



THE
JUBILEE CENTRE
FOR CHARACTER & VALUES

UNIVERSITY OF
BIRMINGHAM

Assessing Virtue: Lessons from Subfields of Psychology

Blaine Fowers

These are unpublished conference papers for the 'Can Virtue Be Measured?', held by the Jubilee Centre for Character and Values at Oriel College, Thursday 9th – Saturday 11th January 2014. These papers are works in progress and should not be cited without author's prior permission.

**School of Education
University of Birmingham
Edgbaston
Birmingham**

Assessing Virtue: Lessons from Subfields of Psychology

Blaine J Fowers, Ph.D.
University of Miami
Department of Educational and Psychological Studies
Coral Gables, FL, U.S.A.

The quickening of interest in the virtues across multiple disciplines and the very important real-world implications of virtues have increased the urgency of developing valid and useful assessments. There have been many scattered attempts to measure various virtues, and these attempts have met with varying success (e.g., Cawley, Martin, & Johnson, 2000; Peterson & Seligman, 2004). I have a qualified answer to the question that animates this conference, can virtues be measured? I believe that several important developments suggest that we can be relatively optimistic that a dedicated community of scholars can measure virtue in meaningful ways if we adopt a multi-method approach. I qualify my answer because, as I hope to make plain, measuring virtue is a very complex and daunting task. The successes to date suggest, however, that it is no longer so much a question of whether virtue can be measured, but how well and for what purposes it can be measured.

The time has come for a more programmatic approach to virtue assessment that can pave the way toward a fruitful science of virtue and flourishing. This presentation focuses on four major challenges for the measurement of virtue and suggests proven approaches to addressing these challenges. One primary theme of the presentation is the importance of recognizing that no single study or single research method can address all of the challenges for even a single virtue. Success in virtue assessment will require the systematic effort of many researchers over an extended period of time. This project is much more likely to succeed if investigators can work from a shared understanding of the concept of virtue. Collaboration

among researchers in developing virtue constructs and conducting research using a variety of coordinated methods is also a key to success. This presentation will provide one possible starting point for such an endeavor, and indicate how some of the key challenges can be addressed.

Inasmuch as Aristotle has been the predominant touchstone for recent work on virtue, this presentation will have an Aristotelian bent. Yet I will address virtue in a way that is substantially compatible with any eudaimonistic viewpoint (e.g., Platonic or Stoic).

Eudaimonistic theories take virtues as the character strengths that make it possible for human beings to flourish (Brewer, 2009; Fowers, 2012). The value of a eudaimonistic perspective is that virtue has a specifiable and systematically important role in the good life rather than being isolated as specific traits whose desirability is limited to a presupposed, but unarticulated moral scheme.

Virtue Constructs

Any serious attempt to investigate a domain of interest in the human sciences must begin with a clearly defined construct or constructs. Social scientists use the term “construct” to designate a phenomenon of interest that is not directly observable, but is theorized to explain or predict other phenomena. Constructs are abstract idea about qualities that individuals have such as intelligence, creativity, or anxiety. Because it is impossible to observe constructs directly, we devise measures that can serve as indicators of the construct, such as tests for intelligence or heart rate for anxiety. Indicators of constructs can take many forms, including self-reports, expert ratings, and physiological measurement. On this view, virtue is a construct because it is not observable in the way that mass or velocity are, yet many theorists

see virtue as an important explanatory concept with respect to human actions, relationships, and welfare.

If a construct of interest is vaguely or inadequately defined, there is no possibility for satisfactory measurement. Poor construct definition has characterized research on virtue, beginning with Hartshorne and May's (1928) honesty studies and continuing to the present. Fortunately, scholars have devoted themselves to defining virtue more assiduously (e.g., Fowers, 2005; Peterson & Seligman, 2004) and in ways that allow the construct to be measured. There are two distinct levels of construct definition that must be considered in virtue measurement—the general construct of virtue and the elements that are specific to each virtue.

A General Definition of Virtue

The common sense approach to virtue is to describe and study a specific virtue such as courage or loyalty. Such specific character strengths are concrete, relatively easy to grasp, and they have a place in our ordinary speech about praiseworthy actions. It is essential to transcend common sense ideas in order to develop a systematic conceptual understanding of virtue that can inform the measurement project. The first step in clarifying the definition of virtue is to recognize the distinction between particular virtues (e.g., honesty, generosity) and virtue as a general concept. I believe that we must begin by understanding what we mean by the general concept of virtue itself before we delve deeply into specific virtues. All of the virtues have common features central to their definitions as virtues. If our understanding of virtue in general is inadequate or erroneous, these shortcomings will inevitably lead us to misconceive any and all of the specific virtues.

Common features of virtues. As I describe in detail elsewhere (Fowers, 2005), I suggest that virtues have the following features:

1. Virtues are enacted for the sake of some good (e.g., justice or knowledge).
2. Virtues always show up in behavior.
3. Virtues are accompanied by congruent emotions and motivated spontaneously.
4. Virtues are habitual and therefore consistent over time.
5. Virtues are based on knowledge of how best to act.
6. Practical wisdom is evident in virtues.
7. Virtues are enacted in situations specific to the virtue (i.e., courage is important in situations involving risk).

Comprehensive assessment of virtue will encompass all of these features. Although it is clear that not every study can encompass all of these elements of virtue, the fewer features there are in a measurement approach, the weaker the assessment of virtue will be. The most commonly neglected element in the virtue construct is the most important one, that virtues are defined as the character strengths that make it possible to pursue what is good. This is the core characteristic of virtues in a eudaimonistic theory, but virtues are often assessed as though they are enacted for their own sake rather than for the sake of some human good. It is important to conceptualize virtues in terms of the goods that they help to bring about. A research approach to the human goods can be framed in terms of goal pursuit because goals are concrete versions of a human good (e.g., the goal of a college degree can be part of the pursuit of the good of knowledge). There is an extensive literature on goal pursuit (Covington, 2000; Eccles & Wigfield,

2002; Fowers, Mollica, & Procacci, 2010), which can provide paradigms for studying virtues in the context of goal pursuit.

Character types. Another frequently neglected aspect of virtue is the concept of character types, and it is useful to distinguish five types of overall character. The first and ideal type of character is the virtuous individual, who is able to act consistently in a way that is fitting, given the circumstances. Virtuous character is guided by a clear vision of what is good and admirable, and the individual is spontaneously motivated to act consistently in the service of those goods. This shows up as a clear harmony between the individual's desire and what he or she sees as the proper action or duty. This emotional harmony differentiates the virtuous character from the continent and incontinent, who live in perpetual conflict between duty and desire. The key motivational concept of virtue is the love of what is good. Virtuous individuals cultivate and enact strong characters because they are deeply attracted to worthwhile aims such as democracy, justice, and knowledge.

One contrast with virtue is the pair of character types known as the continent and the incontinent. In both of these character types, the individual knows what is admirable and good, but the individual's understanding of the best kind of action and his or her desire to follow through with admirable action are often at odds with each other. Thus, continent and incontinent individuals struggle with a conflict between duty and desire. Continent individuals are those who know how to act well and decide to do so, even though this decision is contrary to their desires and involves inner conflict. As I have discussed elsewhere (Fowers, 2005, 2008), it is unfortunately common for social scientists to confuse virtuous and continent individuals (e.g., Baumeister & Exline, 2000; Peterson & Seligman, 2004), positioning them as theorists of

moral continence rather than of virtue. Individuals who know what they should do but frequently cannot bring themselves to do it have incontinent characters. Incontinent individuals frequently feel guilt, remorse, and shame about their inability to act in the ways they think are fitting.

The fourth type is the vicious character. The term 'vicious' does not necessarily connote hostility or mean-spiritedness, simply that one consistently chooses to act ignobly, such as a characteristically greedy, deceitful, exploitive, or self-indulgent person. The vicious person feels entirely justified in their ignoble actions because they have a misguided sense of what is good. Vicious individuals pursue ends such as acquisitiveness, domination of others, and self-indulgence as goods in themselves. They do this by construing themselves and the world in a way that makes their actions seem appropriate and justified to them. Most social scientists do not recognize the vicious character, which is a serious flaw because their characterizations of a good life may fit a vicious character as well as a virtuous one (Fowers, 2008; Kristjansson, 2013).

The final type is the beastly character, which includes individuals who have become enslaved to desires or habits that suppress or destroy their basic human capacities to reason, make choices, and participate in good relationships with others. The prototype of beastliness is a drug addict whose entire being is devoted to serving the addiction.

These character types are a way of looking at an individual as a whole. Clearly, it is also important to recognize, study, and foster individual virtues.

Specific Virtue Constructs

There are, of course, many specific virtues that we can study, and there is no canonical list. It seems wise, however, to focus our initial investigations primarily on the virtues that

appear on virtually every list. These include virtues such as courage, honesty, generosity, and justice. Gratitude, kindness, and compassion are also frequently studied virtues, due to the modern emphasis on the interpersonal nature of morality, and the modern interest in the reduction of suffering (Taylor, 1989). It is important for researchers to carefully define the virtues they wish to study, and to include the seven elements of virtue in the previous section in their definitions and measurement as far as possible.

A focus on the situations that call for particular virtues may be especially helpful to virtue researchers as arranging such situations experimentally or focusing on the situations in non-experimental research may prove more enlightening than less targeted studies. For example, investigators interested in courage should focus their attention on situations in which research participants face some degree of risk in which something important is at stake. The risk could be physical, emotional, or social, and the stakes ought to be roughly commensurate with the risk. Similarly, the virtue of compassion could be most fruitfully studied in a context in which some form of human frailty is evident, and justice could be studied when there is some distribution of resources, rewards, burdens, or rights at stake.

Trait/Situation Issues

There has been a century-long debate in psychology over whether behavior is primarily influenced by traits or situations, with personality psychologists arguing for the importance of traits and social psychologists arguing for the influence of situations. Thankfully, this debate appears to be nearing its end in the conclusion that both traits and situations are important influences (Fleeson, 2001, 2007; Molenaar & Campbell, 2009). The most interesting aspect of recent work on this question has indicated that traits and situations are not simply independent

contributors to behavior. Traits and situations appear to interact in that certain situations tend to amplify traits. For example, social situations tend to amplify the expression of extroversion. This suggests that it is important to consider both traits and situations in every instance of behavior. The work on the trait/situation question has focused on personality traits, but it can be applied to the interactions of virtue traits and situations.

One of the most common arguments against the existence of virtues is that behavior is primarily situationally determined. The claim that virtues do not exist because behavior can be explained by situational factors was advanced by Doris (2002) in an impressive discussion of over 100 experimental studies that seemed to support his conclusion. Doris' position is not unique. Rather, it is representative of the views held by many psychologists and philosophers that character and virtue are either non-existent or unhelpful in explaining behavior (Harman, 2009; Ross & Nisbett, 1991). Many authors have disputed Doris's viewpoint (Homiak, 2011; Kupperman, 2001; Lefevor & Fowers, 2012; Sreenivasan, 2002), and the weight of argument seems to run counter to his claims. There are also two forms of empirical evidence that undermine the situationist viewpoint. The most common type of study Doris cited in his argument was experiments on helping behavior. There have been hundreds of such studies, and they very consistently show that a number of situational factors encourage helping behavior (e.g., factors that increase positive mood). My colleagues and I have just completed a meta-analysis (a study that combines the empirical results of many studies) of 283 experimental studies designed to investigate situational influences on helping behavior. We found that although situational factors do promote helping behavior, helping was also common in control conditions that designed specifically to include no situational encouragements for

helping (Lefevor, Fowers, Ahn, Cohen, & Lang, 2013). Nearly half as many people help in the control conditions as in the conditions with situational inducements to helping.

Doris and others have focused on the fact that situational factors increase the likelihood of helping, which is a very helpful piece of the puzzle. Yet he and the authors of all of those experimental studies have entirely neglected the fact that a significant proportion of participants in the control condition also engaged in helping behavior. The two likeliest explanations for non-situationally encouraged helping are there is simply a statistical likelihood of helping behavior among humans by nature or that there is a subset of people who have the trait of helpfulness; that is, as a matter of character.

The second source of evidence that behavior cannot be entirely explained by situational factors has been presented by Fleeson and his colleagues (Fleeson, 2001; Fleeson & Gallagher, 2009; Meindl, Jayawickreme, Furr, & Fleeson, 2013). They developed a density distribution model of traits that views traits as a distribution of behaviors that may manifest more or less of a trait. They suggest that individuals manifest behavior at nearly all levels of all traits. People differ in how frequently they manifest each level of each trait. The task is then to assess how often and how strongly a person manifests the trait over time. These researchers have found that there is a good deal of within-person variation in how much a trait shows up in behavior, which corroborates the research on the importance of situational factors in behavior (Fleeson, 2001; Fleeson & Gallagher, 2009). In spite of this within-person variability, the average level of trait expression and the variation in an individual's trait expression are remarkably consistent over time. These studies have also demonstrated very strong consistency in individuals' average level of trait expression across time.

The published work from this group of researchers has concerned personality traits up to this point, but they have conducted several unpublished studies of virtue. For virtue traits, Meindl et al. found correlations between time periods of .66 - .94. Using their behavioral measures, the average correlation between social desirability and virtue scales was .28, and the only statistically significant correlation was between impression management and compassion.

Although virtue researchers all seem to recognize that virtues are traits, one of the most common weaknesses in virtue assessment is that virtues are measured by simple summary self-report scales. Of course, the items can be written in such a way that they reflect stable traits, but no one-time assessment can document a trait unambiguously. True traits must show up across time and situations. Experience sampling provides a much stronger method for assessing the stability of traits across time and situations. This is particularly important in the assessment of virtues, a key feature of which is a habitual inclination to act according to the virtue that is often referred to as “second nature.” In experience sampling methods, research participants are prompted to respond to questions on at least five occasions to allow the individual assessment of change or stability of a particular time interval (Bolger & Laurenceau, 2013). Experience sampling studies typically sample at least daily for one to three weeks. This approach to studying virtues solves two vexing problems. The first is the ability to demonstrate the stability of a virtue over time and across situations. The second is that the individual is the unit of analysis, which is the appropriate unit for virtue measurement. Although the experimental method has many important uses in general and for the assessment of virtue, the method generally relies on the assessment of differences between groups of people in different conditions. Virtues are characteristics of individuals, not groups, which means that

experimental research on situational influences on behavior is inherently unsuitable to assess virtue. In technical terms, situationist research commits the ecological fallacy, in which group differences are interpreted as individual differences (Robinson, 1950).

Experience sampling also allows us to measure the influence of situations and traits in the same study. Moreover, it is possible to assess the degree to which traits and situations interact. This is particularly important in understanding character types inasmuch as one of the best ways to detect differences between the character types is in the ways that situations influence behavior. Whereas a virtuous character type should be consistent in acting well across situations, a continent character will be somewhat less consistent, and an incontinent character will frequently be swayed by situational factors.

Veridicality Issues

To date, the vast majority of research on virtues has used direct global self-report methods, particularly in the form of questionnaires about specific virtues. This popularity is parallel to global self-reports in the personality field (Robins, Tracy, & Sherman, 2007). This is, of course, a very expeditious method for studying virtue, with the primary advantage of opening “a pipeline to prodigious amounts of unique information about the target of assessment” (Paulhus & Vazire, 2007, p. 235). The downside of self-report methods is that they are particularly vulnerable to questions of veridicality, and this form of research alone cannot allay doubts about the actuality of a virtue.

Challenges to Veridicality

The most obvious challenge to the truth value of self-report questionnaires is the social desirability response bias. When individuals respond to questionnaire items about themselves,

there is a common tendency to give the most socially desirable response. This bias has also been an issue in research on personality (Bäckström, 2007). Social desirability may be seen as particularly pernicious in self-report measures of virtue because traits such as courage, generosity, and justice have widespread approbation. There are two very well documented forms of social desirability (Paulhus, 1984), impression management and self-deception. Impression management occurs when an individual responds in such a way to influence another person to see him or her in the best possible light. Self-deception occurs when the individual is fully convinced that he or she has the socially desirable characteristic.

There is some disagreement about whether social desirability is a problem in personality measurement, however. McCrae and Costa (1983) suggested that “social desirability scales are better interpreted as measures of substantive traits than as indicators of response bias” (p. 882). This viewpoint undercuts the worry that responses to virtue measures could be influenced by a social desirability response bias. At the same time, Paulhus and Vezire (2007) note that “it is only prudent to be skeptical about respondents’ claims about their personality [or virtue]—especially on highly evaluative traits” (p. 235). Even if one accepts McCrae and Costa’s view that social desirability is a trait, then an important issue remains because it would be necessary to demonstrate the discriminant validity (distinctness) of any virtue trait from the trait of social desirability.

A second challenge to the veridicality of self-reports of virtue is the self-enhancing positive illusions that appear to be widespread in normal cognition (Alicke & Sedikides, 2009; Taylor & Brown, 1988). There is a great deal of evidence that normal individuals tend to see themselves in unrealistically positive ways (Fowers & Lang, in press). These illusions include

seeing oneself as having very positive attributes, being unrealistically optimistic about one's personal future, and having more control over life circumstances than is reasonably possible. These positive illusions are relatively mild and appear to contribute to individuals' psychological and physical well-being (e.g., Dufner et al., 2012; Why & Yuang, 2011). To the extent that individuals have positive illusions, self-reports of virtues are likely to be inflated.

A third challenge to the truth value of self-reports of virtue is the confirmation bias, which is the tendency to seek, interpret, and retain information that confirms what one already believes or that one would prefer to believe (Greenwald, 1980). This process is generally conceived as an unwitting selectivity in seeking information, searching memory, and responding to memory (Nickerson, 1998). When individuals think of themselves as honest or courageous, there is a tendency to recall only those actions that confirm this self-perception.

Finally, the items in self-report measures of virtues tend to be worded in trait terms—appropriately so—but this means that the respondent must summarize his or her self-perceptions over a very long period of time. Such summarization exacerbates the likelihood and extent of distortion due to the three challenges already noted.

Addressing Veridicality Concerns

There are a variety of ways to address the concerns outlined in the previous section with all of the measurement methods available to researchers. There are a number of ways to assess traits, including virtues, each with its own advantages and disadvantages. The methods include self-report measures, other-report measures, expert ratings, experience sampling, and experimental approaches. It is important for investigators of virtue to incorporate the best

methods for reducing bias in their results for the assessment method they choose for each particular study.

Global self-report assessment. As already indicated, self-report scales are quite vulnerable to all four challenges to veridicality. Self-report measurement has the advantages of being an inexpensive and time-efficient method for assessing traits, so it is likely to remain a very common form of virtue assessment. In addition, Paulhus and Vezire (2007) point out that the individual may be the best-qualified assessors of his or her personality due to having much more information than anyone else about their private and public behavior and intrapsychic information across contexts. This may be even truer of assessing virtue traits because extremely important elements of virtue include information about their cognition, emotion, motivation, and goals. These are vital elements of virtues to which other observers may not be privy. The social desirability biases are the most problematic of the challenges. Researchers have found that self-deception is not a threat to the veracity of virtue, so I will focus on impression management (Meindl et al., 2013). There are five approaches to addressing impression management.

First, the researcher can test the degree to which the scale is vulnerable to the social desirability biases by examining the relationships between the virtue scale and measures of social desirability. For example, correlations between the scales of the Virtue in Action Inventory and social desirability have been relatively low (Peterson & Seligman, 2004; Ruch et al., 2010). If the correlations are sufficiently low, it can alleviate concerns about the potentially confounding effects of social desirability.

Second, the influence of social desirability can be removed statistically by including measures of social desirability in studies of virtue and partialing them out of the trait measure. The problem with statistical control is that it can reduce the criterion related validity of the virtue scale because some of the variance that social desirability measures and virtue scales share may provide valuable information about virtue (Meindl et al., 2013; Paulhus & Vezire, 2007).

Third, one can use what is known as an ipsative format in which respondents choose between responses that are comparable in terms of desirability. This is a method that addresses social desirability through item construction.

The fourth method for reducing the impression management bias also involves item construction. The scale items can be phrased so that they do not have the positive or negative valence that induces people high in social desirability to respond in a biased way. This has the advantage of reducing social desirability without giving up the content of the scale items. Bäckström, Björklund, and Larsson (2009) dramatically reduced the relationships between impression management and measures of the Big Five Personality traits. Their neutrally valenced scales were strongly correlated (.70 - .83) with the standard measures of the traits, but had much weaker relationships to social desirability than the standard measures.

Finally, Meindl et al. (2013) created virtue items that balanced the social desirability of a virtue description with euphemism/dysphemism statements. They used euphemisms to make socially undesirable characteristics such as deception more palatable (e.g., “linguistically creative”) and dysphemisms to make socially desirable characteristics, such as being concerned about others’ welfare, less attractive (e.g., “evolutionarily unfit”). These researchers found that

the use of this method reduced social desirability for measures of four virtues: honesty, fairness, compassion, and courage).

The other three challenges to veridicality are very difficult to manage in self-report scales of virtue. Positive illusions are difficult to control because the standard measures of unrealistically positive self-perceptions are discrepancies between self-ratings on positive attributes and ratings of the average person. To the degree that an individual has virtues, it would be accurate for the individual to rate himself or herself more highly on positive attributes. Confirmation biases will be difficult to remove from self-report indices of virtue because self-report measures are based on recall, which is the primary cognitive process where confirmation bias emerges. Similarly, distortions due to summarizing one's actions over a significant time period are inherent in self-report scales of traits.

Other-report. Having a respondent describe another individual's virtues could alleviate some of the concerns about veridicality. Barranti, Hawkins, Helzer, Fleeson, and Furr (2013) found that ratings from other people (friends, family members, and co-workers) indicated some consensus on the virtue traits of fairness, compassion, honesty, and temperance. Consensus was stronger for adults in the community than for university students. In another study, Hawkins, Fancullo, and Furr (2013) found that other-reports successfully differentiated between two people on traits of altruism and tender-mindedness. The other-reports were correlated with self-reports of tender-mindedness but not altruism.

A German language version of the Virtues in Action Inventory was correlated with peer ratings on the 24 strengths it contains (Ruch et al., 2010). The correlations of these two forms of measurement were substantial, ranging from .29 to .69, with a median of .40.

This method has also been used with individuals reporting on their spouse's virtues, including other-centeredness, generosity, admiration, teamwork, shared vision, and loyalty (Hawkins, Fowers, Carroll, & Yang, 2007). The construct validity of this Marital Virtue Profile was supported by the fact that it was a better predictor of marital quality scores than standard communication and conflict measures. Subsequent studies have found that the Marital Virtue Profile also predicts good relationship problem solving (Veldorale-Brogan, Lambert, Fincham, & DeWall, 2012) and that marital virtues mediate the relationship between individual well-being and relationship quality (Veldorale-Brogan, Bradford, & Vail, 2010). Although it is likely that the challenges to veridicality are reduced with other reports of virtue, there remains some question about the degree to which such other reports are subject to social desirability and positive illusions because research has indicated that both factors are active in relationship measurement (e.g., Fowers, Lyons, Montel, & Shaked, 2001). Other reports could also be obtained from friends, family members, co-workers, supervisors, and so forth, with increasing interpersonal distance reducing the likelihood that social desirability and positive illusions would influence the reports.

Expert rating. One way to eliminate the influence of all four challenges to veridicality would be to develop a rating system with which one could train experts in rating virtues. To my knowledge, this has not been attempted. This method has two significant challenges, however. The development of expert rating scales and training in them are very time-consuming and difficult endeavors. In addition, the opportunity for observing individuals sufficiently to rate them on virtues would be difficult to arrange. Behavioral rating scales are generally used with

relatively short samples of behavior, and this practice would likely be insufficient to assess virtues.

Experience sampling assessment. I will only briefly describe experience sampling because that is the focus of Eranda Jayawickreme's talk, one of the trait researchers at Wake Forest University who are pioneering experience sampling for virtue traits. These researchers have conducted a number of studies in which they have participants respond to questions about virtuous behaviors multiple times per day for nine or more days (e.g., Meindl et al., 2013). This allows a more direct assessment of the degree to which an individual manifests a virtue trait. One of the methods Fleeson and his colleagues use to reduce veridicality biases is that they ask participants to respond to behavioral statements that are indicators of the virtue trait, and the descriptions typically refer only to the previous three hours. The short period of time assessed in these reports helps to reduce the problem that emerges in global self-reports: integrating and simplifying a great deal of information. Using this method, they have been able to reliably measure honesty, compassion, and fairness (Meindl et al., 2013).

Experimental research. The concerns with veridicality can be entirely obviated by assessing behavior directly. Some research on virtue has been conducted experimentally, most notably the Hartshorne and May (1928) studies of honesty in children, the results of which have been in dispute ever since they were reported. Some virtues lend themselves to experimental test, including honesty, kindness, generosity, and justice. Experimental conditions could be arranged to provide opportunities for individuals to demonstrate virtuous actions, particularly by experimental conditions that instantiate the situational factors that call for the virtue (e.g., risk and courage). The shortcoming of the experimental method is that it is generally short in

duration and it would be difficult to assess the consistency of virtuous behavior. In one noteworthy experiment, honesty was assessed by providing an opportunity to cheat for financial gain (Greene & Paxton, 2009). The investigators found that a significant proportion of their participants did not take the opportunity to cheat, and that the honest individuals made their decisions more quickly than those who cheated. Moreover, they found that the control network of the brain was not activated in the honest individuals, but it was in those who cheated (assessed with functional magnetic resonance imaging). The fact that honest individuals responded more quickly and without apparent thought suggests that the honest responses were spontaneously motivated, one of the seven key features of virtue. Interestingly, Greene and Paxton did not discuss their results in terms of virtue even though a virtue framework provides the most comprehensive and compelling explanation of their results (Fowers, in press), suggesting that they did not fully appreciate what they had found. This fascinating study could provide one experimental paradigm for future studies of virtues.

Construct Validity Issues

Construct validity is the central problem in most social science measurement because it is the degree to which a given assessment measure is a useful indicator for a theoretically identified construct. To reiterate, a construct is an abstract, theorized quality of an individual such as extraversion or courage. Construct validity is the overarching term used to encompass a number of ways that measures can be assessed for their soundness. Construct validity is a matter of degree, and it is never final. We can have more or less confidence that a measure is valid based on the results of assessments for specific uses. I briefly discuss the five most

important forms of construct validity for virtue assessment: content, criterion, convergent, discriminate, and incremental validities.

Content Validity

Content validity refers to the question of whether a scale covers the relevant content for the construct. This form of validity is assessed through expert judgments about whether the items in the scale are appropriate for the construct and cover the essential elements of the construct. When developing a measure of a virtue, it is important to have a panel of experts in that virtue examine the items to evaluate the content validity of the new scale. This is at least as important in virtue measurement as in any other area because the make-up of virtues is not well-established at this point. Accepting the criteria listed in the first section of this paper or some other analogous set of criteria as guidelines for the measurement of virtues would go a long way toward clarifying and establishing standards for virtue assessment in the community of scholars interested in this topic.

Criterion Validity

Good assessment measures allow us to predict other criteria. For example, a math aptitude test should be related to subsequent math learning and performance. In the case of virtues, eudaimonic theory suggests that virtue traits should predict the degree to which an individual experiences eudaimonia. A number of scales in The Virtues in Action Inventory correlate with eudaimonic well-being (Peterson, Ruch, Beermann, Park, & Seligman, 2007). There are a number of measures of eudaimonic well-being that can serve as criteria for virtue scales, such as the Psychological Well-Being Scale (Ryff, 1989), the meaning scale from the Orientations to Happiness Scale (Peterson, Park, & Seligman, 2005), the Eudaimonic Motivation

scale (Huta & Ryan, 2010), and the Questionnaire for Eudaimonic Well-Being Scale (Waterman et al., 2010). Because virtues are defined as the character strengths that make it possible to pursue worthwhile ends, virtue measures ought to predict aspects of goal pursuit, such as efficacy and success in goal pursuit (Palys & Little, 1983).

Another important criterion with respect to which virtue scales should be assessed is the degree to which an individual has a constitutive orientation to goal pursuit. This is important because of the contrast Aristotle (1999) makes in the opening paragraph of the *NE* between means-end activities and constitutive activities. In means-end activities, the means and the end are separable. The end is what is important and the means are disposable. In constitutive activities, the means partly constitute the ends, which renders the means valuable because it is inseparable from the end. Virtues are characteristically constitutive activities because the way that one acts virtuously is inseparable from the aim of acting virtuously. The more characteristically virtuous a person is, the more one would expect that individual to adopt a constitutive orientation to his or her goals (Fowers, Mollica, & Procacci, 2010).

A third category of criterion validity assessments could include behaviors. One would expect the trait of generosity to predict acts of kindness or helping, for example. It is important to keep in mind that, in general, relationships between self-reports and behavior tend to be quite modest (Paulhus & Vazire, 2007).

Convergent Validity

A third form of construct validity is the degree to which a measure correlates with measures with which it is theorized to relate. It is reasonable to expect virtue traits to be related to many well-established constructs. Demonstrating these relationships will help to

support the convergent validity of virtue measures. There are many studies of pro-social behavior, including helping behavior (e.g., Eisenberg et al., 2002), and virtues should be positively related to relevant pro-social behaviors. It is also reasonable to expect virtues such as generosity, justice, and friendship to be positively related to measures of relationship quality in friendships, family relationships, and romantic relationships. Individual differences in authenticity (Wood, Linley, Maltby, Baliousis, & Joseph, 2008) and moral identity (Aquino & Reed, 2002) should be related to virtues as well. Virtues should also be related to hedonic well-being scales such as life satisfaction and happiness. Some virtue measures have been found to correlate with hedonic well-being scales, including the Virtues in Action scales (e.g., Park, Peterson, & Seligman, 2004) and a kindness scale (Otake, Shimai, Tanaka-Matsumi, Otsui, & Fredrickson, 2006). Finally, measures that could indicate inclinations to vicious behavior, such as Machiavellianism (Christie & Geis, 1970), could reasonably be expected to have negative associations with virtues, particularly with honesty and justice. These are just a few suggestions about how the construct validity of virtue measures can be assessed in the area of convergent validity. This is potentially a very broad area of research, and each virtue would have content-related constructs with which it should converge based on the specifics of the virtue.

Discriminant Validity

Assessments can also be evaluated by gauging whether they are independent from measures of constructs with which they should be unrelated. A simple, but important example would be intelligence, which should be unrelated or only slightly related to virtue traits. It also seems reasonable that virtues should not differ across gender, race, ethnicity, or political or

religious affiliation. There are no doubt many other personal or demographic attributes that should be independent of the degree to which individuals manifest virtue.

Incremental Validity

One of the challenges of demonstrating the existence of virtues is that purportedly virtuous actions and relationships between self-reports of virtue and other measures can be attributed to third variables such as personality traits or situational factors. Incremental validity is the degree to which a measure captures some aspects of a phenomenon above and beyond other potential explanations for it. For example, it would be important to document that a measure of courage can account for risk taking above and beyond impulsiveness or thrill-seeking. There will be specific incremental validity issues for each virtue that need to be addressed to bolster our confidence that we are measuring the virtue. One question common to many virtue traits are whether they can be shown to have incremental validity with respect to personality scales.

Conclusion

The project of measuring virtue is a daunting one, partly because there has been and continues to be significant resistance to this project in the social sciences and in philosophy. Beginning with Allport (1921) psychology has viewed virtue as an inappropriately evaluative term, preferring a focus on personality, which was conceived as a descriptive rather than evaluative term. Recent developments in psychology has called this viewpoint into question by highlighting the evaluative aspects of the most prominent concepts of personality (Danziger, 1990; McCrae & John 1992; Richardson, Fowers, & Guignon, 1999). Many social scientists have become interested in empirically investigating virtues. This interest has been accompanied by

significant and intriguing developments in the conceptualization and measurement of virtue traits. So there is reason for optimism about the assessment of virtue.

At the same time, the project of measuring virtue is a daunting one, as I have outlined in this paper. The general construct of virtue has not been very well developed and there are many competing views. In addition, there are a number of individual virtues that have been studied in a somewhat haphazard and halting manner, with little consistency across research groups. There remain questions about how virtue traits relate to situational factors. Most of the research to date has been conducted with summary self-report scales that are seriously threatened by questions of veridicality. The construct validity of virtue measures remains largely unexamined for most virtue measures.

I have also outlined a number of reasons for hope in virtue assessment. There are robust conceptions of general and specific virtues from which researchers can draw. This conference is one manifestation of the interest in developing greater coordination and collaboration in virtue assessment. There is also increasing evidence that the situationist critique of virtue traits is unsustainable, opening the way for social scientists to believe that virtue is a worthy focus of study. Although self-report measures of virtue have important shortcomings, it is likely that this method will remain prominent among researchers due to its low financial and effort costs. I outlined four methods for dealing with veridicality issues in self-report measures that can reduce the degree of concern. In addition, there are other methods of virtue assessment, including other-report, experience sampling, and experimental approaches that are less subject to reporting biases. I also outlined how the assessment of construct validity of virtue measures can proceed.

I am cautiously optimistic about the prospects for valid and useful assessments of virtues. The success of this project will hinge on the degree to which a collaborative community of researchers emerges that can coalesce around a definition of the construct(s) to be studied and will persist in that investigative process. Because no single study or method can definitively document the existence of virtues, multiple studies with varying methods are needed, and this will proceed best if there is significant coordination and collaboration across research teams. In short, the requirement for virtue researchers is to practice the virtues of friendship, generosity, and justice in the study of virtue for the sake of the goods of knowledge and how that knowledge can help to create a better world. Although daunting, this project is one of the most worthy endeavors social scientists, educators, and philosophers can undertake.

References

- Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology* 20, 1-48.
- Allport, G. W. (1921). Personality and character. *Psychological Bulletin*, 18, 441–455.
- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83, 1423-1440.
- Aristotle. (1999). *Nicomachean Ethics*. (M. Ostwald, trans.) Upper Saddle River, NJ: Prentice Hall.
- Bäckström, M. (2007). Higher-order factors in a five-factor personality inventory and its relation to social desirability. *European Journal of Psychological Assessment*, 23, 63–70.
- Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, 43, 335-344.
- Barranti, M., Hawkins, A., Helzer, E., Fleeson, W., and Furr, R. M. (2013). Is morality in the eye of the beholder: Inter-judge consensus on moral traits.
- Baumeister, R. F., & Exline, J. J. (1999). Virtue, personality, and social relations: Self-control as the moral muscle. *Journal of Personality*, 67, 1165-1194.
- Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York: Guilford.
- Brewer, T. (2009). *The retrieval of ethics*. Oxford, England: Oxford University Press.

- Cawley, M. J., III, Martin, J. E., & Johnson, J. A. (2000). A virtues approach to personality. *Personality and Individual Differences, 28*, 997–1013.
- Christie, R. & Geis, F. C. (1970). *Studies in Machiavellianism*. New York: Academic Press.
- Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology, 51* (1), 171-200.
- Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research*. Cambridge: Cambridge University Press.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*, 71-75.
- Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge: Cambridge University Press.
- Dufner, M., Denissen, J. A., van Zalk, M., Matthes, B., Meeus, W. J., van Aken, M. G., & Sedikides, C. (2012). Positive intelligence illusions: On the relation between intellectual self-enhancement and psychological adjustment. *Journal of Personality 80*, 537-572.
- Eccles, J. S. & Wigfield, A. (2002). Motivational, beliefs, values, and goals. *Annual Review of Psychology, 53*, 109-132.
- Eisenberg, N., & Guthrie, I. K., Cumberland, A., Murphy, B. C., Shepard, S. A., Zhou, Q., & Carlo, G. (2002). Prosocial development in early adulthood: A longitudinal study. *Journal of Personality and Social Psychology, 82*, 993-1006.
- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*, 1011-1027.

- Fleeson, W. (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality, 75*, 825-862.
- Fleeson, W., & Gallagher, P. (2009). The implications of big-five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology 97*, 1097-1114.
- Fowers, B. J. (in press). A phenomenological reformulation of psychological science: Resources and prospects. In M. Altman & H. Pederson (Eds.) *The horizons of authenticity: Essays in honor of Charles Guignon's work on phenomenology, existentialism, and moral psychology*. New York: Springer.
- Fowers, B. J. (2012). An Aristotelian framework for the human good. *Journal of Theoretical and Philosophical Psychology, 32*, 10-23.
- Fowers, B. J. (2008). From continence to virtue: Recovering goodness, character unity, and character types for positive psychology. *Theory & Psychology, 18*, 629-653.
- Fowers, B. J. (2005). *Virtue and psychology: Pursuing excellence in ordinary practices*. Washington, DC: APA Press.
- Fowers, B. J., & Lang, S. (in press). Positive illusions. In D. Aldwin (Ed.) *Encyclopedia of mental health, (2nd ed.)*. New York: Elsevier.
- Fowers, B. J., Lyons, E., Montel, K., & Shaked, N. (2001). Positive illusions about marriage among married and single individuals. *Journal of Family Psychology, 15*, 95-109.

- Fowers, B. J., Mollica, C. O., & Procacci, E. N. (2010). Constitutive and instrumental goal orientations and their relations with eudaimonic and hedonic well-being. *Journal of Positive Psychology, 5*, 139-153.
- Greene, J. D., & Paxton, J. M. 2009. "Patterns of neural activity associated with honest and dishonest moral decisions." *PNAS* 106: 12506-12511.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35*, 603-618.
- Harman, G. (2009). Skepticism about character traits. *Journal of Ethics, 13*, 235-242.
- Hartshorne, H., & May, M. A. (1928). *Studies in the nature of character. Vol. 1: Studies in deceit*. New York: Macmillan.
- Hawkins, A., Facuillo, J., & Furr, R. M. (2013). Perception of benevolent traits: Significant agreement between self and others.
- Hawkins, A. J., Fowers, B. J., Carroll, J. S., & Yang, C. (2007). Conceptualizing and measuring marital virtues. In S. Hofferth & L. Casper (Eds.) *Handbook of measurement issues in family research*. Hillsdale, NJ: Erlbaum.
- Homiak, M. (2011). Moral character. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2011 ed.). Retrieved from <http://plato.stanford.edu/archives/spr2011/entries/moral-character/>

- Huta, V., & Ryan, R. M. (2010). Pursuing pleasure versus virtue: The differential and overlapping well-being benefits of hedonic and eudaimonic motives. *Journal of Happiness Studies*, *11*, 735-762.
- Kristjansson, K. (2013). *Virtues and vices in positive psychology: A philosophical critique*. Cambridge, England: Cambridge University Press.
- Kupperman, J. J. (2001). The indispensability of character. *Philosophy*, *76*, 239-250.
- Lefevor, G. T., & Fowers, B. J. (March, 2012). *Character and situation: How to transcend the traditional dichotomy as an Aristotelian*. Paper presented at the Mid-Winter Meeting of the Society for Theoretical and Philosophical Psychology, Austin, TX.
- Lefevor, G. T., Fowers, B. J., Ahn, S., Cohen, L., & Lang, S. (2013). *Is character plausible? A meta-analytic investigation of the strengths of situations on helping behavior*. Manuscript in preparation.
- 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., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175-215.
- Meindl, P., Jayawickreme, E., Furr, R. M., & Fleeson, W. (2013). *A foundation beam for studying morality from a personological point of view*. Manuscript in preparation.
- Molenaar, P. C. M., & Campbell, P. G. (2009) The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*, 112-117.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175-220.

- Otake, K., Shimai, S., Tanaka-Matsumi, J., Otsui, K., & Fredrickson, B. L. (2006). Happy people become happier through kindness: A counting kindnesses intervention. *Journal of Happiness Studies, 7*, 361-375.
- Palys, T. S., & Little, B. R. (1983). Perceived life satisfaction and the organization of personal projects systems. *Journal of Personality and Social Psychology, 44*(6), 1221-1230.
- Park, N., Peterson, C., & Seligman, M.E. P. (2004). Strengths of character and well-being. *Journal of Social and Clinical Psychology, 23*, 603–619.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598–609.
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224–237). New York: Guilford Press.
- Peabody, D. (1967). Trait inferences: Evaluative and descriptive Peterson, C., Park, N., & Seligman, M. E. P. (2005). Orientation to happiness and life satisfaction: The full life versus the empty life. *Journal of Happiness Studies, 6*, 25-41.
- Peterson, C., Ruch, W., Beermann, U., Park, N., & Seligman, M.E.P. (2007). Strengths of character, orientation to happiness, and life satisfaction. *Journal of Positive Psychology, 2*, 149–156.
- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Washington, D.C.: American Psychological Association.
- Richardson, F. C., Fowers, B. J., & Guignon, C. (1999). *Re-envisioning psychology: Moral dimensions of theory and practice*. San Francisco: Jossey-Bass.

- Robins, R. W., Tracy, J. L., & Sherman, J. W. (2007). What kinds of methods do personality psychologists use? A survey of journal editors and editorial board members. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 673–678). New York: Guilford Press.
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*, 351–57.
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation*. Philadelphia: Temple University Press.
- Ruch, W., Proyer, R. T., Harzer, C., Park, N., Peterson, C., & Seligman, M. E. P. (2010). Values in Action Inventory of Strengths (VIA-IS): Adaptation and validation of the German version and the development of a peer rating form. *Journal of Individual Differences, 31*, 138-149.
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology, 57*, 1069-1081.
- Sreenivasan, G. (2002). Errors about errors: Virtue theory and trait attribution. *Mind, 111*, 47-68.
- Taylor, C. (1989). *Sources of the self*. Cambridge, MA: Harvard University Press.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social-psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

- Veldorale-Brogan, A., Bradford, K., & Vail, A. (2009). Marital virtues and their relationship to individual functioning, communication, and relationship adjustment. *Journal of Positive Psychology, 5*, 281-293.
- Veldorale-Brogan, A., Lambert, N. M., Fincham, F. D., & DeWall, C. N. (2012). The virtue of problem-solving: Perceived partner virtues as predictors of problem-solving efficacy. *Personal Relationships, 20*, 511-523.
- Waterman, A. S., Schwartz, S. J., Zamboanga, B. L., Ravert, R. D., Williams, M. K., Agocha, B., Kim, S. Y., & Donellan, M. B. (2010). The Questionnaire for Eudaimonic Well-Being: Psychometric properties, demographic comparisons, and evidence of validity. *Journal of Positive Psychology, 5*, 41-61.
- Why, Y., & Huang, R. (2011). Positive illusions and its association with cardiovascular functions. *International Journal of Psychophysiology 81(3)*, 305-311.
- Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and empirical conceptualization and the development of the authenticity scale. *Journal of Counseling Psychology, 55*, 385-399.