

Running Head: CHALLENGES FOR PERSONALITY PARADIGM

Cultural, Contextual, and Measurement Challenges
for the Paradigm of Personality Science

Gerard Saucier and Amber Gayle Thalmayer

University of Oregon

Draft of chapter for publication in *The Handbook of Personality at Work* (N. Christiansen & R. Tett, Eds.). New York: Routledge.

Address correspondence to:
Gerard Saucier
Department of Psychology
1227 University of Oregon
Eugene OR 97403-1227
gsaucier@uoregon.edu
tel. 541 346 4927
fax 541 346 4911

August 2012

The previous chapters in this volume each engage with aspects of personality science. In this chapter, we focus on personality science itself. We advance the position that contemporary personality science presents a dual picture. In part, personality science has accrued a stock of well-accepted knowledge about patterns in human behavior that make for a sort of consensual “paradigm.” But there are also major unresolved matters. On these disputed matters, one can delineate opposing views: a received and conservative but probably not entirely correct view, and an alternative and seemingly radical view that is probably just as likely to be correct. Reflecting the lively relevance of industrial-organizational psychology for core personality science, numerous chapters in this volume intersect with aspects of this alternative view.

The Consensual Paradigm

There are some things we know about personality. Humans vary psychologically from one another in ways that make different individuals best suited to different social roles (Hogan, 1983; Hansen, this volume; Murphy, Deckert, & Hunter, this volume). Individual psychological differences have likely been maintained in our species due to trade-offs between costs and benefits for different levels of reactivity and behavioral tendencies (Nettle, 2006). The importance of such individual psychological differences to people and societies everywhere is made clear by the ubiquitous presence of words to describe personality attributes across languages and cultures (Saucier, Thalmayer, & Bel-Bechar, 2012). While there are cultural differences in the extent to which individuals are likely to describe themselves using attributes (Heine, 2012), people everywhere describe others with attribute terms indigenous to their native language. This likely speaks to the universal human lifestyle of functioning in groups (Dunbar,

1998), and the ubiquity of differences in individuals' abilities to fulfill specific roles (Hogan & Blickert, this volume).

Attempts to make sense of the ways that people differ from one another extend into ancient history, for example in catalogs of virtues, systems of astrology, or Galen's four humors. Early in the twentieth century, such efforts were framed in clinical terms, for example in Carl Jung's work on personal growth through the expression of different aspects of typology (Reynierse, this volume). For nearly 100 years, there have been attempts to fashion empirically grounded personality inventories to capture broad ranges of attributes (Prewett, Tett, & Christiansen, this volume). But the role of experts in determining which content to include inevitably led to some biases, for example an undue emphasis initially on aspects of abnormal psychology (e.g. the MMPI).

The lexical rationale, that the most salient differences between persons will be encoded in the natural language, established an empirical procedure that could be deployed across cultures to help identify the most important dimensions of personality attributes (Goldberg, 1981; Saucier, 2009). Lexical studies have helped us to understand the ways that diverse attributes group together into higher order dimensions. And after evidence from the first handful of languages was available, they led to the establishment of the useful Big Five model (Digman, 1996; De Fruyt & Wille, this volume), which is the basis of the popular NEO-PI-R (Costa & McCrae, 1992) and other inventories. Lexical studies have also spawned other models, for example the six-factor structure of the HEXACO-PI (Lee & Ashton, 2004). Although it is debatable whether these lexicon-inspired structures are sufficient as models of personality (e.g., Block, 1995, 2010), and whether the cross-cultural replicability standard they imply is too strict (e.g., Rolland, 2002), there is little debate that they provide a central reference-point for the field.

We know that personality attributes, operationalized as scores on personality inventories (but likely generalizable to other sources of data), have the capacity to predict career choices and outcomes (Murphy et al. this volume). Indeed, personality attributes have predictive ability for virtually every important life outcome studied, including health, longevity, marriage and divorce, social relationships, well-being, and creative expression (Ozer & Benet-Martinez, 2005). They appear to have the same predictive ability as SES or cognitive ability with respect to mortality, divorce and occupational attainment (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), and they “predict job performance as well as any other procedure, and...outperform most other predictors” (Hogan & Blickert, this volume, p. X).

There are widely accepted criteria for evaluating personality measures; many of these consensual criteria are reflected in the review of Prewett, Tett, and Christiansen (this volume). Measures built on some combination of empirical and rational (or theoretical) considerations are judged favorably. Assessments are now customarily presented in terms of positions on a continuum rather than in terms of binary “types.” Moreover, validity and reliability are consensually regarded as each having some importance.

Personality psychologists would show wide agreement on this point as well: Individual psychological differences can be assessed by more than one method. Drawing on an earlier classification by Cattell, Funder (2010) distinguishes between (S)elf-report, (I)nformant, (L)ife outcome, and (B)ehavioral data. Data gathered in different ways can importantly enrich our understanding of individuals, and might help us make measurement more cross-culturally applicable (Heine, Buchtel, & Norenzayan, 2010).

Contemporary personality science evinces trust in self-report scores on personality variables, even if expert positions range from unconditional trust to a more suspicious “trust but

verify” stance. To illustrate this range, the issue of faking on personality tests is mentioned as a concern by several authors in this volume (e.g. McFarland; Prewett et al.), whereas Hogan (this volume) argues that to the extent that people exaggerate or diminish their characteristics, they primarily display their awareness of social expectations and their ability to conform to them. Hogan’s argument is consistent with findings (Uziel, 2010) that impression management is associated with well-being, job performance, health and interpersonal adjustment.

Most accept that there will be cultural differences in how personality operates. McCrae (2002), for example, documents that average variances on NEO-PI-R scores are “geographically ordered” with European populations having higher variability than Asian populations. Similarly uncontroversial are findings of cultural differences in the extent to which personality attributes are seen as fixed and enduring, versus malleable and influenced by the situation (Kanagawa, Cross, & Markus, 2001; Heine, 2012). In individualistic cultures, it is considered desirable to have a consistent identity across situations – in fact, personality consistency is associated with social skills, being well-liked, and well-being (Suh, 2002). But in collectivistic cultures, being consistent across situations is not associated with such rewards (Suh, 2002). Instead, people in less individualistic societies appear to value adapting their behavior and identities to the role they play -- clinging to a single idea of the self, regardless of context, is viewed as a lack of social skills or maturity. The theoretical framing for these findings of cultural differences is not, however, consensual. Prominently, there is no agreement on whether anything about the cultural environment actually has a *causal* effect on personality tendencies.

As Nye and Roberts (this volume) point out, current evidence suggests that personality attributes are characterized both by stability and by change over time. The consensus is that people everywhere have *some* consistency in their behavioral tendencies, and that change can

also occur. What remains subject to considerable debate and controversy is what the real sources of change in scores over time might be. McCrae (1993) suggests they may be due to little more than measurement error and maturation, whereas Nye and Roberts suggest substantive sources of change.

With these last few points, the reader can see some signs of dispute in the field, beginning to creep in about the edges of our account. In the next section, we describe a useful conceptual framework for understanding scientific controversies more broadly. We then turn to deal directly with main lines of dispute within the field.

Paradigms and Anomalies

Kuhn (1970) provided a much-cited account of paradigms in science, and of some processes by which paradigms shift over time. Briefly, a “paradigm” consists of notions shared by scientists in a field about what phenomena should be studied, how they should be studied (with what methods and/or equipment), what questions scientists should be answering, and how results are to be interpreted. It is basic to the shared worldview or mindset of scientists in a field at a particular time. Arguably, in our view, such paradigms are a wider phenomenon within the ideational worlds humans create and inhabit. That is, something like paradigms are not unique to science but might be observed in cultural, political, and religious models of the world as well (as in Wallace, 1961; Westen, 1985). But scientific paradigms differ from such non-scientific models in an important way: Scientific paradigms are explicitly open to revision, and indeed are presented as tentative and potentially imperfect models.

Imperfections within a paradigm are part of the mechanism of change. Certain observations (or findings) do not fit the paradigm, are “anomalies” for it. As awareness of anomalies increases, so will dissatisfaction with the paradigm. When an alternative – or

“revised” -- paradigm can account for the anomalies, one might then see a paradigm shift, whether this shift be rapid, gradual, or part of a generational shift.

In the Kuhnian approach, then, a fairly typical situation in a scientific field is this: One finds areas of consensus, and a received, popular model that incorporates this consensus but also includes other elements that are more debatable. The ultimate fate of the received model depends on the quantity and importance of the anomalies that build up around it as scientists attempt to apply it.

It can be debated whether personality science has a paradigm in the same sense that physics does. There is less commonly held theory here than one sees in the hard sciences, and arguably the field is a collection of small quasi-paradigms vying for a higher status (a characterization that may fit social psychology even more squarely than personality). We leave that issue aside here, and use the term ‘paradigm’ to cover the consensual as well as the most popularly shared notions in the field (i.e., the received model).

In this chapter, we lay out first what we perceive to be the disputable aspects of the received model and next the major anomalies that arise with respect to them. Then, we suggest what a radically different model might look like, a model opposed to the received model in as many respects as possible, but possibly able to deal better with the anomalies. We stress that, in fact, we do not know of anyone who has proposed *in toto* this exact “negative benchmark” model, knitting together all the elements we suggest it might have. But we identify some proposals found in this volume that go in the direction of doing so. Our aim in organizing views in this way is to stimulate competition between reasonable models, and improve personality science. Obviously, it is possible that the outcome of such competition may be a hybrid that includes the best features of each view, while excluding the least supportable features of each.

The Received View in Personality Science

There are a number of now commonly held assumptions in personality science that may not be entirely justifiable. The popularity of most of these assumptions date back a few decades, to what was arguably a previous era of rapid paradigm shift in the field. This era was associated with the resolution of the personality-situation debate and the widespread adoption of a five-factor model. The following are widely held but debatable assumptions.

1. *Questionnaire responses reflect disinterested, factual descriptions of persons' attributes.* On this assumption, when researchers and assessors present individuals with items describing some range of tendencies or attributes, with a request to describe the self or some other person, respondents oblige by providing the most truthful account they can. That is, they take the opportunity to file a realistic, veridical report (like some journalist or scientist) about how things are, and not about how they ideally wish things to be. Support for this assumption is provided by evidence for *some* accuracy in personality description (Funder, 1995), and for seeing response biases such as social desirability as a sort of “red herring” (Ones et al., 1996). Based on this assumption we can interpret questionnaire responses as pure indicators of traits, of truly objective basic tendencies of persons, with no need to adjust the method or monitor response protocols to deal with non-veridical responding.

2. *Self-report is the primary data source, and so the main use of informant data is to demonstrate the validity of self-report data.* To support this assumption, one can argue that the self is the best expert on the person's tendencies and attributes, having exposure in all times and situations (except when asleep or perhaps extremely intoxicated). Informants (that is, knowledgeable acquaintances of the person targeted in the description) at best have a far more limited exposure, so should be regarded as secondary sources; informant reports can be (and

frequently are) used to establish that self-reports have validity. This assumption, if well-founded, would be a tremendous practical boon to personality science, because self-reports are quick and convenient to obtain. It can be linked to a very long tradition in psychology (going back to Wundt and Ebbinghaus) of taking introspection quite seriously; consistent with that tradition, the assumption is that much of the truth of personality can be introspected by the possessor of the traits.

3. *Whatever the target of description, all respondents (no matter the context or culture) compare the target to the same objective standard.* In their initial instructions, some assessment instruments ask the respondent specifically to compare the target (usually self) to some kind of typical person. Therefore, one would be high on a trait if higher than typical, low on a trait if lower than typical. Other instruments leave the reference-group implicit -- that is, they do not specify what the comparison or reference group should be. This seems more common where the items consist of statements with which one can variously agree or disagree. In either case, the context is usually unspecified, and there is no recommendation to compare oneself to members of other populations and cultures (which would of course pose challenges for those who have not travelled the world). Essentially, it is assumed that trait levels exist in an absolute sense and not mainly in a comparative sense, so that ratings will not shift if the reference group is shifted.

4. *People normally behave consistently across situations, so response instructions regarding contexts are unnecessary.* In instructions for completing personality assessment, respondents are often asked to generalize across the widest range of instances, if they are given any instructions at all on the matter of situations. Evidence that there is some cross-situational consistency might be taken to justify ignoring situations, either in items or instructions. If so, broad and sometimes ambiguous trait-descriptive adjectives applied to someone's behavior

across all situations would be an acceptable (in fact, admirable for their efficiency) mode of assessment. As an extension, people might be assumed to behave consistently across time, so that (in a sample being assessed across time) the time of administration of the measure becomes immaterial, as per the next assumption.

5. *The traits emerging from personality assessment are hard-wired and essentially unchangeable “essences of persons” unaffected by culture and context.* Supporting this assumption, at least to a degree, personality attributes measured by contemporary questionnaires are usually found to be substantially heritable and not much affected by variation between environments shared within families (Turkheimer, 2000). They are also routinely found to be relatively stable across time (Caspi, Roberts, & Shiner, 2005). It has been difficult to identify the sources of across-time variation in these attributes, such that it has appeared reasonable to attribute most of the observed across-time variation to measurement error or to generic maturation processes (McCrae, 1993; Costa & McCrae, 2006). Taking all these findings together, it has seemed reasonable to posit that personality traits are basic tendencies influenced by biology but not by culture, context, or other aspects of environment (McCrae & Costa, 1999).

6. *The general nature of these “essences of persons,” and their structure/organization, can be identified suitably in European-origin populations, and generalized to the rest of the world.* Many hard-wired psychological processes, like those involved in sensation and psychophysical phenomena, would not be expected to vary much between populations around the world, so it would not matter much which population is sampled to develop a model. More debatable is the extension of this assumption to social behavior and to the personality domain. Some evidence does suggest this to be reasonable. Personality measures developed in the West

(e.g., the NEO-PI-R, developed in the American context) do appear to be reasonably translatable and able to function rather well in at least many cultural contexts.

7. *Studies of the natural language of personality in diverse cultural settings have supported a structural model with five factors.* This position is easy to support from literature appearing in the early 1990s, when the “diverse cultural settings” in which studies of personality language (lexical studies) had been conducted included only America, Germany, and the Netherlands. One could add citations to select later studies -- a couple of Slavic-language (Polish, Czech) studies and of studies of Italian and Turkish descriptors – to support this view. Certainly, converging results from seven languages must be taken seriously. And perhaps even more seriously because two of those languages – German and English – have been the prime scientific languages of psychology in its (approximately) 130-year history, these being thus “preferred languages” in another way.

8. *If one measures those five factors, one has a sufficient measure of personality.* This assumes, in part, the finality and comprehensiveness of a five-factor model. It assumes that additional factors beyond these five are unnecessary. It might also be taken to assume the unimportance of that specific variance not captured by the broad factors but captured uniquely by facets or subcomponents of these factors (for different views, see O’Neill & Paunonen, this volume; Wood, Nye, & Saucier, 2010).

9. *A top-tier indicator of measurement quality is coefficient Alpha (internal consistency), referencing convergence among indicators.* Reliability is by definition the absence of measurement error, and this is certainly one good index of the quality of a measure. The most convenient (and thus, predictably, the most commonly cited) index of reliability involves measuring how consistent is the scale (or construct) across subsets of its items (or indicators). Of

course, reliability is no end in itself. But the other commonly cited index of measurement quality, validity, is assumed to have reliability as a prerequisite, with higher reliability enabling higher validity. So a central defining quality for any trait measure will be its level of internal consistency (as measured with coefficient Alpha). Because internal consistency (like any form of reliability) is often defined in terms of the absence of (measurement) error, one might believe that highly internally consistent measures should give a better approximation to the truth. And, by this reasoning, the prime criterion in scale development should understandably be the maximizing of internal consistency. In scale development, factor analysis gives a guide toward maximizing internal consistency (i.e., select that set of items loading high on a factor, and reject the rest).

10. *In evaluating a personality scale or inventory, it is enough to test against chance rather than making rigorous comparison to competing scales or inventories.* In examining the validity of a scale or inventory, the key things are absolute indices, for example, whether self-ratings produce significant correlations with ratings by others, or whether there is significant capacity to predict external outcomes. Once these significant effects are established, the scale/inventory is “validated.” If competing scales or inventories are to be studied, they are primarily useful in showing that the present measure measures what it purports to measure (e.g., a scale labeled “sociability” correlates with a previous scale labeled “sociability”).

We have presented the foregoing ten assumptions consecutively and without objection, so as to afford the reader a chance to sense how they fit together. Let us encapsulate them as a linked theoretical narrative. Factual information about persons’ traits is available from self-report questionnaire responses, wherein respondents use a common objective standard to reference their various cross-situationally consistent patterns. The factuality of traits scored from these

inventories is demonstrated by the correlations of these self-ratings with those by informants, as well as by indications of the traits' genetic heritability. These biologically based traits are relatively impervious to context and culture. So, the nature of these traits as observed in a Western setting (e.g., crucially, in a five-factor array) gives strong indication for how they will be observed anywhere. A key measurement quality is one aided and abetted by factor analysis -- internal consistency -- although documenting the validity of a scale or inventory (though not necessarily the comparative validity in relation to other scales/inventories) is also important. One can see the phrase "internal consistency" as a hallmark throughout this received view; behavioral patterns are consistent across contexts and cultures, stemming from traits not only internal to the person but internal to the genome, and internal consistency on the psychometric level is also highly valued.

For those well-acquainted with contemporary personality science, these assertions are unremarkable. Indeed, despite their debatability, they have the quality of being obvious conventions. As a result, a speech at a scientific meeting that laid out these points would likely induce considerable boredom in the audience. The alternative points of view we describe next, however, go against convention, and may jar the gentle reader out of his or her chair.

A Strong Alternative View

Here we define an alternative view, by aggregating sensible alternatives to the ten assumptions just listed. In a sense, we are using the ten assumptions as a "negative benchmark" and laying out an alternative in systematic contradistinction. Doing so is important because, as pointed out below, the received view has generated anomalies and the sensible alternatives may take better account of evidence, avoid these anomalies, and therefore be building blocks in a better paradigm.

Taken together, these alternatives form a patchwork construction, a loosely strung-together set of potentially viable alternatives to elements in the received view. After delineating them, we discuss the degree to which they form a coherent set.

1. *Personality descriptions are a compound of fact and value, but this tends not to compromise the validity of scores.* If personality ratings were disinterested and factual, reports by different observers would strongly converge, socially desirable responding would not be a concern, and personality language would contain mainly the evaluation-free descriptors necessary for purely objective descriptions. But accuracy in personality ratings is far from perfect; there is a rather surprising degree of divergence between ratings of the same target, between self and various observers (Clifton, Turkheimer, & Oltmanns, 2005). This suggests that value preferences of the respondent influence responses. These value-preferences may lead the respondent to enhance the favorability of the description (so as to match an ideal or the way things ought to be), or to diminish its favorability (which may be overall in line with an ideal of modesty). Thus, it is unsurprising that personality descriptors in the natural language form a clearly bimodal distribution with respect to favorability: Most descriptors are either clearly favorable or clearly unfavorable (see, for example, Goldberg, 1978, Figure 1). The same is true of questionnaire items, attributes, or scale names. Personality descriptions as a rule evaluate, that is, invoke a value. Descriptions lie somewhere on a continuum between purely objective description and description in terms of how one ought to be or would ideally be. The fact versus value quotient in personality descriptions clearly depends partly on the assessment situation; where high stakes are involved, responses tend to be different (Levashina & Campion, 2006), as there is more motivation to assert value over fact. Socially desirable responding certainly does occur; however, to competently assert value requires knowledge of the consensual relative value

associated with a description. Those individuals who, more than others, choose to assert value over fact and provided a biased description may be cheating (a bad thing) but are also demonstrating their knowledge of social norms (a good thing) -- there may be others who attempt to cheat but fail to do so effectively, because they are unable to communicate a fully desirable impression, not fully knowing what one would look like. As a result the presence of this tendency overall tends not to compromise the validity of scores, and controlling for socially desirable responding does not generally improve predictive validity (consistent with Ones et al., 1996).

2. *The prime source of data for human behavioral patterns, core features of personality, is informants (knowledgeable acquaintances).* If self-reports were the best type of data, they would predict outcomes better than other types of data, in any cultural context. But there are indications (see Connelly, this volume) that reports by knowledgeable acquaintances (informants, observers) are more predictively valid than self-report, at least for many kinds of attributes (i.e., excluding those that give particular primacy to a subjective viewpoint, such as affective states, and beliefs and attitudes). Informant data offers some inherent advantages. First, it is possible to aggregate raters and thus improve the reliability of the joint ratings. Second, diverse informant reports about the same individual can be heterogeneous, which would lead to less favorable enhancement than in self-reports. Moreover, informant reports appear to be a more universally generalizable mode of personality assessment; there are parts of the world where self-report is a relatively unnatural activity, whereas talk about others (including gossip) appears to be more universal. Therefore, it is reasonable to give priority to informants, especially those with extensive knowledge of the subjective states and preferences of the target of description, or as Connelly (this volume, p. 7) puts it, “unique access to internally held information.” On this

view, for tapping behavioral patterns relevant to work and other settings, self-report data is useful but secondary.

3. *Standards of comparison for self- and informant-ratings vary by culture and context.* If respondents in all settings used the same objective comparison standards, expert ratings of characteristics of various populations should converge with aggregated ratings of laypersons. However, this has not consistently occurred. Experts judge East Asians to be more collectivistic than North Americans, but cross-cultural comparisons failed to reveal this pattern (Heine, Lehman, Peng, & Greenholz, 2002). Expert views of national characteristics have been found uncorrelated with means of samples from different nations (McCrae & Terracciano & 2006). And counterintuitively, gender differences in personality have been found to be greater in European than in non-Western populations, despite greater gender equity in the European settings (Costa, Terracciano, & McCrae, 2001). As Heine (Heine et al., 2002; Heine, Buchtel, & Norenzayan, 2008) has pointed out, discrepancies like these can easily arise due to reference-group effects. Populations differ in their standards of comparison. A 5-foot-6 male might be considered tall in Malaysia but short in Norway; it may be more difficult to reach the standard of being judged 'very hard-working' in a high-pressure achievement-oriented society than in a low-pressure society where there is more focus on enjoying life. Another example comes from Nye and Roberts (this volume): a college student may see him or herself differently in comparison to peers before and after getting a first job; the peer-group standard changed. Standards of comparison for ratings should be expected to vary by culture and context. Given this hazard, the best approach is to use multiple methods including objective behavior-count data, and to use clearly anchored and easily translated response scales that, as much as possible, require people to count (e.g., how often this happens, if not specifically how many times it has) or make an

explicit choice between competing alternatives. These approaches should generally serve to suppress or eliminate reference-group effects (Heine, 2012).

4. *Personality tendencies in part operate across situations and in part are situation-specific.* Although there is evidence for both, the evidence for consistency (stability) across time is more compelling than the evidence for consistency across situations (Mischel & Peake, 1982). There are anomalies for a strict assumption of cross-situational consistency. There are moreover important context-specific traits (e.g., coping and emotion regulation). Indications that self-descriptions are affected by the stakes in the assessment situation is another indicator of situational specificity. Some personality characteristics are best observed and assessed in specific contexts (a key premise of Trait Activation Theory; Tett et al., this volume; for example, extraversion at parties or purely social events, or conscientiousness from interacting with demands at work or at school. In sum, the premise of cross-situational consistency can easily be taken too far. A better working assumption would allow for both cross-situational consistency and situation-specificity in behavioral patterns.

5. *Although personality tendencies are partially genetically based, they represent in part “software” subject to influence by cultural and contextual factors.* If personality traits were hard-wired and unchangeable essences unaffected by culture and context, scientists would be unable to detect coherent sources of personality change in the environment, in the roles individuals adopt, or the belief/value systems that exist in cultures or subcultures. Systematic effects of environments, roles, beliefs, values, and other cultural components are anomalies for a “hard-wired essentialism” view of traits. But such anomalies are observed. For example, personality change can be predicted by individuals engaging and investing in new roles (Lodi-Smith & Roberts, 2012), antecedent variation in attitudinal beliefs (Saucier, 2012), and by the

extent to which individuals value specific attributes (Weisberg, De Young, & Simpson, 2012). Moreover, bicultural individuals may answer questionnaires significantly differently according to which of their two cultural identities is made salient (Hong et al., 2000). These anomalies indicate systematic patterns in how personality changes.

6. Models of personality dispositions bear some imprint of the cultural setting in which they were developed. If the nature and structure of personality dispositions were identified well enough in studies of European-origin populations, then generalizing to the rest of the world would pose no problem. It has been difficult, however, to establish measurement invariance across cultures for personality measures (Allik et al., 2012; Poortinga, van de Vijver, & Van Hemert, 2002). Reference-group effects (described above) and differential self-enhancement tendencies (Xie, Roy, & Chen, 2006) make comparisons more difficult. Studies of the natural language of personality indicate that, when the indigenous structure of various lexicon are investigated independently and then compared, only a core of one or two factors truly arises independently in any language (Saucier et al., 2012), the rest being more subject to variation. Personality models embody, almost inevitably, some cultural presuppositions. The best way to eliminate this ethnocentric bias is to develop “culturally decentered” models and measures by anchoring them in investigations that from the beginning involve a culturally, internationally diverse set of populations. This kind of development process is exemplified by the model of “social axioms” (Leung et al., 2002).

7. Studies of the lexicons of languages do not necessarily support the contemporary five-factor model of personality, although studies of languages having their origins in northern Europe do. If studies of the natural language of personality in diverse cultural settings supported a structural model with five factors, then studies of the natural language (lexical studies), no

matter the language or culture, should lead to the same five factors by an equally direct route. A less demanding, but still relevant, test would involve showing that translated measures of these five factors work about equally well in all populations. Neither of these projected outcomes has materialized. Approximately half of the lexical studies to date have not yielded the full set of five factors where one would expect them, in the five-factor solution (Saucier & Simonds, 2006). And those finding something like a Big Five structure typically diverge in some way from the American template; Agreeableness and Intellect/Openness factors seem especially prone to vary between studies. A factor fully characterizable as 'Openness to Experience' has been difficult to find in most languages (McCrae, 1990), although content related to originality (Saucier, 2009) or unconventionality (Ashton et al., 2004) arises with more frequency. Once one moves to a broader base of languages, beyond those with origins in northern Europe, a six-factor model (Ashton et al., 2004) appears more replicable than the Big Five. Furthermore, when one takes a popular questionnaire measure of the five factors and translates it to numerous languages, one does see some variation across cultural settings in structure, and some attenuated reliability for some scales (Piedmont et al., 2002; Rolland, 2002). Failures to establish full measurement invariance across populations for this questionnaire (Allik et al., 2012; Poortinga et al., 2002) suggest that comparison of means between culturally different populations is hazardous. Given these numerous anomalies, the Big Five structural configuration should be regarded as somewhat culture-specific. Studies of lexicons appear to converge on no more than two large personality factors (Saucier et al., 2012) that, given their ubiquity across languages, might be less subject to cross-cultural variation. These "Big Two" factors (social self-regulation and dynamism) map onto well-established higher-order factors of the Big Five (Digman, 1997; De Young, 2006). And they yield a more parsimonious and theoretically tractable model than the Big Five; the

obvious drawback is that models with five or six factors (not to mention 16 or 30 “facets”) should routinely be superior in prediction contexts (Saucier et al., 2012). The Big Two cap the structural hierarchy of personality dispositions, but the best practical work is probably done much farther down the hierarchy, indeed perhaps at a level well below any Big Five or Six (O’Neill & Paunonen, this volume).

8. *Five factors capture a great deal of the personality-attribute domain, but it takes more than five factors to be comprehensive of this domain.* If the five factors were a sufficiently comprehensive content model for personality attributes, several outcomes would be observed. No characteristics of importance would be found outside the five content domains (i.e., having a low multiple correlation with the five factors). No model with more than five orthogonal factors would be workable (any additional factor would fall apart under tests or scrutiny). And measures of more specific traits would be easily subsumable under one or more of the five factors, with specific subcomponents (or “facets”) adding little to prediction beyond what the five factors offer. The reality is otherwise. It is not difficult to detect some important characteristics beyond the Big Five (Saucier & Goldberg, 1998; Paunonen & Jackson, 2000). Six-factor models (Ashton et al., 2004; Hogan & Hogan, 1997; Jackson, Paunonen, & Tremblay, 2000) appear quite workable. The more specific “facet” level does appear to account for substantial variance beyond the broad factors (O’Neill & Paunonen, this volume). These outcomes are anomalous for a strong five-factor theory. The alternative view leaves open how many factors would be comprehensive. To take account of points 6 and 7 above, we should bear in mind that a comprehensive model for one language and cultural setting may not exactly match the comprehensive model for another. There is likely culture-specific content that is important and necessary in one part of the world, but not in another. As we increase comprehensiveness, we are almost certainly decreasing cross-

cultural replicability and compatibility. The value of parsimony (and theoretical clarity) pushes us in one direction toward fewer factors, whereas the value of comprehensiveness pushes us in the other direction, toward more factors. It is not impossible that decades from now scientists will think of personality as more than five types of attributes (perhaps even more than 16 or 30) theoretically organized at the broadest level into less than five major groupings.

9. *Internal consistency is a useful, but not a top-tier, indicator of measurement quality.* If internal consistency (indexed by Cronbach's coefficient Alpha) were truly the best standard for evaluating measurement quality, on par with validity, higher internal consistency would guarantee monotonically higher validity. No alternative measure of homogeneity would prove particularly useful. Against this assumption, experts (including Cronbach [Cronbach & Shavelson, 2004] himself) have come to critique an "abusive" overemphasis on internal consistency (Schmitt, 1996). Thorndike (1967, p. 214) commented that "exclusive preoccupation with item internal consistency may lead to an undue narrowing of the scope" of a measure, a decrease of predictive validity consistent with the attenuation paradox identified by Loevinger (1954). In fact, short, content-heterogeneous scales often perform as well as longer more homogeneous scales, despite much lower internal consistency (Thalmayer, Saucier, & Eigenhuis, 2011). Overemphasis on maximizing Alpha easily leads to redundancy or "parallelism" in scales that attenuates rather than enhances predictive capabilities (Loevinger, 1954). Another slight to the cult of internal consistency comes from Item Response Theory, which not only gives importance to heterogeneity in item-difficulty levels (so as to measure effectively at diverse levels of the construct), but sets unidimensionality rather than high internal consistency as a prerequisite for model estimation. Indeed, unidimensionality offers more theoretical clarity: It communicates that a single interpretation can be given to the set of items, as they converge on

one thing only. Internal consistency, in contrast, merely tells us the degree to which test variance is attributable to something beyond specific item variance (i.e., item uniqueness, which is treated as measurement error); that “something” could be many things (i.e., multidimensional; Cortina, 1993). Retest stability may be more informative than internal consistency, in that it can be applied even to one-item measures or linear composites. Presented with a choice between a reliable measure and one that lacked reliability but had much stronger validity, most professionals would choose the latter.

10. *The most important test of the value of a measurement instrument comes when it competes against other, partly similar measurement instruments.* Prewett, Tett, and Christiansen (this volume) observe that there are many inventories in the marketplace, but few studies comparing them. Goldberg has pointed out that the field of personality assessment has no “consumer’s union” to test competing products against each other and make recommendations about which to use and which to avoid. From a commercial standpoint, comparative validity tests might be a hazard to be avoided: Odds are good (at least one in two) that some other instrument(s) will do better than yours, which could have deleterious impacts on future marketing – why take the risk? The current norm, that personality scales and inventories are developed with little testing against other measures, allows each instrument to exist in a little world of its own. Fortunately, some recent studies cut against the grain. Grucza and Goldberg (2007) compared 11 multiscale inventories for their capacity to quantitatively predict diverse outcomes. Thalmayer, Saucier, and Eigenhuis (2011) and Credé et al. (2012) each compared brief and medium-length inventories (at least eight in each study). Going beyond comparative prediction, Prewett et al. (this volume) compare 10 inventories in frequent use in I/O settings with respect to six desiderata (including three approaches to validity). Such comparisons are

uniquely informative, and will ultimately strengthen the field by helping to optimize and standardize measurement practices. It becomes easier to decide which is the fastest horse if the horses are actually permitted to run on the same track at the same time.

The Alternative View: Less Coherent Than the Received View?

Let us now summarize the major points in the alternative view (see the right column of the Table). Descriptions of personality patterns are a compound of fact and value. A prime source of data regarding these patterns is knowledgeable acquaintances. Standards of comparison for self- and informant-ratings vary by culture, and may even vary between situations. Although personality tendencies in part operate across situations, in part they are situation-specific. Although personality tendencies are partially genetically based, they represent in part “software” subject to influence by cultural and contextual factors. Therefore, models of personality dispositions – regarding their structure, their measurement, and their associations – inevitably bear some imprint of the cultural setting in which they were developed. As a prime example, studies of the lexicons of languages do not necessarily support the contemporary five-factor model of personality, although studies of languages having their origins in northern Europe do. Five factors are acknowledged to capture a great deal of the personality-attribute domain, but not all of it. With respect to measurement, unidimensionality and retest stability are as important as internal consistency, and validity is even more important. Rather than evaluating each measure in an isolated way, on its own, the most important test of the value of a measurement instrument comes when a well-controlled study, or rigorous review, forces it to compete against other, partly similar measurement instruments.

What do these pieces have in common? Of what common theoretical narrative might they be a part? There are recurrent key concepts: fluidity, standards, comparisons, complexity, bias

reduction, and differentiation of central from peripheral components. In what makes the alternative view distinct from the received view, several kinds of themes can be discerned.

In this alternative view, basic elements of personality are seen as more fluid, subject to variation and “imprints” based on the group and environment in which they occur. Fundamental to this variation are “standards,” because different groups and environments have different standards. Human behavioral patterns are seen as partially a product of the cultural and cognitive models that incorporate these standards, product of their “software” as well as their “hardware.” On this view, personality becomes more anchored in the relational world, with less attempt to reduce it to a core set of biological processes. What Hogan and Blicke (this volume) call “reputation” comes to the fore. By “reputation” here we mean something of a proposed agreement about how the facts of a person’s behavioral history comport with a set of value/evaluative standards, more than the results of “impartial jury deliberations” trying to arrive at the objective truth about someone. Thus, the proposed view of personality and character is more congruent with how it might be observed in traditional societies: the output of convergences in gossip-like processes, in which values are asserted as well as facts, and in which the purported facts exist against a background of asserted or implied values.

In this alternative view, there is an overt attempt to escape from ethnocentric bias in a rapidly globalizing world in which people have more exposure to cultural patterns outside their own, and in which there is no guarantee that the Western way (which has dominated psychology heretofore) will always be pre-eminent. This approach avoids making the personalities of Western, educated, industrialized, rich, democratic people (“WEIRD” as per Henrich, Heine, & Norenzayan, 2010) the central concern of personality psychology. Instead, there is an attempt to capture patterns operating across a broader variety of contexts. Indeed, the patterns of the

“weird” group may turn out to be too atypical to form the optimal template for an optimal model of personality and character.

This view embraces complexity. One moves away from a small set of relatively simple standards (five universal factors, internal consistency, binary decisions as to scale validity). Both personality and personality measurement are viewed more complexly. What is important about the ways that people vary psychologically from one another cannot be reduced to a few endogenous phenomena that drive all else. And a good measure proves itself by meeting diverse standards, not only those based on classical test theory and factor analysis. Under this approach it is more challenging to be comprehensive. Scientists need to look beyond a few basic factors, and the assumption of universal cross-situational consistency must be set aside, allowing for context- and culture-specific phenomena.

In this alternative view, personality science becomes more of a science *of* comparisons *about* comparisons. This science compares models and measures of personality patterns, perceptions of which arise as we compare people to one another and to any of various standards. There is more “fluidity”: These standards are not necessarily fixed across time and situations; some of the standards are situation-specific. Some of the variation in standards might have to do with facts varying from place to place, but some may be due to evaluative standards varying.

The alternative view is not, however, a wildly relativistic approach in which all possible observations and views are accorded equal importance. Attention is directed to differentiating relatively universal from more culturally and contextually varying aspects of personality. More broadly, an implied overarching meme is the division of phenomena into relatively central and peripheral elements, with the latter not excluded from consideration. Examples of possible “central” phenomena are those established to be more rather than less universal, or the attributes

that multiple informants can agree on, broad and parsimonious personality factors, the components of psychological patterns that are cross-situationally consistent, and the components that can be linked to genetic influences or to biological correlates.

Thus the “patchwork” alternative view is not incoherent. We believe the evidence that has accumulated against the received view need not make the field more chaotic. The anomalies for the received view push us consistently in certain directions, those defined here by the alternative view.

This alternative view does not conform fully to any theoretical framework currently available. But some frameworks proposed in this volume intersect with it. Indeed, in our view, frameworks developed within industrial-organizational psychology are currently underappreciated within the science of personality proper; personality psychologists have much to learn from I/O psychologists. Let us review next some key intersections.

In Socio-Analytic Theory (Hogan & Blicke, this volume) there is a critique of overly “intrapsychic” approaches and an emphasis on reputation (as collective perception), both of which resonate with much of the alternative view (and with a “personality as social perception” view; Srivastava, 2010). In Socio-Analytic Theory there is also a clear implication that change in roles may be a causal factor in personality change (see also Nye & Roberts, this volume). Moreover, in Socio-Analytic Theory acceptance and status are taken to be prime motives underlying personality variation, and these are also key themes in the Social Self-Regulation and Dynamism factors that we have described as recurrent across cultural settings. Connelly’s (this volume) review of the observer (in our terminology, “informant”) perspective on personality, identifies strengths of informant data, and the limits of cross-situational consistency, in a manner consonant with the alternative view. The multilevel constructs identified by Narayan and

Ployhart (this volume) as emergent phenomena include organizational culture, which is defined in terms of “observable norms and values” (p. X) – which function as standards -- that are transmitted between members; we would reframe organizational culture as the context of a specific organization where distinct standards apply. In Trait Activation Theory (Tett et al., this volume), such standards are evident in “work demands” component, in which constraints and potential consequences are imposed. More broadly, we believe this theory (or a slightly more generalized adaptation of it) has much to offer as a model of how dispositions and situations relate and interact; the received view includes no such model. Finally, we note that the “watershed period” in research on personality and work in the late 1980s and early 1990s, as identified by Christiansen and Tett (this volume), was simultaneously the period in which what we term “the received view” became consolidated within personality science.

Conclusions

Marking the progress of personality as a science, there are a number of empirically based understandings that are now consensually held. These form the core of an emergent “paradigm of personality science.” Going beyond this core, we identify matters of controversy, on which we distinguish the received view and an alternative view.

In contemporary personality science, the received view has tended toward a fairly simple model emphasizing fixed and cross-situationally consistent traits identified in Western populations. The alternative view embraces complexity and attempts to reduce ethnocentric bias; it gives attention to peripheral (relativistic, context-specific) phenomena while differentiating them from central (universal, consistent, etc.) phenomena. The alternative view sees more fluidity in personality attributes, sees them emerging from social comparisons, and attends to the ways in which perception of attributes is influenced by standards that vary across situations and

cultural settings. Differing approaches to the creation and evaluation of measures follow in part from some of the same contrasts.

The alternative view delineated here does not correspond exactly to any theoretical framework previously proposed. We see it as emerging organically from anomalies that have arisen within the received view. It brings together various perspectives that seem to help make sense of these anomalies. We suggest the next paradigm shift in personality science will be in the direction of this alternative view, even if all components of this view detailed here are not ultimately retained. When such a shift occurs, it will have important implications for understanding and prediction in the field of “personality at work.” Indeed, as chapters in this volume make clear, research on work-relevant personality has vitally participated in and contributed to the current paradigm of personality science. It may likewise participate and contribute to future shifts in that paradigm.

References

- Allik, J., Massoudi, K., Realo, A., & Rossier, J. (2012). Personality and culture: Cross-cultural psychology at the next crossroads. *Swiss Journal of Psychology, 71*, 5-12.
- Ashton, M. & Lee, K. (2010) On the cross-language replicability of personality factors. *Journal of Research in Personality, 44*, 436-441.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*, 187-215.
- Block, J. (2010). The five-factor framing of personality and beyond: Some ruminations. *Psychological Inquiry, 21*, 2-25.
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development, stability, and change. *Annual Review of Psychology, 56*, 453-484.
- Clifton, A., Turkheimer, E., & Oltmanns, T. F. (2005). Self - and peer perspectives on pathological personality traits and interpersonal problems. *Psychological Assessment, 17*, 123-131.
- Cortina, J. M. (1993). What is coefficient Alpha: An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.
- Costa, P. T., & McCrae, R. R. (2006). Age changes in personality and their origins: Comment on Roberts, Walton, and Viechtbauer (2006). *Psychological Bulletin, 132*, 26-28.
- Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*, 322-331.

- Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology, 102*, 874-888.
- Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient Alpha and successor procedures. *Educational and Psychological Measurement, 64*, 391-418.
- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*, 1138-1151.
- Digman, J.M. (1996) The curious history of the Five Factor model. In Wiggins, J.S. (Ed.), *The five factor model of personality: Theoretical perspectives*. New York: Guilford.
- Digman, J. M. (1997). Higher order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246-1256.
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology, 6*, 178-190.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-73.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670.
- Funder, D. (2010). *The personality puzzle*. New York: Norton & Co.
- Goldberg, L. R. (1978). The differential attribution of trait-descriptive terms to oneself as compared to well-liked, neutral, and disliked others: A psychometric analysis. *Journal of Personality and Social Psychology, 36*, 1012-1028.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.) *Review of Personality and Social Psychology: Vol. 2* (pp. 141-165). Beverly Hills, CA: Sage.

- Gruza, R. A., & Goldberg, L. R. (2007). The comparative validity of 11 modern personality inventories: Predictions of behavioral acts, informant reports, and clinical indicators. *Journal of Personality Assessment, 89*, 167-187.
- Heine, S. J. (2012). *Cultural psychology, 2nd edition*. New York: Norton.
- Heine, S. J., Buchtel, E. E., & Norenzayan, A. (2008). Research report: What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychological Science, 19*, 4, 309-313.
- Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (January 01, 2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology, 82*, 6, 903-18.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83.
- Hogan, R. (1983). A socioanalytic theory of personality. *Nebraska Symposium on Motivation*.
- Hogan, R. & Hogan, J. (1997). *Hogan personality inventory manual*. Tulsa, OK: Hogan Assessment Systems.
- Hong, Y., Morris, M. W., Chiu, C., & Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist, 55*, 709-720.
- Jackson, D. N., Paunonen, S., Tremblay, P. F. (2000). *Six Factor Personality Questionnaire*. Port Huron, MI: Sigma Assessment Systems.
- Kanagawa, C., Cross, S. E., & Markus, H. R. (2001). "Who Am I?" The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin, 27*, 90-103.

- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Lee, K., & Ashton, M.C. (2004). The HEXACO Personality Inventory: A new measure of the major dimensions of personality. *Multivariate Behavioral Research*, *39*, 329-258.
- Leung, K. Bond, M. H. de Carrasquel, S. R. Munoz, C. Hernandez, M. Murakami, F. (2002). Social axioms: the search for universal dimensions of general beliefs about how the world functions. *Journal of Cross-Cultural Psychology*, *33*, 286-302.
- Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment*, *14*, 299-316.
- Lodi-Smith, J., & Roberts, B. W. (2012). Concurrent and prospective relationships between social engagement and personality traits in older adulthood. *Psychology and Aging*.
- Loevinger, J. (1954). The attenuation paradox in test theory. *Psychological Bulletin*, *51*, 493-504.
- McCrae, R. R. (1990). Traits and trait names: How well is Openness represented in natural languages? *European Journal of Personality*, *4*, 119-129.
- McCrae, R. R. (1993). Moderated analyses of longitudinal personality stability. *Journal of Personality and Social Psychology*, *65*, 577-585.
- McCrae, R. R. (2002). NEO-PI-R data from 36 cultures: Further intercultural comparisons. In R. R. McCrae & J. Allik (Eds.), *The five-factor model of personality across cultures*. New York: Kluwer Academic / Plenum.
- McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed.) (pp. 139-153). New York: Guilford.

- McCrae, R. R. & Terracciano, A. (2006). National character and personality. *Current Directions in Psychological Science*, 15, 156-161.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mischel, W., & Peake, P. K. (1982). Beyond déjà vu in the search for cross-situational consistency. *Psychological Review*, 89, 730-755.
- Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, 61, 622-631.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660-679.
- Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, 57, 401-21.
- Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the Big Five? Plenty! *Journal of Personality*, 68, 821-835.
- Piedmont, R. L., Bain, E., McCrae, R. R., & Costa, P. T. (2002). The applicability of the Five-Factor model in a sub-Saharan culture: The NEO-PI-R in Shona. In R. R. McCrae & J. Allik (Eds.). *The five-factor model of personality across cultures* (pp. 155-173). New York: Kluwer Academic/Plenum.
- Poortinga, Y. H., van de Vijver, F. J. R., & Van Hemert, D. A. (2002). Cross-cultural equivalence of the Big Five: A tentative interpretation of the evidence. In R. R. McCrae & J. Allik (Eds.). *The five-factor model of personality across cultures* (pp. 281-302). New York: Kluwer Academic/Plenum.

- Rolland, J.-P. (2002). The cross-cultural generalizability of the Five-Factor model of personality. In R. R. McCrae & J. Allik (Eds.). *The five-factor model of personality across cultures* (pp. 7-28). New York: Kluwer Academic / Plenum.
- Saucier, G. (2009). Recurrent personality dimensions in inclusive lexical studies: Indications for a Big Six structure. *Journal of Personality, 77*, 1577-1614.
- Saucier, G. (2012). *Isms dimensions: Toward a more comprehensive and integrative model of belief-system components*. Manuscript submitted for publication.
- Saucier, G., & Goldberg, L. R. (1998). What is beyond the Big Five? *Journal of Personality, 66*, 495-524.
- Saucier, G. & Simonds, J. (2006). The structure of personality and temperament. In Mroczek, D. K. & Little, T. D. (Eds.), *Handbook of personality development* (pp. 109-128). Mahwah, NJ: Erlbaum.
- Saucier, G., Thalmayer, A.G., & Bel-Bahar, T. (2012). *Human attribute concepts: Relative ubiquity across twelve mutually isolated languages*. Manuscript submitted for publication.
- Saucier, G., Thalmayer, A.G., Payne, D. L., Carlson, R., Sanogo, L., Ole-Kotikash, L., Church, A.T., Katigbak, M. S., Szarota, P., Szirmák, Z., & Zhou, X. (2012). *A basic bivariate structure of personality attributes evident across nine languages*. Manuscript submitted for publication.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*, 350-353.
- Srivastava, S. (2010). The five-factor model describes the structure of social perceptions. *Psychological Inquiry, 21*, 69-75.

- Suh, E. M. (2002). Culture, identity consistency, and subjective well-being. *Journal of Personality and Social Psychology*, 83, 1378-91.
- Thalmayer, A. G. Saucier, G. & Eigenhuis, A. (2011). The comparative validity of brief- to medium-length big five and big six personality inventories. *Psychological Assessment*, 23, 995-1009.
- Thorndike, R. L. (1967). The analysis and selection of test items. In D. N. Jackson & S. Messick (Eds.), *Problems in human assessment* (pp. 201-216). New York: McGraw-Hill.
- Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, 9, 160-164.
- Uziel, L. (2010). Rethinking social desirability scales: From impression management to interpersonally oriented self-control. *Perspectives on Psychological Science*, 5, 243-262.
- Wallace, A. F. C. (1961). *Culture and personality*. New York: Random House.
- Weisberg, Y., De Young, C., & Simpson, J. A. (2012, July). *The influence of ideal personality and significant others on personality change*. Paper, 16th European Conference on Personality, Trieste, Italy.
- Westen, D. (1985). *Self and society: Narcissism, collectivism, and the development of morals*. New York: Cambridge University Press.
- Wood, D., Nye, C. D., & Saucier, G. (2010). Identification and measurement of a more comprehensive set of person-descriptive trait markers from the English lexicon. *Journal of Research in Personality*, 44, 258-272
- Xie, J. L., Roy, J.-P., & Chen, Z. (2006). Cultural and individual differences in self-rating behavior: An extension and refinement of the cultural relativity hypothesis. *Journal of Organizational Behavior*, 27, 341-364.

Table 1

	The received view in personality science	An alternative view
1	Questionnaire responses reflect disinterested, factual descriptions of persons' attributes	Personality descriptions are a compound of fact and value (but this tends not to compromise predictive validity of scores)
2	Self-report is the primary data source (informant data mainly demonstrates the validity of self-report data)	The prime source of data for human behavioral patterns is knowledgeable informants
3	Across contexts and cultures, we compare targets to the same objective standard	Standards of comparison vary by culture and context
4	People behave consistently across situations; instructions regarding context are not needed	Personality tendencies are in part situation-specific
5	Personality traits are hard-wired and essentially unchangeable "essences"	Personality tendencies are based partially in genetics, but also "software" influenced by cultural and contextual factors
6	The most important personality traits and their structure/organization, can be identified in European-origin populations, and generalized to the rest of the world	Models of personality dispositions bear some imprint of the cultural setting in which they were developed
7	Studies of the natural language of personality in diverse cultures support a structural model with five factors	Studies of the lexicons of northern European languages support the five-factor model, but studies in other languages generally do not
8	The five factors provide a sufficient representation of personality	Five factors capture much of the personality domain, but are not comprehensive
9	Coefficient Alpha (internal consistency) is on par with validity as important to evaluations of measurement quality	Unidimensionality and retest stability are as important as internal consistency, and validity is more important
10	In evaluating a personality scale or inventory, it is adequate to test against chance to determine if it has validity	The value of a measurement instrument is best assessed in competition against partly similar instruments

