A virtues approach to personality ☆

Michael J. Cawley III, James E. Martin, John A. Johnson*

Department of Psychology, The Pennsylvania State University, University Park, PA, USA

Abstract

The structure of virtue was investigated through the development and construct validation of the Virtues Scale (VS), a 140-item self-report measure of virtues. A factor analysis of responses from 390 participants revealed four factors: Empathy, Order, Resourcefulness, and Serenity. Four virtue subscales constructed from the highest loading items on each factor were correlated with the Revised NEO Personality Inventory (NEO PI-R) scales in two additional samples ($ns = 181$ and 143). One of these samples also completed the DIT measure of Kohlbergian moral development. Meaningful, replicated correlations between the virtue subscales and personality scales and complete lack of relationships between the virtues scales and the DIT indicate that virtue is a function of personality rather than moral reasoning and cognitive development. © 2000 Elsevier Science Ltd. All rights reserved.

1. Introduction

The concept of virtue has been a powerful explanatory term in moral philosophy (MacIntyre, 1981). Virtue derives from the Greek concept *areté* (αρετη), which means "excellence". To the Greeks, a person of *areté* was someone who effectively used all of his or her faculties to achieve real results (A non-peer reviewed internet reference, Hooker, 1996). *Areté* means possessing the life skills we need to achieve our highest human potentials. In contrast to duty ethics, which answers the question, "What ought I to do?" by stating a rule

one is obligated to follow, virtue ethics answers the question, "What kind of person ought I be?" by recommending a virtue worth cultivating (e.g., "be patient").

Personality psychology, given its concern with human potential, motivation, and will, can be regarded as a modern, scientific outgrowth of moral philosophy (Averill, 1980, 1997). Nonetheless, with rare exceptions (e.g., Bertocci & Millard, 1963), the virtues have been largely ignored in modern personality theory. An examination of recent texts in personality psychology (Burger, 1993; Engler, 1991; Liebert & Spiegler, 1990; and McAdams, 1990) revealed that none of the texts included the term "virtue" in their index. Partly to establish its scientific validity, the field has moved away from the normative concept of character towards the morally neutral concept of personality (Allport, 1937; Conley, 1985).

This study sought to reintroduce the question of virtue to modern psychological discussion. The development of a Virtues Scale (VS) provides a tool for understanding the structure of virtue and its relationship to psychological concepts such as personality and moral development. This study describes a self-report rating scale of individuals' self-evaluations of virtues. The Virtues Scale is evaluated and interpreted in terms of reliability and correlations with measures of personality and moral development.

While the term "virtue" may not appear in personality texts, personality trait models such as the Five Factor Model (FFM; Costa & McCrae, 1992) include trait terms (e.g., agreeable, conscientious, honest, modest) which could be understood as virtues. The Virtues Scale was correlated with the Costa and McCrae (1992) NEO PI-R scale in order to explore the relationship between virtue and personality.

Recent discussion in moral philosophy has restored virtue to its central place in our understanding of human practice, ethics, and experience. MacIntyre (1981) raises a number of philosophical considerations that may be useful for a psychological examination of the concept of virtue. He argues that the importance of virtue has been de-emphasized by those modern moral theorists who emphasize duty and principle over virtue. He states that a coherent moral theory requires both an ethic of virtue and an ethic of principle. He also argues that as cultures change new conceptions of the virtues emerge. Thus he makes two arguments that are central to this project. He both makes a strong case for the importance of the concept of virtue, and then also points out the current conceptual confusion regarding the catalogue of virtues and their relation to other moral concepts such as duty. This study addresses these concerns through a psychological examination of the concept of virtue and the development of a Virtues Scale.

While much of modern moral developmental theory (Kurtines & Gewirtz, 1984, 1995; Lickona, 1976; Rest, 1979, 1986; etc.) has followed Kohlberg (1976, 1981, 1984, etc.) and avoided the virtues, recent philosophical (MacIntyre, 1981; Meilaender, 1984; etc.) and psychological critics (Kilpatrick, 1992; Vitz, 1990; etc.) of Kohlberg have renewed interest in a psychology of virtue. Kohlberg's (1976, 1981, 1984, etc.) model is based predominantly upon a Kantian ethic of duty. Kohlberg (1981) ignores the ethic of virtue, dismissing it as the "bag of virtues" (p.2) approach. His critics (Kilpatrick, 1992; MacIntyre, 1981; Meilaender 1984; Vitz, 1990; etc.) argue that both an ethic of duty and an ethic of virtue are necessary for understanding moral development. Kohlberg's critics have suggested that his model of moral development overemphasizes the reasoning component of moral development to the neglect of virtue and upbringing. Lewis (1947) makes this same point in his comment that:

I had sooner play cards against a man who was quite skeptical about ethics, but bred to believe that 'a gentleman does not cheat', than against an irreproachable moral philosopher who had been brought up among sharpers. (p.34)

In this context, the Virtues Scale would provide a psychological measure of the ethic of virtue in contrast to the Kohlbergian tradition in developmental psychology.

It seems that for a time psychology abandoned a deeper and richer conception of the human person by turning from character to personality. The character, or virtues approach, to the human person addressed issues regarding human agency and personal responsibility. According to this approach, each person is responsible for the development of his or her character through the habitual practice of the virtues. Moral events, or dilemmas, are not seen as unrelated, isolated episodes but instead are seen as parts of a continuous story in which one tests and cultivates ones character. In contrast, the personality approach seems to provide a snapshot of an individual who is told that "there are no right or wrong answers" as they begin their surveys.

Three major arguments have been used to discredit the virtues approach in personality psychology: the definitional argument, the famous Hartshorne and May studies, and concerns about social desirability.

The definitional argument can be traced back to Allport (1937), who while recognizing the historical importance of past work on character, then defined it out of his classic definition of personality. Indeed, Allport may be seen as the decisive moment for psychology in the turn away from virtue. He notes that many writers view character as:

...the aspect of personality that engenders stability and dependability, that is responsible for sustained effort in the face of obstacles, or works for remote ends rather than those that are nearer in time but of less worth. (p.51)

Allport goes on to suggest that character is a term which is more relevant to ethics than to psychology. He does not seem to consider the view that it may be relevant to both disciplines. He concludes that:

Character is personality evaluated, and personality is character devaluated. Since character is an unnecessary concept for psychology, the term will not appear again in this volume, excepting in quotations from other writers, or in a clear historical context. (p.52)

Thus, when Allport and Odbert (1936) developed their comprehensive catalog of 18,000 human traits from the unabridged dictionary, they deliberately excluded evaluative trait terms. Thus, by a priori definition, evaluative terms, character, and virtue were left out of important early models of personality. Nonetheless, evaluative terms such as Agreeableness and Conscientiousness have found their way back into the modern FFM of personality — despite its origin in Allport and Odbert's work (John, Angleitner, & Ostendorf, 1988).

The famous Hartshorne and May studies (1928; Hartshorne, May, & Maller, 1929; and Hartshorne, May, & Shuttleworth, 1930) of character education and children are often cited as presenting evidence against the virtues approach in psychology. Hartshorne and May concluded that their results did not support the position that character traits are consistent

across all settings. However, several reanalyses (Maller, 1934; Burton, 1963; Rushton, 1984) of the Hartshorne and May studies make the argument that these studies actually support the virtues approach in psychology. Hartshorne and May (1928) actually found that certain high morale schools and teachers did produce students who behaved better, and that teacher ratings of trustworthiness did correlate with behavioral measures of honesty. Vitz (1990) argues that "Hartshorne and May's studies taken at face value are consistent with a general honesty factor as part of personality or character, but with specific and situational qualifications" (p.717).

Loevinger (1985) noted that their third volume of studies actually concluded that there is a high correlation between honest behavior and character integration. Upon reanalyzing the results of the Hartshorne and May studies, Rushton (1984) pointed out the high correlations, typically of the order of 0.50 and 0.60, between teacher ratings of children's honesty and the more reliable Hartshorne and May measures based on combining behavioral tests. Rushton thus showed that Hartshorne and May's results have been consistently misrepresented in the psychological literature. Rushton concluded that:

> ...not only did total scores within the battery of altruism tests and measures yield evidence of consistency, but so too did measures of self-control, persistence, honesty, and moral knowledge. Indeed there was evidence for a pervasive general factor of moral character. (p.273)

Concerns about social desirability have sometimes been raised to discredit the virtues approach in personality psychology. Since virtues are traits that are viewed as desirable by much of society, a measure of virtues might be seen as vulnerable to a social desirability response bias. However, in their review of the social desirability literature, McCrae and Costa (1983) noted that:

> Many psychologists still regard correlations with social desirability (SD) scales as evidence of the invalidity of measures, despite 20 years of research showing that this interpretation is usually unjustified. (p. 882)

They argued that:

> social desirability scales are better interpreted as measures of substantive traits than as indicators of response bias and that they are of little use as suppressor variables in correcting scores from other scales. (p. 882)

Thus it seems that social desirability scales measure more substance than style (see also Block, 1965; Nicholson & Hogan, 1990). Johnson (Johnson, 1990; Johnson & Horner, 1990) found that even responses to socially desirable items that are unlikely to be literally true show convergent and discriminant validity.

In this study, Costa and McCrae's (1992) five-factor model (FFM) of personality as measured by their personality inventory, the NEO PI-R, was used as a representative model and measure of personality for two reasons. First, the FFM has become the dominant model of recent personality trait research (Wiggins & Trapnell, 1997), and second, the Costa and McCrae personality model as measured by the NEO PI-R has been shown to be the most

successful measure of the five-factor model in recent personality trait research (Johnson, in press).

Personality psychology has largely followed Allport's (1937) definition of personality, which excludes evaluative traits such as virtues. However, the five-factor model includes evaluative dimensions such as Agreeableness and Conscientiousness. Thus, it seems that these theorists must now either redefine personality so as to allow evaluative terms (including concepts of character and virtue) or recognize that the five-factor models conflate personality with character and virtue.

Concerning the structure of virtue, the hypothesis of this study was that the factor analysis of the 140 virtue items would yield a reasonable and meaningful factor structure. It was anticipated that the number of virtue factors would be between three and nine (Goldberg & Digman, 1994). A factor structure result less than three or greater than nine would diminish the usefulness of the virtues as explanatory concepts.

Concerning the relationship between virtue and personality, the hypothesis of this study was that the virtue factors derived from the factor analysis would not correlate so highly (above 0.80) with the personality scales of the NEO PI-R so as to be subsumed by them. It was anticipated that there would be significant correlations between the virtue factors and personality measures of similar constructs. It was also anticipated that the virtue factors would not correlate with personality measures of different or unrelated constructs.

In summary then, this study seeks to reintroduce the concept of virtue into modern psychology through the development and validation of the Virtues Scale, which is contrasted with non-virtue measures of personality and moral development. The factor analytic data demonstrate that the Virtues Scale can be utilized to explore the notion of virtue as a psychological concept. The correlational data demonstrate that the Virtues Scale can be used to explore theoretical issues concerning the interrelationships between virtue, personality and moral development. In addition, these data support the view that the concept of virtue is not merely a subsidiary component or redundant version of personality and/or moral development, but that the notion of virtue may stand alone as an important psychological concept in its own right.

2. Method

2.1. Participants

Participants in this research were undergraduate students from psychology courses at the Pennsylvania State University. Phase I involved 390 students (163 male and 227 female). Phase II included 181 students (72 male and 109 female). Phase III included 143 students (45 male and 98 female).

2.2. Measures

2.2.1. The Virtues Scale (VS)
The construction, development and validation of the Virtues Scale are discussed in the

dissertation of the first author (Cawley, 1997). The test construction approach followed the lexical tradition described by Brokken (1978) and Angleitner, Ostendorf, and John (1990). These lexical researchers used heuristic criterion sentences to determine whether an adjective found in a dictionary was personality relevant. For Brokken (1978, p.17), an adjective was considered a personality descriptive adjective if the following sentence made sense when the ellipsis was replaced by the adjective: "He (She) is ... by nature." Angleitner et al. (1990, p.94) used the following heuristic criterion sentence to judge whether an adjective was personality relevant: "How [adjective] am I?" In the present study, the first author examined *The New Merriam-Webster Dictionary* (1989) for virtue terms, using the following two heuristic criterion sentences: "What ought I to be?" or "What ought I to do?" (Mayo, 1958). The term was recorded as a virtue if it expressed the concept of *areté* when replacing the ellipsis in one of the respective responses to the two questions: "Be..." or "Show...". Thus, *patient* was recorded as an adjectival form of virtue because it fitted the sentence, "Be patient", while *patience* was recorded as a noun form of virtue because it fitted the sentence, "Show patience".

This procedure yielded 140 unique virtue terms (adjectival and noun forms were considered to reflect the same virtue). These 140 virtue terms were developed into 140 Virtues Scale items by creating three-sentence clusters containing words and phrases from each term's definitions in *The New Merriam-Webster Dictionary* (1989) and *The Random House Dictionary* (1980) and also its entries in *The Merriam-Webster Thesaurus* (1989) and *Roget's Pocket Thesaurus* (1980). The decision to use sentence clusters rather than single virtue terms followed an admonition from Goldberg (Goldberg, 1999; Goldberg & Kilkowski, 1985). Goldberg's research indicates that the abstractness and ambiguity of isolated adjectives and nouns limits their utility in conveying nuances of description. Consequently, Goldberg recommends embedding an adjective or noun in a context that specifies the word's meaning.

The full text of the Virtues Scale is available on the World Wide Web at: http://cac.psu.edu/ ~j5j/test/VS.html. The following is a sample questionnaire item for the virtue term *prudence*:

I am wisely cautious in practical affairs.
I am able to make correct decisions based on my good judgement, and common sense.
The people who know me best would describe me as prudent, discreet, and sensible.

For each item, subjects indicated a score along a seven point Likert scale indicating "Least like you really are" (1) to "Most like you really are" (7). The Virtues Scale asks subjects to report their "REAL virtues". Subjects are asked to respond as follows: "Make sure that you answer the way YOU REALLY ARE and not the way you ideally should be."

2.2.2. The Revised NEO Personality Inventory (NEO PI-R)

The NEO PI-R (Costa & McCrae, 1992) is a self-administered, paper and pencil inventory comprised of 240 agree-disagree statements. The NEO PI-R provides scores for five primary domains: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Each of the five domains contains six facets. Many validation studies are noted in the Revised NEO Personality Inventory (NEO PI-R) Manual (1992).

2.2.3. The Defining Issues Test (DIT)

Rest's (1979) Defining Issues Test provides a psychological measure of the ethic of duty.

Anastasi (1988) describes the DIT as "an instrument that combines Kohlberg's general theoretical orientation with psychometrically sound test construction procedures." Anastasi notes that the DIT:

> presents a series of moral-dilemma stories similar to those designed by Kohlberg. For each dilemma, respondents choose what they consider the proper action and also rate the importance of stated issues to be considered in resolving the dilemma. Rather than placing individuals into stages, the test generates several quantitative scores indicating the person's degree of sophistication in understanding and applying abstract moral principles. Empirical data on validity, reliability, and other psychometric features of the test have been accumulated through an extensive and continuing research program. (p. 583)

2.3. Procedure

Participants were provided with survey booklets, answer sheets, and a number two pencil. They were given as much time as they needed, At the conclusion of the study, the subjects were debriefed and given an information sheet describing the study.

3. Results

3.1. Phase I results: factor analysis

An exploratory factor analysis of the 140 Virtues Scale items was conducted to determine the underlying factor structure of this instrument. This factor analysis was based on the responses of 390 students. Factors were extracted by the "Maximum Likelihood" (ML) procedure and rotated using the Varimax rotation. The scree test suggested that a four factor solution was most appropriate for the data. Follow-up factor analyses utilizing three, four, and five factor solutions confirmed that the four factor solution was the most interpretable. Factors were interpreted and labeled according to the content of the highest loading items. The four factors combined accounted for 36.4% of the variance. Factor I, Empathy, showed an eigenvalue of 27.22, accounting for 19.4% of the variance. Factor II, Order, had an eigenvalue of 10.25, accounting for 7.3% of the variance. Factor III, Resourcefulness, possessed an eigenvalue of 8.15, accounting for 5.8% of the variance. Factor IV, Serenity, had an eigenvalue of 5.30, accounting for 3.8% of the variance.

3.2. Factor subscale construction

Four subscales were developed based upon these four factors. Subscale items were selected based on three criteria. First, items were selected for each factor based upon each item's factor loading scores. The top twenty items for each factor based on factor loadings were initially kept and scale reliabilities were calculated for each. Second, in order to develop brief and

reliable subscales, the lowest-loading items on each scale were dropped one at a time and reliabilities recalculated. Third, in order to retain independence of the four factors, one item (Courage) was dropped from subscale III because it also loaded high on Factor IV.

The items for the four virtue factor subscales are presented in Table 1. The four virtue scale factor subscale scores are calculated by taking the sum of the individual item scores for each subscale. The four subscales have demonstrated excellent reliability across all three phases of data collection as noted in Table 2.

3.3. Phase II and III results: personality correlates

The moderate correlations between the four Virtues Scale factor subscales and the five primary NEO PI-R personality scales are presented in Table 3. On the one hand, these moderate correlations indicate that the VS factors are not merely alternate measures of the NEO personality scales. On the other hand, the strongest correlations between the NEO scales and the VS factor subscales are sensible and meaningful, supporting the construct validity of the Virtues Scale. Moreover, Table 3 shows that the highest NEO-VS correlations from phase II were also the highest NEO-VS correlations in phase III. The coefficient of congruence between the two correlation matrices was found to be 1.00. The following paragraphs single out the NEO correlates that characterize each VS factor subscale. Each pair of Pearson coefficients represents the correlations from phases II and III, respectively. All correlations reported below are significant at the $p < 0.01$ level (two-tailed) except the $r = 0.22$ correlation, which is significant at the $p < 0.05$ level.

Table 1
Scale items for four virtue factor subscales

Factor I Empathy	Factor II Order	Factor III Resourceful	Factor IV Serenity
Empathy	Order	Resourceful	Serene
Concern	Discipline	Purposeful	Meek
Understanding	Serious	Perseverance	Forbearance
Considerate	Decent	Persistence	Forgiveness
Friendly	Deliberate	Confidence	Peaceful
Sympathy	Scrupulous	Sagacity	Patient
Affable	Earnest	Self-esteem	Merciful
Sensitive	Self-control	Fortitude	
Charity	Self-denial	Intelligence	
Compassion	Abstinence	Zealous	
Liberal	Obedient	Independent	
Gracious	Conservative		
Courtesy	Cautious		
	Careful		
	Tidy		
	Austere		
	Clean		

Table 2
Reliabilities for the four virtue factor subscales

Alphas	Factor I Empathy	Factor II Order	Factor III Resourceful	Factor IV Serenity
Phase I ($n = 390$)	0.93	0.90	0.87	0.85
Phase II ($n = 181$)	0.93	0.91	0.87	0.84
Phase III ($n = 143$)	0.93	0.87	0.80	0.80

VS Empathy is most strongly related to NEO domain scales Agreeableness ($rs = 0.48$, 0.48) and Extraversion ($rs = 0.40$, 0.26), and, within these domains, to NEO facets A3 Altruism ($rs = 0.66$, 0.57) and E1 Warmth ($rs = 0.64$, 0.56). According to the professional manual for the NEO-PI-R (Costa & McCrae, 1992), persons with high scores on these domain and facet scales tend to show affection, friendliness, a willingness to help others, and an active concern for the others' welfare.

VS Order is most strongly related to the NEO domain scale Conscientiousness ($rs = 0.63$, 0.53), specifically to its facets C2 Order ($rs = 0.57$, 0.39), C3 Dutifulness ($rs = 0.50$, 0.41), C5 Self-Discipline ($rs = 0.51$, 0.40), and C6 Deliberation ($rs = 0.55$, 0.41). Costa and McCrae (1992) note that "Conscientiousness is an aspect of what was once called character; high C scorers are scrupulous, punctual, and reliable" (p. 16). Persons with high scores on the C2, C3, C5, and C6 facet scales tend to keep things in their proper places, to adhere to their ethical principles, to fulfill their moral obligations, to carry tasks through to completion, and to think carefully before acting (Costa & McCrae, 1992).

VS Resourcefulness is most strongly related to NEO domain scales Neuroticism (as a negative correlation) ($rs =$ 0.60, 0.37) and Conscientiousness ($rs = 0.49$, 0.42), and to NEO facets E3 Assertiveness ($rs = 0.56$, 0.34), C1 Competence ($rs = 0.52$, 0.47), C4 Achievement Striving ($rs = 0.48$, 0.46), C5 Self-Discipline ($rs = 0.48$, 0.35), N3 Depression (as a negative correlation) ($rs =$ 0.53, 0.38), and N6 Vulnerability (negatively) ($rs =$ 0.65, 0.41). Persons with high scores on these domain and facet scales tend to be emotionally stable, purposeful, strong-willed, diligent, determined, forceful, effective, and unlikely to become discouraged or hopeless (Costa & McCrae, 1992).

VS Serenity is most strongly related to NEO domain scales Agreeableness ($rs = 0.45$, 0.44) and Neuroticism (as a negative correlation) ($rs =$ 0.38, 0.22), and to NEO facets A4 Compliance ($rs = 0.53$, 0.48), and N2 Angry Hostility (negatively) ($rs =$ 0.63, 0.50). Persons with high scores on these domain and facet scales tend to be sympathetic, calm, meek, mild, forgiving, easygoing, and slow to anger (Costa & McCrae, 1992). Johnson and Ostendorf (1993) compare this personality pattern to the Taoist conception of *femininity*, considered one of the highest virtues in that philosophical tradition.

The correlations between the four Virtues Scale factors and the Defining Issues Test (DIT) scores are presented in Table 4. There are no significant correlations between these two measures. This suggests that these two measures assess two separate domains within moral psychology. Table 4 also presents the correlations between the DIT and NEO-PI-R personality

Table 3
Correlations between the four virtue factor subscales and the NEO PI-R personality scales

NEO PI-R scale	Virtues scale factor subscales[a]							
	Empathy Phase		Order Phase		Resourceful Phase		Serenity Phase	
	II	III	II	III	II	III	II	III
Neuroticism	0.10	0.08	0.22**	0.09	0.60**	0.37**	0.38**	0.22*
N1: Anxiety	0.18*	0.24**	0.03	0.12	0.48**	0.26**	0.20**	0.06
N2: Angry hostility	0.16*	0.19*	0.22**	0.14	0.28**	0.19*	0.63**	0.50**
N3: Depression	0.11	0.03	0.07	0.06	0.53**	0.38**	0.18	0.14
N4: Self-consciousness	0.17*	0.17*	0.02	0.00	0.48**	0.27**	0.10	0.05
N5: Impulsiveness	0.11	0.01	0.47**	0.29**	0.28**	0.10	0.32**	0.23**
N6: Vulnerability	0.05	0.20*	0.17	0.04	0.65**	0.41**	0.31**	0.04
Extraversion	0.40**	0.26**	0.14	0.12	0.34**	0.30**	0.07	0.05
E1: Warmth	0.64**	0.56**	0.05	0.21	0.22**	0.30**	0.27**	0.30**
E2: Gregariousness	0.27**	0.18*	0.24**	0.10	0.01	0.01	0.02	0.07
E3: Assertiveness	0.10	0.14	0.09	0.05	0.56**	0.34**	0.01	0.20*
E4: Activity	0.14	0.03	0.11	0.10	0.40**	0.17*	0.03	0.07
E5: Excitement-seeking	0.12	0.05	0.36**	0.04	0.09	0.10	0.08	0.09
E6: Positive emotions	0.48**	0.35**	0.19*	0.06	0.15*	0.26**	0.13	0.21*
Openness to experience	0.34**	0.04	0.22**	0.37**	0.08	0.20*	0.04	0.09
O1: Fantasy	0.19**	0.02	0.43**	0.40**	0.19*	0.03	0.12	0.21*
O2: Aesthetics	0.31**	0.04	0.03	0.21**	0.08	0.12	0.10	0.14
O3: Feelings	0.50**	0.26**	0.16*	0.14	0.02	0.17*	0.02	0.05
O4: Actions	0.08	0.03	0.21**	0.30**	0.12	0.16	0.02	0.13
O5: Ideas	0.07	0.05	0.03	0.18*	0.26**	0.29**	0.08	0.02
O6: Values	0.19*	0.03	0.15	0.29**	0.00	0.10	0.05	0.05
Agreeableness	0.48**	0.48**	0.20**	0.31**	0.04	0.08	0.45**	0.44**
A1: Trust	0.17*	0.23**	0.18*	0.12	0.16*	0.21*	0.35**	0.26**
A2: Straightforwardness	0.28**	0.30**	0.27**	0.25**	0.05	0.03	0.30**	0.23**
A3: Altruism	0.66**	0.57**	0.16*	0.35**	0.15*	0.19*	0.32**	0.35**
A4: Compliance	0.22**	0.33**	0.20**	0.31**	0.06	0.03	0.54**	0.48**
A5: Modesty	0.26**	0.25**	0.06	0.18*	0.27**	0.12	0.18*	0.27**
A6: Tender-mindedness	0.48**	0.26**	0.09	0.03	0.09	0.05	0.13	0.20*
Conscientiousness	0.03	0.20**	0.63**	0.53**	0.49**	0.42**	0.25**	0.18*
C1: Competence	0.06	0.01	0.39**	0.28**	0.52**	0.47**	0.22**	0.07
C2: Order	0.00	0.12	0.57**	0.39**	0.25**	0.19*	0.11	0.00
C3: Dutifulness	0.07	0.28**	0.50**	0.41**	0.39**	0.28**	0.31**	0.18*
C4: Achievement striving	0.06	0.24**	0.45**	0.38**	0.48**	0.46**	0.09	0.13
C5: Self-discipline	0.01	0.19*	0.51**	0.40**	0.48**	0.35**	0.23**	0.17
C6: Deliberation	0.01	0.04	0.55**	0.41**	0.18*	0.10	0.24**	0.23**

[a] *Note*. $N = 181$ for Phase II; $N = 143$ for Phase III.
* $P < 0.05$; ** $P < 0.01$ (two-tailed).

Table 4
Correlations between the four Virtue Scale factors and the Defining Issues Test scores

DIT scores	Virtues Scale factor subscales				NEO-PI-R domain scales[a]				
	Empathy	Order	Resourceful	Serenity	N	E	O	A	C
Level 2	0.02	0.00	0.08	0.05	0.03	0.01	0.13	0.02	0.05
Level 3	0.02	0.03	0.03	0.13	0.02	0.03	0.05	0.01	0.19*
Level 4	0.04	0.15	0.03	0.03	0.08	0.16*	0.27**	0.00	0.12
Level 5A	0.04	0.00	0.02	0.00	0.10	0.03	0.25**	0.13	0.13
Level 5b	0.02	0.05	0.10	0.05	0.01	0.09	0.22**	0.04	0.10
Level 6	0.07	0.01	0.07	0.09	0.07	0.05	0.03	0.10	0.04
DIT P score	0.02	0.02	0.07	0.04	0.11	0.07	0.30**	0.06	0.08

[a] *Note.* $N = 181$. NEO-PI-R domain scores are as follows: N = Neuroticism, E = Extraversion, O = Openness to Experience, A = Agreeableness, and C = Conscientiousness.
* $P < 0.05$; ** $P < 0.01$.

scores. Of the five main NEO scales, only Openness to Experience correlated consistently with the DIT, with higher DIT levels associated with higher levels of Openness.

4. Discussion

The results support the preliminary statements offered in the introductory section concerning a psychological theory of the virtues. It seems that the reintroduction of virtue ethics into psychology can prove as fruitful as it has in philosophy, and that a psychology of the virtues, involving the dimensions Empathy, Order, Resourcefulness, and Serenity, can serve to complement the current models of moral development. The results also support the view that a psychology of virtue is more closely related to personality than to Kohlbergian moral development. The results suggest that a psychology of virtue can address aspects of moral psychology that are not sufficiently explained by moral developmental psychology alone.

4.1. The structure of virtue

Concerning the structure of virtue, the hypothesis that the factor analysis of the 140 virtue items would yield a reasonable and meaningful factor structure was supported by the four factor model of the structure of virtue derived from the factor analysis. VS Empathy correlates most highly with NEO PI-R Agreeableness and NEO PI-R Extraversion. Empathy describes the aspect of virtue involved with empathetically caring for and serving other people. VS Order correlates most highly with NEO PI-R Conscientious. Order describes the aspect of virtue involved with conscientiously maintaining order and discipline over one's self and one's behavior. VS Resourcefulness correlates most highly with NEO PI-R Neuroticism and NEO PI-R Conscientiousness. Resourcefulness describes the aspect of virtue involved with successfully developing and utilizing one's resources with confidence and a strong sense of purpose. VS Serenity correlates most highly with NEO PI-R Agreeableness and negatively with

NEO PI-R Neuroticism. Serenity describes the aspect of virtue involved with maintaining an attitude of serenity, meekness, and peace in the face of interactions with others and the world.

The four factors of the Virtues Scale each describe a distinct, coherent, and meaningful psychological virtue. Each is recognizable as a virtue, and none are duplicates of a non-virtue concept. In contrast to rationally developed schemes (Bertocci & Millard, 1963) or models developed for a particular context such as the workplace (Rust, 1998, 1999), the present factor structure has emerged from a comprehensive search of virtue terms in the natural language. The lexical tradition of culling a dictionary for natural language terms is based on the premise that all significant aspects of a human domain (in this case, the domain of virtue) become encoded in the language (John et al., 1988; Saucier & Goldberg, 1996). Whether we actually found all virtue terms in the dictionary may be tested some day by other researchers, just as Norman (1967) tested the adequacy of Allport and Odbert's (1936) lexical search.

Norman (1967) suggests that words for widespread, distinctive, socially significant, perceptible variations between persons find their way into the natural language. However, Norman and other lexical researchers say little about *why* traits are widespread, distinctive, socially significant, and perceptible enough to become encoded in language. One explanation comes from evolutionary ethics (Campbell, 1975; Waddington, 1967), which contends that natural selection attuned us to perceive and place emotional significance on moral traits that affected group cohesion and survival. Working in this tradition, Hogan (1973, 1983, 1996; Hogan, Johnson, & Emler, 1978) argues that moral dimensions are the most important part of personality.

4.2. Virtue and personality

Concerning the hypothesis of nonidentity of virtue and personality, we found that the virtue factors derived from the factor analysis correlated with the personality scales of the NEO PI-R well below the level of the scales' reliabilities. Thus, these data seem to support the idea that personality and virtue may be viewed as separate though related concepts. The personality factors most closely related to virtue are Conscientiousness, Agreeableness, and, to a lesser degree, the inverse of Neuroticism. Since both the virtues and personality approaches are concerned with enduring human dispositions, there are some moderate correlations between these constructs. However, these data make clear that the Virtues Scale factors are not merely personality factors or facets. Indeed, one might raise the question as to whether the NEO PI-R factors of Agreeableness and Conscientiousness might be broken down and reconceptualized as perhaps another model of psychological virtue.

McCrae and John (1992) acknowledge the evaluative nature of Agreeableness (A) and Conscientiousness (C), by noting that:

> Like A, C is a highly evaluated dimension; indeed, A and C are the classic dimensions of character, describing "good" versus "evil" and "strong-willed" versus "weak-willed" individuals. Perhaps it was these moral overtones that often led scientific psychologists to ignore these factors, but in fact, both represent objectively observable dimensions of

individual differences. Some people are thorough, neat, well-organized, diligent, and achievement-oriented, whereas others are not, and self-reports of these characteristics can be validated by peer or spouse ratings (McCrae & Costa, 1987). (p. 197)

McCrae and John (1992) also acknowledge that there are two components of Conscientiousness (C): an inhibitive view and a proactive view. They note that:

A number of different conceptions of C have been offered. Tellegen's Constraint and Hogan's Prudence reflect an inhibitive view of C as a dimension that holds impulsive behavior in check. Digman and Takemoto-Chock's Will to Achieve represents a proactive view of C as a dimension that organizes and directs behavior. The term Conscientiousness combines both aspects, because it can mean either governed by conscience or diligent and thorough. Empirically, both kinds of traits seem to covary. (p.197)

Perhaps the virtues factor Order represents the inhibitive, non-impulsive aspect of Conscientiousness as a virtue, and the virtues factor Resourcefulness represents the proactive, diligent aspect of Conscientiousness as a virtue (see also Johnson & Ostendorf, 1993).

The results of this study raise the issue of the status of personality and character. Historically, notions related to character helped form the early core of personality theory. With the advent of Allport (1937), character was divorced from personality. Thus personality was seen as a non-normative, non-evaluative construct. However, with the arrival of the five-factor models of personality, character returns in the guise of Agreeableness and Conscientiousness. However, the five-factor models of personality do not explicitly address issues of character and virtue, and as a result the structure of virtue remains unclear. Perhaps, the combination of both evaluative and non-evaluative terms in past five-factor personality research factor analyses has obscured the true structure of evaluative terms. This study looks directly and explicitly at the question of the virtues through the use of a factor analysis of normative, evaluative virtue trait terms. The results offer a four-factor model of virtue (Empathy, Order, Resourcefulness, and Serenity) which are not easily explained by past research in personality. At this point it appears that personality and virtues psychology represent separate but related domains of inquiry. However, in keeping with the historical roots of personality theory, the virtues might be brought under the rubric of personality through a broadened definition of it that would include both evaluative and non-evaluative terms.

One relatively recent area of applied research that concerns both the domains of personality and virtue is integrity testing for employment. Rust (1998, 1999), for example, describes the validation of Orpheus, a combined personality and integrity questionnaire. Five of the Orpheus scales were designed to represent the FFM in the context of the work environment. Seven additional scales were constructed to assess integrity traits believed to be important in the workplace: proficiency vs carelessness, good vs poor work orientation, patience vs anger, fair-mindedness vs resentfulness, loyalty vs disloyalty, disclosure vs lying, and initiative vs inertia. Researchers working in this area have suggested that *integrity* is a higher-order construct incorporating the FFM domains Agreeableness, Conscientiousness, and (the low end of) Neuroticism (Hogan & Ones, 1997; Sackett & Wanek, 1996). That suggestion is supported by the VS factor subscales' correlations with these three domains in the present study.

4.3. Virtue and moral development

Concerning the relationship between virtue and moral development, the hypothesis of this study — that the virtue factors derived from the factor analysis would not all correlate so highly (above 0.80) with the DIT P score as to be subsumed by it — was supported. The DIT P score and the Kohlbergian stage scores did not account for any of the four virtue factors. It seems clear from the data that the Virtues Scale and the DIT assess different moral domains. Given that the DIT P score correlated with none of the four Virtues Scale factors, we shall argue that the DIT P score does not measure all of the aspects of moral psychology that are accounted for by the Virtues Scale. Instead, the DIT P score seems to be a narrow measure of Enlightenment Liberalism. This view is supported by the correlation between the DIT P score and the NEO Openness score, which has been found to be correlated with various measures of Liberalism (McCrae, 1996). Hogan (1970) and Emler, Renwick, and Malone (1983) likewise demonstrated a link between Kohlberg's stage model and liberalism. It seems that the progression from the Virtues tradition to Kantian ethics to Kohlberg's Stage Model to Rest's DIT P score have left moral psychology with only one small slice of the moral pie. The Virtues Scale model seeks to serve the whole pie. As Kohlberg's critics have noted, the enthronement of openness and tolerance as the only virtues of modern moral developmental theory has led to a situation where moral psychology is dominated by situation ethics and moral relativism. The virtue and moral development results support the notion that a virtue-based psychological conception of the human person is distinct from, and much broader and richer, than that of modern moral developmental theory.

The independence of this measure of the virtues and the personality measures from the more cognitive DIT measure of moral development may also reflect the independence of the mental (cognitive–intellectual) and moral (emotional–motivational) domains in psychology and philosophy (Averill, 1980). Averill observes that the mental domain evolved from studies of epistemology, while the moral domain (including personality) evolved from studies of virtue ethics, motivation, will, and emotion. Thus, from Averill's observation, one would expect a measure of virtue to be more strongly related to measures of personality than to measures of cognitive moral development. Additional empirical data on the relationships among virtue, personality, moral cognitive development, and epistemological style can be found in Cawley (1997).

Thus, the results do seem to support the notion that a psychology of virtue can address aspects of moral psychology that are not sufficiently explained by moral developmental psychology alone. The virtues approach does indeed appear to consider aspects of moral psychology which are neglected by the Kohlbergian moral development approach. Perhaps, it is now time to reach back to the pre-Kohlberg character literature and reconsider it in light of this four-factor model of virtue.

In conclusion, this study has sought to reintroduce the concept of the virtues to modern psychology through the development of a reliable and valid Virtues Scale. This study presents a four-factor model of virtue involving: Empathy, Order, Resourcefulness, and Serenity. The psychological concept of virtue resonates with personality theory, and seems to complement moral developmental theory. While the virtue factors do correlate with related personality traits, they are not identical to them and capture an evaluative dimension of personality or

character that has been long neglected. The virtue factors seem to map a separate moral domain from that of the Kohlbergian moral development approach, since no meaningful and significant relationships were found between the four virtues scales and the DIT measure of moral development. This study concludes that the virtues approach to moral psychology appears fruitful and that the Virtues Scale is a reliable and valid measure of one perspective on the psychology of virtues.

References

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, *47*, 211.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: Macmillan.

Angleitner, A., Ostendorf, F., & John, O. P. (1990). Towards a taxonomy of personality descriptors in German: A psycho-lexical study. *European Journal of Personality*, *4*, 89–118.

Averill, J. R. (1980). The emotions. In E. Staub, *Personality: Basic aspects and current research*. Englewood Cliffs, NJ: Prentice Hall.

Averill, J. R. (1997). The emotions: An integrative approach. In R. Hogan, J. A. Johnson, & S. R. Briggs, *Handbook of personality psychology* (pp. 513–541). San Diego, CA: Academic Press.

Bertocci, P. A., & Millard, R. M. (1963). *Personality and the good: Psychological and ethical perspectives*. New York: David McKay.

Block, J. (1965). *The challenge of response sets*. New York: Appleton-Century-Crofts.

Brokken, F. B. (1978). *The language of personality*. Meppel, The Netherlands: Krips.

Burger, J. M. (1993). *Personality* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Burton, R. V. (1963). Generality of honesty reconsidered. *Psychological Review*, *70*, 481–499.

Campbell, D. T. (1975). On the conflicts between biological and social evolution and between psychology and moral tradition. *American Psychologist*, *30*, 1103–1126.

Cawley M.J. (1997). *The Virtues Scale: A psychological examination of the structure of virtue and the relationships between virtue, personality, moral development, and epistemological style*. Doctoral dissertation, The Pennsylvania State University, 1997. Dissertation Abstracts International.

Conley, J. J. (1985). A personality theory of adulthood and aging. In R. Hogan, & W. H. Jones, *Perspectives in personality*, vol. 1 (pp. 81–115). Greenwich, CT: JAI Press.

Costa Jr, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R™)* and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

Emler, N. P., Renwick, S., & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology*, *45*, 1073–1080.

Engler, B. (1991). *Personality theories* (3rd ed.). Boston, MA: Houghton Mifflin.

Goldberg L.R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In Mervielde I., Deary I., De Fruyt F., Ostendorf F. (Ed.), *Personality psychology in Europe* (vol. 7, pp. 7–28). The Netherlands: Tilburg University Press.

Goldberg, L. R., & Digman, J. M. (1994). Revealing structure in the data: Principles of exploratory factor analysis. In S. Strack, & M. Lorr, *Differentiating normal and abnormal personality* (pp. 216–242). New York: Springer.

Goldberg, L. R., & Kilkowski, J. M. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology*, *48*, 82–98.

Hartshorne, H., & May, M. A. (1928). Studies in the nature of character. *Studies in deceit*, vol. 1. New York: Macmillan.

Hartshorne, H., May, M. A., & Maller, J. B. (1929). Studies in the nature of character. *Studies in self-control*, vol. 2. New York: Macmillan.

Hartshorne, H., May, M. A., & Shuttleworth, F. K. (1930). Studies in the nature of character. *Studies in the organization of character*, *vol. 3*. New York: Macmillan.

Hogan, J., & Ones, D. (1997). Conscientiousness and integrity at work. In R. Hogan, J. A. Johnson, & S. Briggs, *Handbook of personality psychology* (pp. 849–870). San Diego, CA: Academic Press.

Hogan, R. (1970). A dimension of moral judgement. *Journal of Consulting and Clinical Psychology*, *35*, 205–212.

Hogan, R. (1973). Moral conduct and moral character: A psychological perspective. *Psychological Bulletin*, *79*, 217–232.

Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page, *Nebraska Symposium on Motivation 1982 — Personality: Current theory and research* (pp. 55–89). Lincoln, NE: University of Nebraska Press.

Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins, *The five-factor model of personality: Theoretical perspectives* (pp. 163–179). New York: Guildford.

Hogan, R., Johnson, J. A., & Emler, N. P. (1978). A socioanalytic theory of moral development. In W. Damon, *New directions for child development. Volume 2: Moral development* (pp. 1–18). San Francisco, CA: Jossey-Bass.

Hooker R. 1996. *Areté*. Retrieved April 7, 1999 from the World Wide Web: http://www.wsu.edu/ ∼ dee/ GLOSSARY/ARETE.HTM.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality*, *2*, 171–203.

Johnson, J. A. (1990). Unlikely virtues provide multivariate substantive information about personality. *Paper presented at the 2nd Annual Meeting of the American Psychological Society, Dallas, TX*.

Johnson J.A. Predicting observers' ratings of the big five from the CPI, HPI, and NEO-PI-R: A comparative validity study. *European Journal of Personality* (in press).

Johnson, J. A., & Horner, K. L. (1990). Personality inventory item responses need not veridically reflect "actual behavior" to be valid. *Paper presented at the 61st Annual Meeting of the Eastern Psychological Association, Philadelphia, PA*.

Johnson, J. A., & Ostendorf, F. (1993). Clarification of the five-factor model with the abridged big five dimensional circumplex. *Journal of Personality and Social Psychology*, *65*, 563–576.

Kilpatrick, W. K. (1992). *Why Johnny can't tell right from wrong*. NY: Simon & Schuster.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive developmental approach. In T. Lickona, *Moral development and behavior: Theory, research, and social issues* (pp. 31–53). New York: Holt, Rinehart, & Winston.

Kohlberg, L. (1981). *The meaning and measurement of moral development*. Worcester, MA: Clark University Press.

Kohlberg, L. (1984). Essays on moral development. *The psychology of moral development: The nature and validity of moral stages*, *vol. 2*. San Francisco, CA: Harper & Row.

Kurtines, W. M., & Gewirtz, J. L. (1984). *Morality, moral behavior, and moral development*. New York: Wiley.

Kurtines, W. M., & Gewirtz, J. L. (1995). *Moral development: An introduction, 1*. Boston, MA: Allyn & Bacon.

Lewis, C. S. (1947). *The abolition of man*. New York: Macmillan.

Lickona, T. (1976). *Moral development and behavior: Theory, research, and social issues*. New York: Holt, Rinehart, & Winston.

Liebert, R. M., & Spiegler, M. D. (1990). *Personality: Strategies and issues* (6th ed.). Pacific Grove, CA: Brooks/ Cole.

Loevinger, J. (1985). A century of character development. In S. Koch, & D. E. Leary, *A century of psychology as science*. New York: McGraw-Hill.

MacIntyre, A. (1981). *After virtue*. London, UK: Duckworth.

Maller, J. B. (1934). General and specific factors in character. *Journal of Social Psychology*, *5*, 97–102.

Mayo, B. (1958). *Ethics and the moral life*. London: Macmillan.

McAdams, D. P. (1990). *The person: An introduction to personality psychology*. San Diego, CA: Harcourt Brace Jovanovich.

McCrae, R. R. (1996). Social consequences of experiential openess. *Psychological Bulletin*, *120*, 323–337.

McCrae, R. R., & Costa Jr, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, *51*, 882–888.

McCrae, R. R., & Costa Jr, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81–90.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175–215.

Meilaender, G. C. (1984). *The theory and practice of virtue*. Notre Dame, IN: Notre Dame Press.

Nicholson, R. A., & Hogan, R. (1990). The construct validity of social desirability. *American Psychologist*, *45*, 290–292.

Norman, W. T. (1967). *2,800 personality trait descriptors: Normative operating characteristics for a university population*. University of Michigan: Department of Psychology.

Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.

Rest, J. R. (1986). *Moral development: Advances in theory and research*. New York: Praeger Press.

(1980) *Roget's Pocket Thesaurus*. Boston: Houghton Mifflin.

Rushton, J. P. (1984). The altruistic personality. In E. Straub, D. BarTal, J. Karylowski, & J. Reykowski, *Development and maintenance of prosocial behavior* (pp. 271–290). New York: Plenum Press.

Rust, J. (1998). The validation of the Orpheus minor scales in a working population. *Social Behavior and Personality*, *26*, 399–406.

Rust, J. (1999). Discriminant validity of the 'big five' personality traits in employment settings. *Social Behavior and Personality*, *27*, 99–108.

Sackett, P. R., & Wanek, J. E. (1996). New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness and reliability for personnel selection. *Personnel Psychology*, *49*, 787–830.

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins, *The five-factor model of personality: Theoretical perspectives* (pp. 21–50). New York: Guilford.

The Merriam-Webster Thesaurus (1989). New York: Pocket Books.

The New Merriam-Webster Dictionary (1989). Springfield, MA: Merriam-Webster.

The Random House Dictionary (1980). New York: Random House.

Vitz, P. C. (1990). The use of stories in moral development. *American Psychologist*, *45*, 709–720.

Waddington, C. H. (1967). *The ethical animal*. Chicago: University of Chicago Press.

Wiggins, J. S., & Trapnell, P. D. (1997). Personality structure: The return of the big five. In R. Hogan, J. A. Johnson, & S. Briggs, *Handbook of personality psychology* (pp. 737–765). San Diego, CA: Academic Press.