement were accomplished by computing the internal consistency of both the original 80-item CMS and the first revised version. Item-to-total correlations were used in conjunction with Cronbach's alpha (Cronbach, 1950) to discard additional items that (1) did not demonstrate substantial item-to-total correlation, and/or (2) did not enhance the overall internal consistency of the CMS. The internal consistency of the refined version of the CMS was also assessed using this method.

An evaluation of the psychometric properties of the revised 60-item CMS yielded the following data for a second sample of 249 subjects (119 men, 130 women) whose CMS responses contained no missing values: (1) the range of scores on the revised CMS was 129 to 275 (compared to a theoretical range of 60 to 300); (2) the means were 191.43 for the men, 207.53 for the women, and 199.84 for all subjects combined; (3) the standard deviations were 25.11 for men, 29.31 for women, and 28.49 overall; (4) the average item-to-total correlation (controlling for social desirability) was .38; (5) Cronbach's alpha was .89 for the men, .91 for the women, and .91 overall; and (6) in a subsample of 32 subjects, the refined CMS yielded a two-month test-retest reliability coefficient of .94.

A further revision of the CMS was accomplished by eliminating items from the 60-item version with item-to-total correlations less than .30. Fifteen items were deleted through this procedure, resulting in a 45-item version of the CMS.1 In addition to being briefer, this version of the CMS still demonstrates very high internal consistency (alpha = .88, n = 503, mixed-sex sample). CMS items on both the 60- and 45-item versions are presented in Table 1, along with the corresponding item-to-total correlations for each item. Although not used in the studies reported in this paper, it is this 45-item version that we intend to use in future research involving the CMS.

EXPLORATIONS OF THE REVISED CMS ITEM STRUCTURE

The high internal consistency of the revised CMS suggests that it is appropriate to treat the scale as a unitary measure of adherence to conventional morality. On the other hand, the conception of conventional morality on which the scale is based is clearly multifaceted. Because different classes of behavior (lust, stealing, greed, etc.) are all reflected in the various scale items and could form relatively stable factors in the CMS, an exploratory principal components factor analysis with varimax rotation was applied to both the 60- and 45-item versions of the CMS.

An initial factor analysis (n = 249) and a subsequent cross-replication (n = 503) was performed on the 60-item CMS. Up to 19 factors yielded eigenvalues greater than 1 in this procedure, making it difficult to determine which of the principal components were important in terms of contributing meaningful and systematic variance. Examination of the scree plots for both analyses indicated that the first principle component

---

1. Two mixed-sex samples (n = 249, n = 503) were used for computing item-to-total correlations. Items correlating less than .30 using either sample were deleted from the 60-item version of the CMS to create the 45-item version.
<table>
<thead>
<tr>
<th>CMS ITEM</th>
<th>60-ITEM CMS</th>
<th>45-ITEM CMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Money is high on the list of things I consider important.</td>
<td>0.27</td>
<td>dropped</td>
</tr>
<tr>
<td>2. I like to read erotic books or magazines.</td>
<td>0.39</td>
<td>0.47</td>
</tr>
<tr>
<td>3. I am not the kind of person to hold a grudge. (+)</td>
<td>0.29</td>
<td>dropped</td>
</tr>
<tr>
<td>4. I am opposed to the use of alcohol or other recreational drugs. (+)</td>
<td>0.47</td>
<td>0.44</td>
</tr>
<tr>
<td>5. I have taken things I wanted without paying for them or returning them later</td>
<td>0.37</td>
<td>0.39</td>
</tr>
<tr>
<td>6. My behavior at parties has gotten me into trouble.</td>
<td>0.41</td>
<td>0.39</td>
</tr>
<tr>
<td>7. I don’t resent the fact that other people have a better standard of living than I do. (+)</td>
<td>0.27</td>
<td>dropped</td>
</tr>
<tr>
<td>8. I donate money to charities. (+)</td>
<td>0.31</td>
<td>0.51</td>
</tr>
<tr>
<td>9. I avoid going to social events where a lot of people will be drunk. (+)</td>
<td>0.39</td>
<td>0.46</td>
</tr>
<tr>
<td>10. I am very forgiving of others who have injured or offended me. (+)</td>
<td>0.31</td>
<td>0.37</td>
</tr>
<tr>
<td>11. If I had enough money, I wouldn’t work another day</td>
<td>0.33</td>
<td>0.35</td>
</tr>
<tr>
<td>12. I attend church services at least once a week.</td>
<td>0.44</td>
<td>0.38</td>
</tr>
<tr>
<td>13. I prefer a lifestyle that gives me an almost unlimited amount of leisure time.</td>
<td>0.42</td>
<td>0.34</td>
</tr>
<tr>
<td>14. My lifestyle preferences rule out the possibility of premarital or extramarital sexual behavior. (+)</td>
<td>0.50</td>
<td>0.32</td>
</tr>
<tr>
<td>15. I use profanity in my conversations with friends.</td>
<td>0.56</td>
<td>0.53</td>
</tr>
<tr>
<td>16. There are people in this world I would kill if I thought I could get away with it.</td>
<td>0.43</td>
<td>0.45</td>
</tr>
<tr>
<td>17. I “tune out” most of what my parents have to say to me.</td>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td>18. My pride has kept me from making up with someone I was at odds with.</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>19. I take care of myself and don’t worry too much about other people.</td>
<td>0.39</td>
<td>0.32</td>
</tr>
<tr>
<td>20. Some people get offended at the kind of language I use.</td>
<td>0.49</td>
<td>0.53</td>
</tr>
<tr>
<td>21. I believe that if something feels good and is palatable, you should do it as much as you want.</td>
<td>0.46</td>
<td>0.43</td>
</tr>
<tr>
<td>22. I have avoided people rather than having to apologize to them for something I have done.</td>
<td>0.36</td>
<td>0.38</td>
</tr>
<tr>
<td>23. I envy people who have more than I do.</td>
<td>0.34</td>
<td>0.35</td>
</tr>
</tbody>
</table>

(Table continues on page 318.)
CONVENTIONAL MORALITY

TABLE 1 (Continued)

<table>
<thead>
<tr>
<th>CMS ITEM</th>
<th>60-ITEM CMS</th>
<th>45-ITEM CMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>49. I think the best of people, even when they make a mistake. (+)</td>
<td>.28</td>
<td>dropped</td>
</tr>
<tr>
<td>50. I don’t enjoy looking at pornographic films or magazines. (+)</td>
<td>.44</td>
<td>.50</td>
</tr>
<tr>
<td>51. If I accomplish something, I like to let as many people as possible know about it.</td>
<td>.21</td>
<td>dropped</td>
</tr>
<tr>
<td>52. I am not jealous of people who are more successful than I am. (+)</td>
<td>.27</td>
<td>dropped</td>
</tr>
<tr>
<td>53. If people do something to injure or insult me, I retaliate in kind.</td>
<td>.29</td>
<td>dropped</td>
</tr>
<tr>
<td>54. At parties, I drink more than most of my friends.</td>
<td>.41</td>
<td>.45</td>
</tr>
<tr>
<td>55. I make sure that I get my share of whatever rewards are available.</td>
<td>.41</td>
<td>.47</td>
</tr>
<tr>
<td>56. I am a better person than my friends think I am. (+)</td>
<td>.26</td>
<td>dropped</td>
</tr>
<tr>
<td>57. I am careful not to dress in a sexually provocative way. (+)</td>
<td>.39</td>
<td>.37</td>
</tr>
<tr>
<td>58. I could not kill another person under any circumstances. (+)</td>
<td>.32</td>
<td>.36</td>
</tr>
<tr>
<td>59. I will not take advantage of other people, even when it’s clear that they are trying to take advantage of me. (+)</td>
<td>.48</td>
<td>.46</td>
</tr>
<tr>
<td>60. I like to control other people’s behavior as much as I can.</td>
<td>.36</td>
<td>.30</td>
</tr>
</tbody>
</table>

Note. For all correlations on the 45-item version of the CMS and for correlations with an asterisk (*), n = 503. For all other correlations, n = 249 (see footnote 1). Items marked (+) are positively keyed; all remaining items are negatively keyed.

accounted for approximately 25% of the total factor variance, and that the remaining components contributed minimally in accounting for additional variance. In other words, aside from the first principal component, the scree plot was virtually "flat" with respect to the other components, whose eigenvalues were uniformly low and psychometrically "unimportant." The same scree pattern and comparable variance estimates were obtained for the 45-item version of the CMS (n = 503).

Because all but the first principal component in these factor analyses yielded relatively low eigenvalues and accounted for relatively low percentages of the total factor variance, a multifactor interpretation of the CMS is not a satisfactory one. Given the extremely high internal consistency of the items and the lack of more than one major factor, most psychometrists would readily interpret the CMS as a unitary measure of a single, underlying construct—a construct that we have labeled "adherence to conventional morality." Clearly, the high internal consistency of a set of items tapping such disparate classes of behavior could be explained only in terms of a broader disposition to respond to all of these classes in a similar way.

CONVERGENT AND DISCRIMINANT VALIDITY OF THE CMS

Having established that the CMS provides a reliable, internally consistent measure of a single, unitary disposition that organizes subjects' perceptions and reported actions with respect to a wide range of morally relevant behaviors, we sought next to explore the convergent and discriminant validity of the CMS (Campbell & Fiske, 1959). Accordingly, we asked a sample of 73 subjects (34 men and 39 women) to complete the CMS along with a number of relevant but conceptually distinct personality measures. The resulting correlations of the CMS with these other measures are summarized in Table 2.

Relationship with Social Desirability. As mentioned earlier, the relationship between the CMS and the Crowne-Marlowe Social Desirability Scale is of special theoretical interest. On the one hand, the convergent validity of the CMS would be supported by its significant positive correlation with social desirability, since one's professed adherence to conventional moral behavior constitutes socially desirable responding almost by definition. On the other hand, if such a relationship were established it would then be desirable to partial out the variance common to both measures in order to obtain "purified" CMS scores that are relatively

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Conventional Morality Scale (CMS) Convergent/Divergent Validity Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MALES</td>
</tr>
<tr>
<td>Social desirability</td>
<td>.08</td>
</tr>
<tr>
<td>Disinhibition</td>
<td>-.38**</td>
</tr>
<tr>
<td>General (SSS)</td>
<td>-.16</td>
</tr>
<tr>
<td>Mehrabian empathy</td>
<td>.25**</td>
</tr>
<tr>
<td>MMPI Pd subscale</td>
<td>-.10</td>
</tr>
</tbody>
</table>

Note. Numbers in parentheses are correlations with the refined, 60-item version of the CMS after statistically correcting for social desirability. Numbers in brackets are social desirability correlations with the unrevised, 80-item version of the CMS, n (males) = 34, n (females) = 39, n (total) = 73.

*p < .05
**p < .01
***p < .001.
CONVENTIONAL MORALITY

free of socially desirable response set bias. For these reasons, the Crowne-Marlowe Social Desirability Scale was expected to serve a dual function in the convergent validation and the subsequent conceptual refinement of the CMS.

As the data in brackets in the top row of Table 2 indicate, we found the predicted positive correlation between social desirability and the unrefined, 80-item version of the CMS for the subjects in our sample, both for each sex separately and for both sexes combined. Moreover, the data outside the brackets indicate that our subsequent attempts to minimize this overlapping variance in the refined 60-item CMS were largely successful, since none of the correlations between social desirability and the refined CMS were statistically significant. This outcome was expected given that one of the criteria for deleting items in the revision process was significant correlation with social desirability.

Relation with Disinhibition. Evidence for the convergent validity of the refined, 60-item CMS was provided by its strong negative correlation with the disinhibition subscale of Zuckerman’s (1979) Sensation Seeking Scale (Table 1, second row). The disinhibition subscale of the SSS assesses individual differences in the desire to engage in disinhibited behavior (e.g., “loose” sexual behavior, use of drugs and alcohol, gambling) as one aspect of a general desire for environmental stimulation. Because free validity considerations suggest that this disinhibition measure should be related to the CMS, we had expected that there would be a significant negative correlation between CMS and disinhibition scores.

In contrast, the general subscale of the SSS, which is interpreted as a very broad measure of individual differences in need for environmental stimulation, would not seem to be related to a measure of conventional morality. We therefore predicted that there would be no significant correlations between this SSS subscale and the CMS. The nonsignificant correlations in the third row of Table 2 substantiate this prediction and provide discriminant validity data to complement the convergent validity data already presented.

Relation with Empathy. Because of conventional morality’s emphasis on concern for the welfare of others (e.g., the Golden Rule), we predicted that subjects’ CMS scores would correlate positively with their scores on Mehrabian and Epstein’s (1972) measure of “empathic tendency.” The data in the fourth row of Table 2 reveal that this prediction was confirmed for the men in the sample but not for the women. This apparent difference may be due to a range restriction artifact, since the range of the women’s empathy scores (76) was considerably smaller than that of the men’s (121). It is also possible that the generally weak correlations obtained here reflect a more theoretically interesting influence, that is to the extent subjects who score high on the CMS are self-righteous or intolerant of others who do not share their ethical standards, the predicted positive relationship between CMS score and “empathic tendency” may be attenuated.

Relation with Psychopathic Deviance. Arguing from a social deviance perspective, one of the authors predicted a negative relationship between the subjects’ CMS scores and their scores on the psychopathic deviate (Pd) subscale of the Minnesota Multiphasic Personality Inventory. Arguing that psychopathology is equally probable in individuals who vary in terms of their endorsement of conventional ethical behavior, the other author predicted that no significant correlation would be obtained. As the data in the last row of Table 2 indicate, the first prediction accurately characterized the women in the sample, whereas the second prediction characterized the men. Because this gender difference does not appear to be an artifact of range restriction, we speculate that it reflects the effects of the proverbial “double standard of morality” that assigns women a more socially deviant status than men (and thus fosters their self-identification as socially deviant) when they engage in the same disapproved behaviors.

Relation with Moral “Idealism” versus “Relativism”. A further test of the convergent and discriminant validity of the CMS took advantage of the fact that Forsythe’s (1960) Ethics Position Questionnaire (EPQ) contains relatively orthogonal subscales that respectively measure the degree to which individuals endorse “idealist” or “relativist” conceptions of morality. We predicted that high CM subjects would score significantly higher on the “idealist” subscale of the EPQ than on the “relativist” subscale because of their professed adherence to a moral code (i.e., that based on the Ten Commandments and the seven deadly sins) that has traditionally been regarded as absolute and allowing for few, if any, exceptions. In contrast, we predicted that low CM subjects would either (1) score significantly higher on the “relativist” subscale than on the “idealist” subscale, or (2) reveal no significant differences in their “idealist” and “relativist” subscale scores.

Because the predicted results could be expressed in the form of a statistical interaction, the data were entered into a 2 × 2 (analysis of variance model that treated the subjects’ CM classification (high vs. low, by median split within gender) as a between-subjects variable and treated their EPQ subscale scores as a within-subjects variable. The results of this analysis, depicted in Figure 1, revealed that the predicted interaction

2. Low CMS scores would not necessarily imply that these subjects have no ethical standards—only that their standards are less conventional than those of high CMS subjects. Therefore, the more conservative prediction of no differences between EPQ scale scores for low CM subjects may be the most appropriate.
was highly significant, \( F(1, 71) = 14.74, p < .001 \). In addition, the relevant mean comparisons revealed that the high CM subjects scored significantly higher on the “idealist” subscale than on the “relativist” subscale of the EPQ, \( t(36) = 4.94, p < .001 \), whereas the low CM subjects showed no significant difference in their endorsement of these measures, \( t < 1 \).

These results clearly confirm our prediction that high CM subjects more strongly endorse an idealistic orientation to moral behavior than a relativistic one. In contrast, low CM subjects do not differentially endorse moral idealism and moral relativism, appearing to subscribe to both orientations in roughly the same degree.

Relationship with real–ideal self ratings. The next step in our construct validation procedure was to test for expected differences in the self-concepts of subjects reporting high versus low levels of conventionally moral behavior. Following the precedent established by various “self” theorists (e.g., Duval & Wicklund, 1972; Gergen, 1971; Rogers, 1970; Wylie, 1974), we assumed that dimensions of the self-concept could be studied in terms of subjects’ perceptions of (1) the actual or “real” self, (2) the “ideal” self, and (3) the ideal–real self discrepancy. Given these distinctions, our conception of conventional morality led us to predict that, for morally relevant dimensions of self, both the real and the ideal self ratings of high CM subjects would be higher than those of low CM subjects. This prediction is based on the assumptions that high CM subjects do in fact display higher levels of conventionally moral behavior than low CM subjects, and that their motivation to do this derives at least in part from the higher ideals possessed by high CM subjects regarding morally relevant behavior. On the other hand, our conception of conventional morality provides no compelling basis for predicting differences in the real or ideal self-ratings of high versus low CM subjects on morally irrelevant dimensions of self. Accordingly, we designed our study of the self-concepts of high versus low CM subjects so that we could test the prediction that the moral relevance of the self-rated dimensions would interact with the subjects’ CM status (high versus low) to determine their real and ideal self-ratings.

SUBJECTS AND PROCEDURE

The subjects were 30 male and 27 female undergraduates enrolled in introductory psychology classes at the University of Texas at Arlington. These 57 subjects, in a classroom setting, completed the Conventional Morality Scale and then recorded their responses to the items on a real-ideal self-inventory. This second measure has been used in previous research (Ickes, Wicklund, & Ferris, 1973) and consists of 20 bipolar adjective pairs, with each pair defining the end-points of a 20-point horizontal scale. The adjective pairs were as follows:

1. Courteous–Rude
2. Skilled–Unskilled
3. Competent–Incompetent
4. Pleasant–Disagreeable
5. Creative–Unimaginative
6. Honest–Dishonest
7. Careful–Reckless
8. Trustworthy–Untrustworthy
9. Kind–Cruel
10. Independent–Dependent
11. Courageous–Cowardly
12. Generous–Selfish
13. Tolerant–Intolerant
14. Considerate–Inconsiderate
15. Successful–Unsuccessful
16. Well liked–Disliked
17. Industrious–Lazy
18. Sensitive–Insensitive
19. Optimistic–Pessimistic
20. Intelligent–Stupid

The written instructions indicated that, for each of the dimensions defined by the 20 adjective pairs, the subjects were to circle the point on the scale that best described their own behavior.
that represented their real self and the point that represented their ideal self. This procedure yielded a real-self score, an ideal-self score, and a real-self–ideal-self discrepancy score (ideal score minus real score) for each subject on each adjective pair.

In order to test our predictions regarding morally relevant versus irrelevant dimensions of self, we averaged the subjects' self-ratings on adjective pairs 1, 4, 6, 7, 8, 9, 12, 13, 14, and 17 to create indexes of morally relevant self-ratings, and we averaged their self-ratings on adjective pairs 2, 3, 5, 10, 11, 15, 16, 18, 19, and 20 to create indexes of morally irrelevant self-ratings. As one can see from an examination of the items, the morally relevant adjective pairs reflect different types of morally relevant behavior, whereas the morally irrelevant adjective pairs reflect behaviors that are generally unrelated to conventional morality. The internal consistency of these indexes as measured by Cronbach's alpha (Cronbach, 1951) was quite high (morally relevant real-self index, alpha = .86; morally irrelevant real-self index, alpha = .83; morally relevant ideal-self index, alpha = .83; morally irrelevant ideal-self index, alpha = .89).

The data in this study consisted of the four indexes described above as well as two additional indexes that represented the subjects' average discrepancy scores on each of the two sets of items (morally relevant vs. morally irrelevant). For our purposes, discrepancy scores were defined as the absolute difference between the real and ideal scores for each adjective pair, since there is evidence suggesting that absolute difference scores are more reliable than are other types of discrepancy score measures (Hoge & McCarthy, 1983). The internal consistency of these discrepancy scores was also quite high (morally relevant discrepancy scores, alpha = .85; morally irrelevant discrepancy scores, alpha = .84).

The data were analyzed in a $2 \times (2) \times (2)$ repeated measures factorial design that treated conventional morality classification (high vs. low, based on median split) as the single between-subjects variable and treated the type and moral relevancy of the self-rating index (real-self vs. ideal-self, morally relevant vs. morally irrelevant) as the two within-subjects variables. The discrepancy score indexes were analyzed separately because of their dependency on the real- and ideal-self rating data from which they were derived.

3. Confirmatory factor analysis results indicated that this grouping of items accounted for more variance than two random grouping models tested.
4. A previous analysis of these data including the subjects' gender as an independent variable yielded no significant effects for the gender variable, either as a main effect or in interaction with the other variables. For this reason, the findings reported here are based on an analysis of the data collapsed over gender.
their real-self ratings were significantly higher for morally relevant dimensions of self than for morally irrelevant dimensions of self, \( t (28) = 4.30, p < .001 \). Although not specifically predicted, this finding is consistent with the logic underlying our predictions. It also provides additional evidence for the convergent/divergent validity of the CMS in that CM differences were found for morally relevant self-ratings (convergent validity) but not for morally irrelevant ones (divergent validity).

In addition to the findings just reported, there was an expected main effect for the type of self-rating (real vs. ideal) indicating that both high and low CM subjects had significantly higher ideal-self scores than real-self scores on both morally relevant and irrelevant dimensions of self, \( F (1, 55) = 137.15, p < .001 \). There were no differences in the discrepancy scores for high and low CM subjects, \( t (55) < 1 \).

Taken together, these data support our predictions that high CM subjects would report higher real- and ideal-self ratings than would low CM subjects, but only for morally relevant dimensions of self. The fact that there were no differences in the ideal-real discrepancy scores of high and low CM subjects suggests that both types of subjects fall about equally short of attaining their ideals; the difference lies in the higher moral aspirations and real-self appraisals of high versus low CM subjects. If high CM subjects do indeed display more conventionally moral behavior than low CM subjects, this difference may be attributable in part to the motivating influence of their higher ideals on morally relevant dimensions of self.

**PREDICTIVE VALIDITY OF THE CMS**

An important part of the construct validation of a personality measure is establishing its predictive validity. In this section, evidence is provided for the predictive validity of the CMS in terms of a morally relevant behavior: meeting versus failing to meet an obligation to which one has committed oneself.

A common, and frustrating, problem in psychological research using human subjects is the ever-present possibility that subjects who have agreed to participate in an experiment may not show up. Although there are circumstances in which such behavior is justified, there are many others in which subjects are absent for no apparent reason. In such cases of unjustified absence, the commitment that the subject made to arrive at the appointed time and place of the experiment has not been honored. In other words, "no-show" subjects who do not attempt to contact the experimenter and who offer no excuses for their absence have violated their commitment to participate in the experiment. This reasoning sug-

gests that, to the extent that the CMS taps an individual's level of social responsibility and honesty, scores on this measure should predict subjects' "show" or "no-show" status in experimental research.

**METHOD**

As part of a larger study of dyadic interaction to be reported in a subsequent paper (Tooke & Ickes, in preparation), subjects were selected on the basis of extreme high and extreme low CMS pretest scores such that high-high, low-low, and high-low dyads could be formed. Subjects were then contacted by phone and given the opportunity to participate in the study at a time of their choosing. Potential subjects gave the research assistant who did the scheduling verbal commitments that they would appear when scheduled and participate in the experiment. Because an equal number of high and low CM subjects were ultimately run in the study, it is possible to meaningfully compare the "no-show" rates among high versus low CM subjects. For our purposes, "no-shows" were operationally defined as any subjects who (1) failed to keep their experimental appointment and (2) failed to contact the experimenter at any time to notify him of the reason(s) for the absence. Based on the assumed differences between low and high CM subjects in their concern for others or for moral obligations in general, we predicted that the "no-show" rate would be significantly greater among low CM subjects than among high CM subjects.

**RESULTS AND DISCUSSION**

In order to test this prediction, the number of "shows" and "no-shows" by CM classification was entered into a \( 2 \times 2 \) chi-square analysis. As shown in Table 3, this procedure revealed a significant difference, \( X (1) = 8.54, p = .003 \), indicating that low CM subjects were about three times more likely to be "no-shows" than were high CM subjects.\(^5\)

In order to rule out the possibility that any findings from this study might be an artifact of the correlation between CM and social desirability, separate regression equations were formed for the men and the women that were used to predict CMS scores from scores on the Crown-Marlowe Social Desirability Scale (1964). These predicted CMS scores were then subtracted from the subjects' actual CMS scores to yield a final score

\(^5\) These findings are also based on an analysis of the data collapsed over the subjects' gender—a variable that again yielded no significant effects.
that was empirically uncorrelated ($r = .00$) with social desirability (i.e., a “purified” CMS score). Because it was on the basis of these “purified” CMS scores that subjects were selected and scheduled for the dyadic interaction study mentioned above, it is clear that social desirability was not responsible for the difference in the subject “no-show” rates. In fact, in this case, the correlation between the raw CMS scores and “no-show” status was identical to the correlation between the “purified” CMS scores and “no-show” status ($r = -.24, p < .001$), indicating that social desirability per se had essentially no effect on this behavior.

These data indicate that low CM subjects are indeed less likely than high CM subjects to follow through on the moral obligations they incur. Although a person’s tendency to honor commitments is not made explicit in any of the CMS items, there are some items that appear to be relevant to this behavior and can be useful in predicting a person’s tendency to follow through on commitments that have been freely made (e.g., “I put other people’s needs and preferences ahead of my own,” “I prefer a lifestyle that gives me an almost unlimited amount of leisure time”).

In summary, the differential rates of experimental “no-shows” among low versus high CM subjects adds evidence for the predictive validity of the CMS. An additional, meta-theoretical implication of these data is that low CM subjects may be consistently underrepresented in psychological research requiring the voluntary presence of human subjects. To the extent that low CM individuals selectively fail to report to such studies, the findings of these studies may be limited in their generalizability.

**SOCIODEMOGRAPHIC PREDICTORS OF CMS SCORES**

The data reviewed so far in this paper provide considerable evidence for the reliability and construct validity of the CMS, but they tell us little, if anything, about the source(s) of individual differences on this dimension.

In a preliminary attempt to address this issue, we included in our two large-sample administrations of the CMS a number of questions regarding various sociodemographic variables (ethnicity, number of older and younger siblings, parents’ marital status, etc.) that might serve as background predictors of the subjects’ CMS scores.

Although we have yet to test one of the more intuitively plausible predictors—socioeconomic status—we have found that CMS scores are not significantly associated in our student population with the subjects’ race/ethnicity, number of older and younger siblings (a rough proxy of family size), or the marital status (married vs. divorced) of their parents. In fact, only two significant background predictors—gender and death of parent(s)—have emerged to date.

The data for the first predictor—the subjects’ gender—revealed that the mean of women’s CMS scores was significantly greater ($M = 211.84$), than that of men’s ($M = 195.75$), $t (495) = -6.83, p < .001$. A likely interpretation of this gender difference is that societal expectations require that women display greater adherence than men to the tenets of conventional morality (i.e., the so-called double standard). This explanation goes on to assume the differential socialization of men and women with regard to morally relevant behavior.

The data for the second predictor—the loss (in male subjects only) of one or both parents through death—revealed that the mean CMS score of men with one or both parents deceased was significantly lower ($M = 174.28$), than that of men whose parents were still living ($M = 198.79$), $F (1, 197) = 14.31, p < .001$. Further analyses revealed that the gender of the deceased parent did not moderate this effect. Socialization differences are also likely to account for this finding, although specifying exactly which differences are important (e.g., the long-term effects of unstable attachment and anadic depression vs. the importance of parental example and instruction) may require considerable research effort.

**SUMMARY AND IMPLICATIONS**

Our goals in this research were to construct and empirically validate a measure of adherence to conventional moral standards regarding behaviors such as those prescribed by the Ten Commandments and proscribed by the seven deadly sins. The original 80-item version of the Conventional Morality Scale (CMS) was psychometrically evaluated and refined such that an internally consistent (alpha = .91) and temporally stable (test-retest reliability $= .94$) 60-item version of the measure was obtained. Further revision based on item-to-total correlations yielded a 45-item version of the CMS with comparable internal consistency (alpha $= .88$) that will be used in future research. An exploratory factor analysis with
varimax rotation and a subsequent replication of this procedure with a much larger sample suggested that the CMS is most appropriately interpreted as a unitary measure of adherence to conventional morality.

What is the nature of this unidimensional construct we have defined as conventional morality? The data from the EPQ study suggest that the CMS measures an idealist conception of morality as opposed to a relativist one. Complementing these data, the data from our real-ideal self image study indicate that the self-ratings of high CM subjects and low CM subjects differ only on morally relevant dimensions of self. Taken together, both sets of data triangulate with the discriminant/convergent correlational data in “defining” the dimension underlying the CMS as adherence to conventional morality.

Finally, data were presented that provide some preliminary evidence for the behavioral or predictive validity of the CMS. Specifically, it was demonstrated that low scorers on the CMS were significantly less likely than high CMS scorers to honor a commitment they had freely made (i.e., keeping an appointment to participate in a psychology experiment). This finding raises the important question of whether low CM subjects are generally underrepresented in psychological research requiring the voluntary participation of human subjects.

ISSUES FOR FUTURE RESEARCH

The internal consistency of the 60-item CMS (alpha = .91) is remarkably high considering that its item content intentionally represents several behavioral aspects of conventional morality. This measure of interitem reliability reflects a high degree of interrelatedness in the subjects’ reactions to the various classes of behavior that, taken together, define conventional moral behavior in Western culture. We suggest that to fully understand the personality dimension of adherence to conventional morality, it is necessary to investigate a wide variety of behaviors that are relevant to conventional morality rather than limiting oneself to a small subset of these behaviors. Investigations of this sort should be an important, if demanding, goal of future research.

The motivational implications of differences in the real-ideal self-ratings of high and low CM subjects on morally relevant dimensions also warrant further study. Are the higher moral ideals of high CM subjects responsible, at least in part, for the higher real or actual levels of moral behavior that they report? Does the similar average discrepancy between the real and ideal self-ratings of high and low CM subjects on morally relevant dimensions imply that there is an optimal amount of discrepancy that all subjects experience as psychologically tolerable? If so, do real moral behaviors track the changes in moral ideals, and vice versa, whenever this optimal discrepancy is exceeded? Can either short-term laboratory inductions or longer-term rehabilitative interventions potentiate such effects? All of these questions embody important issues for additional research.

With regard to the etiology of the conventional moral personality, our preliminary investigation of some sociodemographic predictors of CMS scores revealed that (1) men scored significantly lower than women, and (2) men with one or both parents deceased scored significantly lower than men whose parents were both still living. Although these findings clearly indicate the importance of socialization processes in the development of the conventional moral personality, further research is needed to evaluate the relative importance and potential interaction of specific socialization influences.

Beyond these research directions suggested by our current studies of the reliability and validity of the CMS, there are other areas of social psychological research to which the construct of conventional morality is relevant. For example, traditional research in such areas as help-giving (e.g., Berscheid & Walster, 1967; Carlsmit & Gross, 1969; Freedman et al., 1967; McMillan, 1971) and equity behavior (e.g., Austin & Walster, 1974; Simmons & Lerner, 1968; Stoklos & Schopler, 1973; Walster, Berscheid, & Walster, 1973) could be further illuminated by an analysis of personality differences in conventional moral values and the impact of these differences on the tendencies of individuals to behave altruistically or equitably in specific situations. Use of this individual difference strategy would allow researchers to explore the role of conventional morality as a moderating variable in various morally relevant situations (see Snyder & Ickes, 1985).

Finally, the influence of individual differences in conventional morality on interpersonal behavior is another appropriate direction for future research. For example, the social comparison processes of subjects who are either similar or dissimilar in terms of conventional morality scores could be investigated and the effects of these processes examined and compared to those found in previous studies in this area (e.g., Morse & Gergen, 1970). Other investigations might focus on the interaction behavior of different combinations of CMS scorers (i.e., high-high, low-low, and high-low). In a current study (Tooke & Ickes, in preparation), we are exploring both the overt social behavior and the covert thoughts and feelings of high and low CM subjects who have been systematically paired to create these different dyad types (for a description of the research paradigm used, see Ickes, 1982; Ickes & Tooke, 1988; and Ickes et al., 1986).

In summary, the various data reviewed in this paper suggest that the Conventional Morality Scale is a valid and reliable measure of individual differences in moral behavior and self-concept. Because of its emphasis
on conventionally moral behavior (as opposed to moral judgment or philosophical orientation per se), we expect that the CMS will have particular utility in studies that investigate the behavioral manifestations of internalized moral standards.

REFERENCES


Tookey, W. S., & Ickes, W. J. Conventional morality, interpersonal behavior and naturalistic social cognition (in preparation).


