An Organizational Framework for the Psychological Individual Differences: Integrating the Affective, Cognitive, and Conative Domains

Thesis · December 2014

DOI: 10.13140/2.1.4964.1283

1 author:

David Condon
University of Oregon
56 PUBLICATIONS **1,193** CITATIONS

SEE PROFILE

NORTHWESTERN UNIVERSITY

An Organizational Framework for the Psychological Individual Differences: Integrating the Affective, Cognitive, and Conative Domains

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

David M. Condon

EVANSTON, ILLINOIS

December 2014

ABSTRACT

An Organizational Framework for the Psychological Individual Differences:
Integrating the Affective, Cognitive, and Conative Domains

David M. Condon

Recognition of the importance of individual differences dates back to humanity's oldest surviving texts yet the scientific study of individual differences has been surprisingly limited. This paradox is presumed to result from the fact that differential psychology has struggled to graduate beyond pre-paradigmatic status as a science. In part, this has stemmed from the tendency to align idiographic approaches with the largely nomothetic methods of differential psychology under the broad label of "personality" research. The struggle has shifted – and, to some extent, abated – following acceptance of the Big Five taxonomy of personality and the more pressing concern has recently been the need to incorporate findings from additional disciplines of differential psychology.

The purpose of this research was to propose an integrated assessment model – a preliminary paradigm which can be tested against extant and future models of individual differences in terms of predictive utility for a wide range of behaviors. The procedures used to develop this model are described separately by discipline (temperament, cognitive ability and vocational interests) and are supplemented by a methodological study regarding item clusters and complexity. All analyses were based on Synthetic Aperture Personality Assessment sampling procedures and large international samples (*N*s ranged from 24,000 to 97,000 participants representing 170 to 199 countries).

The proposed temperament scales were iteratively derived from factor analyses of the items in 8 widely-used public-domain measures and can be scored at three hierarchical levels (with 3, 5 and 15 factors). The case is made that these scales are well-suited for heterarchical assessment and that the heterarchical organization of personality constructs often reflects the manner in

which personality models are used in everyday settings. The cognitive ability scales represent a validated public-domain pool of items designed to assess several types of ability in unproctored online settings. The vocational interest scales are derived from two public-domain measures and reflect the traditional six-factor interests framework. Collectively, these scales form an efficient multi-dimensional, multi-disciplinary assessment model (the “SAPA Personality Inventory”) which aims to serve as a preliminary testable paradigm for differential psychology research.

ACKNOWLEDGEMENTS

It has often been remarked that an increasing proportion of the scientific endeavors worth pursuing in the 21st Century will require the collaboration of many highly-trained individuals with specialized skills. Psychology is not likely to be an exception to this trend. In fact, the work described here – despite serving officially as my dissertation – simply would not have been completed without invaluable input from many of my colleagues. In some cases, it does not even seem appropriate to describe their involvement as a "contribution" because it was so integral.

Above all others, I am tremendously grateful to William Revelle for his fundamental enthusiasm, his inspiration (just a few small suggestions several years ago laid the groundwork for this entire project), his consistent efforts to keep me motivated (with both the carrot and the occasional stick), and his intellect (it seemed he always knew how to proceed at every turn, often before I even saw the turn ahead).

I would also be remiss if I didn't single out a few of the many others who have passed through our lab over the last several years: Joshua Wilt, for never tiring of my ignorance in the early days of grad school and, more recently, for his encouraging suggestions and insightful observations about prior research overlooked; Jason French, whose theoretical contributions and camaraderie were only surpassed in terms of helpfulness by the fact that he was never too busy to help me figure out "why the damn code won't work!"; Sara Weston, whose creativity and passion for science (in general) and personality (in particular) helped me remember why I was doing this in the first place; and Lorien Elleman, whose presence in the lab during my last year helped me to recognize both how much I had and hadn't yet learned. I am particularly grateful to Lorien for his contribution to the chapter on the conative domain – if I'd had more time to invest on any section, it would've been that chapter in order to do his contribution justice.

Thank you also to both Dan Mroczek and David Uttal for their support, insight, and patience

throughout the dissertation process and my time in graduate school. I have learned a great deal from you both already and look forward to further growth from our collaborations in the future. Finally, there are a number of individuals who have either contributed to this project directly or to my development more generally. These include (but are not limited to) Rick Zinbarg, Katherine Oehlberg, Tim Sweeny, KatieAnn Skogsberg, Marcia Grabowecky, Satoru Suzuki, Jillian Cavanaugh, Ashley Brown, Philipp Doebler, Dan McAdams, Lew Goldberg, Colin DeYoung, and Robin Nusslock – a warm thanks to all of you.

DMC (December 14, 2014)

DEDICATION

For my family, especially my parents.
You may not have fully understood this digression,
but you supported me anyway.

For Sara,
because you woke me from a deep sleep to say
that the future is long and unpredictable
and, quite likely, glorious.

And for Kelsey,
because “the cure for boredom is curiosity and
there is no cure for curiosity.” *

* Thank you, Dorothy, for priorities.

Contents

List of Figures

List of Tables

Chapter 1

The Paradigmatic Science of Individual Differences

> “You should not establish a home with an arrogant man. ...
> The eyes of the slanderer always move around as shiftily as a spindle. You should never remain in his presence. ...
> You should not boast in beer halls like a deceitful man: then your words will be trusted. ...
> The artistic mouth recites words; the harsh mouth brings litigation documents; the sweet mouth gathers sweet herbs. ...
> The imprudent decrees fates; the shameless one piles up things in another’s lap: ‘I am such that I deserve admiration.’ ...
> The negligent one ruins his family. ...
> A loving heart maintains a family; a hateful heart destroys a family.”
>
> *The Instructions of Šuruppag (c. 2600 BCE)*

More than a few lines from the oldest surviving text describe the effects of individual differences in human behavior, and the insights of Sumerian King Šuruppag are, by no means, an isolated example. Individual differences are also addressed in influential works by several ancient Chinese authors, including the *Book of Documents* (Legge, 1879, a.k.a. the *Classic of History* or *Shujing*, c. 5th to 11th centuries BCE), the *Analects* of Confucius (Confucius, 1994, c. 435 BCE), and Liu Shao’s *Classified Characters and Political Abilities* (Shao, 2007, a.k.a. *Ren Wu*, c. 200 CE).

This last example is a twelve chapter volume which explicitly posits a detailed theory of individual differences in temperament and cognitive ability as well as descriptions of suitable methods for observational data collection and application of the theory for political and social benefit (Shao, 2007).

More familiar to Western scholars are examples contributed by several ancient Greeks. These include Galen's theory of temperament (based upon the Four Humors described by Hippocrates (Galen, 1916; Hippocrates and Galen, 1846), c. 460 BCE), Aristotle's *Nicomachean Ethics* (Aristotle, 2000, c. 350 BCE), Cicero's *Tusculan Disputations* (Eysenck, 1983a, c. 45 BCE), and the *Characters* of Theophrastus (Boegehold, 1959; De Raad and Ceulemans, 2001; Theophrastus, 1927, c. 319 BCE). Retrospectives on individual differences often begin with Theophrastus' *Characters* – probably because it provides an ancient example of the typological approach to describing trait constellations – though these conveniently overlook the considerable evidence that Theophrastus' claims regarding the generalizability of his characters were made as a thinly-veiled attempt to shield him from the consequences of overtly satirical descriptions of his political contemporaries (Boegehold, 1959; Theophrastus, 1927).

In any case, it seems reasonable to infer on the basis of these ancient references that individual differences in behavior have been a topic of cross-cultural relevance since the beginning of recorded history. Much more recently, evolutionary theory has suggested that the role for intraspecific individual differences may be even more primal. Some of Darwin's own observations in the first edition of *The Origin of Species* (1859) are particularly relevant (see the subsection titled "Individual Differences" in Chapter II – Variation under nature, and all of Chapter IV – Natural Selection; or the Survival of the Fittest). Darwin initially avoided the suggestion that his theory might also be extended to the human species though the study of individual differences among humans was promptly pursued by others (Fechner, 1860; Galton,

1869) and later by Darwin himself (Darwin, 1871, 1886). "Variation is, after all, the grist for the mill of evolution" (Nettle, 2006) in that random variability is the means by which natural and sexual selection mechanistically proceed, perhaps even among humans in the modern era (Courtiol et al., 2012, 2013).

1.1 The absence of a paradigm

> "What human personality is, everybody knows; but nobody can tell."
>
> *William H. Burnham (1929) in Allport and Vernon (1930)*

Two observations are worth emphasizing about the relationship between evolutionary theory and the ancient written references to individual differences. The first of these is recognition of the possibility that, if evolutionary theory is valid across species, references to the importance of individual differences in human texts may be – both literally and figuratively – an artifact of the capacity for verbal and written communication. In other words, the importance of individual differences is not necessarily (and likely, *is not*) dependent on the presence of language structures. This also implies that the degree of introspective recognition (conscious or unconscious) and/or mutual recognition of individual differences likely varies across and possibly within species.

More explicitly, discussion about the importance of individual differences in several of the earliest texts implies that they were important before writing skills were well-developed (the alternative – that the development of writing occurred simultaneous with recognition of the importance of individual differences – seems implausible). This implication has philosophical consequences for the so-called "Lexical Hypothesis," an idea that was first proposed by Sir Francis Galton (Galton, 1884) and has since served as a foundational assumption in individual differences research. The Lexical Hypothesis essentially states that "those individual differences

that are of the most significance in the daily transactions of persons with each other will eventually become encoded into their language" (Goldberg, 1981). Further elaboration of this topic is given in Chapter 2, but it is worth noting here that at least some important underlying differences pre-dated the existence of words to describe them (and the existence of words themselves), and that this is consistent with the Lexical Hypothesis.

The second observation relates to the large chronological discrepancy between the earliest written references to individual differences and the first incidence of their mention in a *scientific* context. Despite being encoded in written language for millennia, individual differences were rarely the focus of systematic study and classification. This circumstance is markedly different from many other aspects of human experience, most notably those which are now characterized as part of the natural sciences. This combination – the widely acknowledged importance of individual differences among humans and the absence of systematic study of their structure or even definition – has led to a diverse array of partially overlapping, anecdotally-derived lay "theories" (Shamdasani, 2003). While the generations after Darwin and his contemporaries (most prominently, Galton) embraced the suggestion that individual differences and their various subsets should be the target(s) of scientific study, the lack of consensus about the best means of proceeding was impressive and persistent.

Throughout the 20th century, scholars of individual differences from varying theoretical orientations have concurred about this dilemma, describing the situation as: "a chaos [that] does not give unity or definiteness of direction to our study" (Allport, 1921); "little other than a chaos of arbitrary dogmas... with complete lack of agreement" (Jung, 1925 from Shamdasani, 2003); "a deadlock: we cannot advance to agreed conclusions for lack of common terminology; and we cannot achieve such a terminology because of the extreme diversity of views among authorities" (McDougall, 1932); "a chaotic center in personality research" (Cattell, 1940); "no progress seems

to have been made" (Roback, 1952); "most so-called 'theories'... are scientifically unimpressive and technologically worthless" (Meehl, 1978); "one element of the model [of a mature science] has no counterpart in the field of personality: the unanimity of qualified persons in agreeing on a paradigm" (Loevinger, 1987); "there is no agreement on definitions, models, methods, results or indeed anything whatever; all is confusion..." (Eysenck, 1991); and "personality psychology has yet to articulate clearly a comprehensive framework for understanding the whole person" (McAdams and Pals, 2006).

These claims fit quite well with the definition of "pre-paradigmatic science" provided by Thomas Kuhn (Kuhn, 1962). While the terminology introduced in Kuhn's seminal work, *The Structure of Scientific Revolutions* (Kuhn, 1962, 1970), is now commonplace, it has also been subject to considerable reconstruction, re-interpretation and even misinterpretation in the vast secondary literature spawned by his original text (Hoyningen-Huene, 1993). Given this and the relevance of his philosophy of science to the issues at hand, a lengthy quotation from the original text is justified:

> "In the absence of a paradigm or some candidate for paradigm, all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant. As a result, early fact-gathering is a far more nearly random activity than the one that subsequent scientific development makes familiar. Furthermore, in the absence of a reason for seeking some particular form of more recondite information, early fact-gathering is usually restricted to the wealth of data that lie ready to hand. The resulting pool of facts contains those accessible to casual observation and experiment together with some of the more esoteric data retrievable from established crafts like medicine, calendar making, and metallurgy. Because the crafts are one readily accessible source of facts that could not have been casually discovered, technology has often played a vital role in the emergence of new sciences. But though this sort of fact-collecting has been essential to the origin of many significant sciences, [several examples demonstrate that] it produces a morass. One somehow hesitates to call the literature that results scientific. ... [T]he typical natural history often omits from it's immensely circumstantial accounts just those details that later scientists will find sources of important illumination. ... This is the situation that creates the schools characteristic of the early stages of a

> science's development. No natural history can be interpreted in the absence of at least some implicit body of intertwined theoretical and methodological belief that permits selection, evaluation, and criticism. If that body of belief is not already implicit in the collection of facts – in which case more than 'mere fact' are at hand – it must be externally supplied, perhaps by a current metaphysic, by another science, or by personal and historical accident. No wonder, then, that in the early stages of the development of any science different men confronting the same range of phenomena, but not usually all the same particular phenomena, describe and interpret them in different ways. What is surprising, and perhaps also unique in its degree to the fields we call science, is that such initial divergences ever largely disappear. For they do disappear to a very considerable extent and then apparently once and for all. Furthermore, their disappearance is usually caused by the triumph of one of the pre-paradigm schools, which, because of its own characteristic beliefs and preconceptions, emphasized only some special part of the too sizable and inchoate pool of information. ... To be accepted as a paradigm, a theory must seem better than its competitors, but it need not, and in fact never does, explain all the facts with which it can be confronted." (Kuhn, 1962, Chapter 2)

This suggests that the study of individual differences has been in the pre-paradigmatic stage for most of recorded history (though, in fairness, this is offset by the fact that individual differences were rarely considered in a scientific context prior to the mid-19th century). Nevertheless, the very prolonged period of "early fact-gathering" is consequential. Thousands of years of casual observation can produce a morass of unusual depths, one which might well require several generations to resolve. This morass may be exacerbated by the fact that those who come to study individual differences typically do so after decades of personal, informal fact-gathering which is unsupported by paradigmatic scaffolding and infused by exposure to pseudo-scientific lay theories (for discussion of examples, see Cattell et al. (1964); Dahlstrom et al. (1996); Mehl et al. (2006); Thagard (1978)). Resolution is possible however. Kuhn suggests that the pre-paradigmatic era fades with the spreading recognition that one school of thought is theoretically superior to the rest. The question for scholars of individual differences is whether such a resolution will ever come to pass, if it has not already occurred.

The prospect of a paradigm in individual differences research has been directly addressed several times over the last 30 years (Eysenck, 1983b; Loevinger, 1987; Wiggins, 2003), with two dissimilar conclusions. Loevinger (Loevinger, 1987) and Wiggins (Wiggins, 2003) concurred that several paradigms exist, with Loevinger going so far as to claim that "there will always be a multiplicity of paradigms" (Loevinger, 1987, p. 6). Both authors coincidentally identify the same number of active paradigms (five), though only two of these are clearly overlapping – the psychodynamic/psychoanalytic paradigm and the multivariate/psychometric paradigm.

For Eysenck, the vital need for a singular paradigm was a frequent refrain (Eysenck, 1983b, 1985, 1991, 1994, 1997; Eysenck and Eysenck, 1985), though the objectivity of this claim was somewhat discredited by the suggestion that his own structural theory (the P-E-N model, discussed in Chapter 3) was the most obvious choice. Despite the partiality of his conclusions, Eysenck's reflections on the issue of paradigm development included several arguments which remain relevant today. Most notable is his suggestion that it is first necessary to evaluate the degree to which the study of individual differences constitutes a *scientific* endeavor before one can consider paradigm development (Eysenck and Eysenck, 1985; Eysenck, 1985).

1.2 On the scientific quality of individual differences research

> "[O]ne source of a malign compass deviation in the early days has been the very eagerness to assume a true scientific status. For this led to premature regimentation, and indeed slavishness, in following the rules of the older, established sciences when the need was really for invention of methods and trial-and-error exploration of the scientific quality of a new area."
>
> *Raymond B. Cattell (1966)*

The "scientific issue" has been a perennial debate (Jastrow, 1901; Boring, 1923; Anastasi, 1948; Hornstein, 1988, 1992) among those who study individual differences, and it has occasionally

been quite heated. A representative example comes from the 1923 meeting of the American Psychological Association when James McKeen Cattell interrupted the meeting to castigate a fellow member for mentioning Freud's name in the context of scientific discourse (Dallenbach, 1955) (this was far more controversial than it might seem today as psychoanalytic theories of personality organization were increasingly popular among APA members at that time). The essence of the controversy is that some approaches to psychological research claim to be more representative of science than others (by virtue of quantification and generalizability) in contrast to the Kantian view that the quantification of mental events is philosophically impossible (Kant, 1979; Loevinger, 1987). The suggestion that quantification and generalizability are key components of the scientific method (Popper, 1959) causes them to be viewed, by turns, as either a necessity in the study of individual differences (Cattell, 1940) or a mindless and unwarranted conformity (Giorgi, 1975) to the standards of the so-called natural sciences.

The defense against quantification and the search for generalizability rests largely on the belief that hermeneutic and existentialist concerns are fundamental to psychologically-oriented individual differences. The extreme view is that the precisely unique qualities of the individual are key determinants of behavior and that the identification of communality across individuals requires an invalidating lessening of precision. As such, idiographic approaches are not only appropriate but mandated in the study of topics such as identity or unconscious features of the psyche. Research on such topics often enjoys wide appeal by virtue of an emphasis on individuality. This emphasis is not inherently problematic – in fact, examination of individual experiences can offer great utility for both the individuals under examination as well as those who endeavor to understand development. But, strict idiographic study is scientifically problematic because it subordinates the search for a generalizable structure of differences across individuals. When used in isolation, idiographic approaches seldom offer opportunities for theory testing in the Popperian sense (Popper, 1959), mainly because it is not currently possible

to identify, measure, and control every one of the environmental and biological variables underlying individual outcomes.

The esoteric nature of this debate is undeniable; it hinges upon the degree of commonality in differences. Yet Eysenck (1985) implies that inaction (or perhaps even boredom) caused by this esoteric bind is actually the primary obstacle to graduation from the pre-paradigmatic state. The study of individual differences, according to his logic, is dissimilar from paradigmatic sciences in that there there is a surplus of "theories" which are either (a) unlikely to ever enter into the realm of science on the grounds that they are incapable of being used to make testable predictions; or (b) so narrow and methodologically restrictive that they make verifiable predictions of little-to-no relevance. The latter approach sacrifices utility for the sake of rigorous scientific methods (though it should be noted that Eysenck's opposition to strict empiricism was somewhat inconsistent over the last 10 years of his career). If a paradigm in individual differences research were to gain acceptance according to Kuhn's suggestion – by merely outdoing its competitors – it would have to fulfill the basic qualifications of science while remaining broad enough to address issues of demonstrable relevance.

Eysenck is not unique in calling for a middle path. To the contrary, it seems that individual difference scholars ironically relish the typological as an explicative tool. Allport pitted the Realist against the Nominalist (Allport and Odbert, 1936) and Actuarial approaches (Allport, 1940); Meehl (1954) the Statistical versus the Clinical. Raymond Cattell split the field into three camps: those who embraced the multivariate approach, the overly-rigorous "bivariate brass instrument" methodologists, and a loose collection of "numerous quasi-scientific schools which led to that scholastic Tower of Babel" (Cattell, 1966, p. 8). (All of these scholars were admittedly biased towards the differential approach advocated herein.) Cronbach (1957; 1975) was perhaps more objective in his description of the "Tight Little Island" of experimentalists and the united

principalities of the correlationalists' "Holy Roman Empire."

The generalized form of these observations is that individual differences research can be organized along a spectrum according to its "scientific-ness." This spectrum is mainly methodological though does also reflect underlying theory in that the most scientific methods tend to address mechanisms of behavior that are common to the human species while the least scientific approaches tend to deal with idiographic aspects of individual experience. As Kluckhohn and Murray (1948) observed, "every man is, in certain respects, like all other men, like some other men, and like no other man" and these degrees of similarity are reflected in the varied types of research on individual differences in behavior.

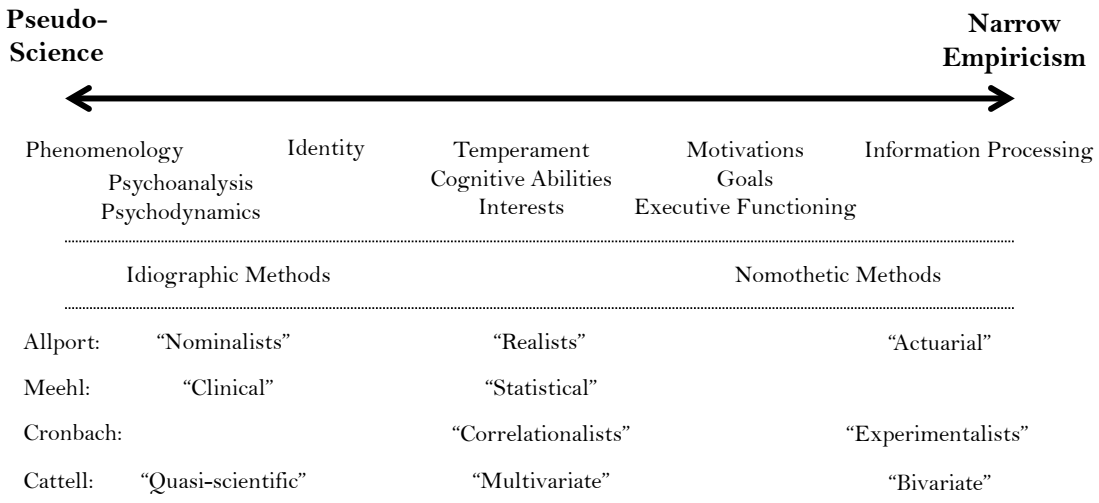


Figure 1.1: The scientific dimension in research on psychological individual differences
Note: This is not intended to exhaustively depict all the major domains of individual differences research nor does it attempt to account for fields of psychological research that seek to describe interpersonal interaction.

Figure 1.1 attempts to capture the essence of this dimension. The ends of this spectrum represent the problematic types described by Eysenck (those which make untestable predictions on the left and those which fail to make predictions of relevance on the right). The labels prescribed by others (Allport, Meehl, Cronbach, and Cattell) are included on the spectrum as

well, though it should be acknowledged that their exact placement would likely lead to some debate among the scholars working in these areas.

The rationale for explicit description of this dimension, despite its imprecision, is to demonstrate the range of "scientific-ness" for research on psychological individual differences. Today, nearly all of the researchers working at various locations on this dimension would describe themselves as "personality psychologists," except perhaps for those "cognitive psychologists" exploring the more generalizable mechanisms of information processing and executive functioning. Those working on the left end of the spectrum might also be referred to as "personologists," though this is uncommon. Research in the middle of this spectrum has traditionally been known as "differential psychology." The placement of these labels on the spectrum is demonstrated in Figure 1.2.

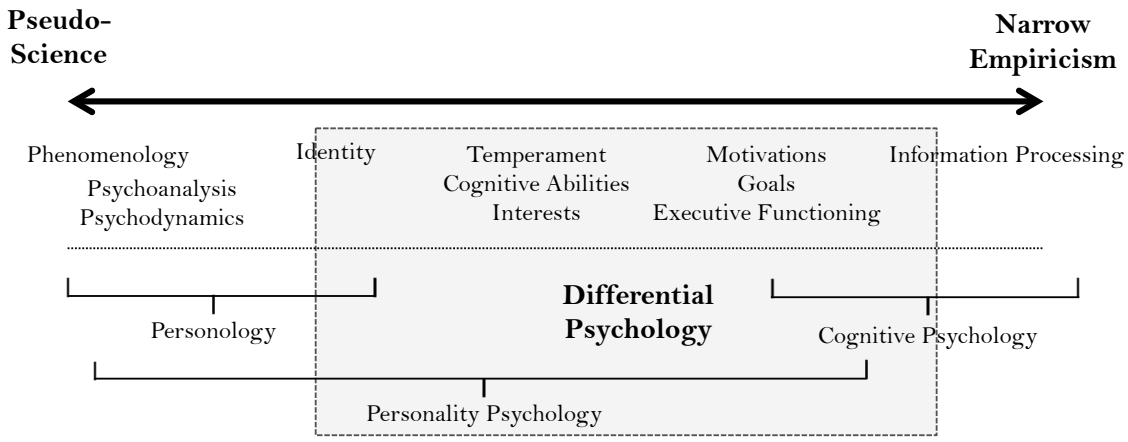


Figure 1.2: Situating Differential Psychology on the scientific dimension

1.3 The need to distinguish personality from differential psychology

> “Although it is true that it is the goal of science to discover rules which permit the association and foretelling of facts, this is not its only aim. It also seeks to reduce the connections discovered to the smallest possible number of mutually independent conceptual elements.”
>
> *Albert Einstein (1941)*

In practice, the “differential psychology” label is used only rarely and even then it is viewed as essentially synonymous with “personality.” The tendency for these terms to be used interchangeably is problematic for two reasons. The first of these is apparent in Figure 1.2; both personality and differential psychology include domains of research which are exclusive of one another. In other words, there are aspects of differential psychology which are not traditionally viewed as part of personality psychology (e.g., neuroanatomical differences) and vice-versa (e.g., case studies of phenomenological experience).

A more nuanced, and perhaps impactful, issue with imbrication of these terms relates to imprecision of the term “personality” more generally. After decades of disagreement about the subtle differences between terms such as character, temperament, and personality (Allport, 1921; Fernald, 1920; Gilliland, 1928; Jastrow, 1915; Klages, 1929; May and Hartshorne, 1927; McDougall, 1929, 1932; Tolman, 1932; Roback, 1927), the burgeoning field united behind Allport’s (1930) view that this area of research should be defined by broad use of the term “personality” to overlay all possible integrative and omnibus interpretations.

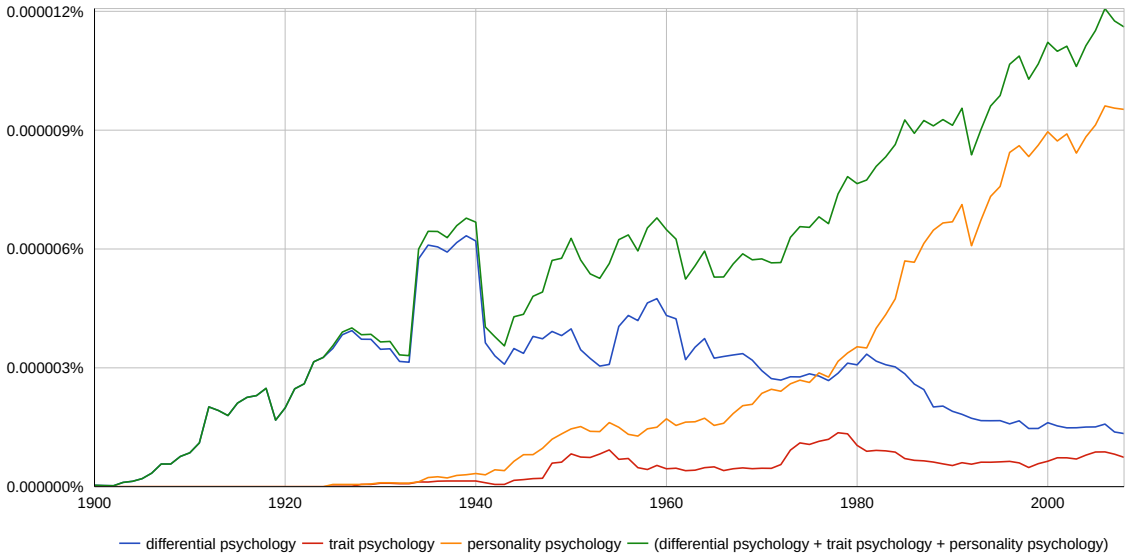
This practical approach is far preferable to the previous ambiguity. However, there is also utility in the application of more specific terminology. “Temperament” for example is generally regarded as the affective component of personality (Allport and Vernon, 1930; Hofstee, 1991; Shiner and DeYoung, 2013) in humans (and other species (Gosling and John, 1999; Gosling, 2001;

Weinstein et al., 2008), though this necessitates further expansion of *person*ality to include non-person animals). Similarly, "traits" are subsumed under the personality label, though they are more specific than temperament. That is, some personality traits might be described as temperamental traits while others would not.

How should "differential psychology" be distinguished from "personality psychology"? Differential psychology seeks to describe and understand individual differences in order to make predictions about behavior. Personality psychologists who disavow the differential psychology approach seek to merely describe patterns of individuality (Lamiell, 1981, 2003), without specification of the extent to which various features of a given signature may be idiosyncratic. As mentioned earlier, these are overlapping but they are *not the same*. The inherently scientific pursuit of (generalizable) predictiveness implies an intention to identify and organize individual differences in terms of their relationships with various outcomes.

To be clear, use of the label "personality psychology" is not inherently flawed, but rather overly vague. This may be the source of its appeal. Figure 1.3 shows the frequency of usage for "personality psychology" and various other two-word phrases (bigrams) in a sample of books written in English with publication dates between 1900 and 2008 and subsequently digitized by Google (approximately 30 million volumes). While the occurrence of these phrases is an admittedly weak indicator of the nature of the research being conducted, the relative frequency of their usage over time is suggestive of long-term trends.

Perhaps the most important trend is the steady increase in usage of the phrase "personality psychology" between 1940 and 1980. Over the same period, usage of "differential psychology" and "trait psychology" was essentially unchanged. After 1980, "personality psychology" continues to increase in frequency while "differential psychology" and "trait psychology" usage decreases. One interpretation of these trends would be that some of the differential psychology



Click on the image to interact with the figure

Figure 1.3: Frequency of usage for labels describing individual differences research
Note: Frequencies based on usage in the corpus of books digitized by Google and written in English (American and British) with publication dates between 1900 and 2008

research conducted since 1980 has been more generically labeled as personality research, though it's also possible that the volume and/or significance of differential psychology research has decreased.

The possibility that a substantial portion of personality research might be more narrowly labeled as differential psychology is supported by two related phenomenon over the last few decades. Hofstee (2007) distinguishes these as the "epistemic" and "ethical" components of the argument against the study of individual differences. The ethical argument often invokes Galton's well-known support of eugenics and makes vague implications that those working in other areas which were influenced by Galton probably maintain the same perspective as British aristocrats at the end of the 19th century. This is no more true for differential psychology than it is for genetics. In the event that the disastrously destructive events of the first half of the 20th century did not provide sufficient rationale for disagreeing with Galton's views about anthropological typologies, it is widely recognized, even among the general public (Condit, 1999; Condit et al., 2001), that strict genetic determinism lacks scientific support. Consider, for example, the differences in height among genetically homogenous populations in North and South Korea (Schwekendiek, 2009; Johnson, 2010b); even for the small number of individual differences where variability is highly influenced by genetics, environmental influences can still contribute substantially. Individual differences in behavior are no exception: they are non-deterministic, multi-factorial by-products of a tremendous number of environmental and genetic variables (Weiss and Lambert, 2011).

The epistemic argument calls into question the value of making predictions (as previously addressed) in a world of limitless individuality and situational specificity (Lamiell, 1981). With regards to nomenclature, it may be that "personality psychology" is preferred over "differential psychology" if the former suggests a greater allowance for contextualism by virtue of simply

being more vague. In other words, it may be the case that *descriptive* personality psychology research is held to looser standards than *predictive* differential psychology research when it comes to accounting for situational factors.

As trivial as these arguments may seem to many scholars, they continue to perpetuate the very constraints lamented by Cattell (1966), Cronbach (1957), and Eysenck (1985). These arguments may serve as legitimate rationale for use of the more generic "personality" label when naming program areas within academic psychology departments or when describing the broad aims and objectives of a research journal, but they should be eschewed by those conducting quantifiable and generalizable research when more specific terminology is an option[1]. While there exist at least two academic societies (ISSID, 2014; SMEP, 2014) and a sizable contingency of scholars, especially outside the United States, who routinely frame their work as differential psychology research when appropriate (Bouchard et al., 1990; Buss, 1991; Chamorro-Premuzic et al., 2011; DeYoung, 2010a; Johnson, 2007; Kanai and Rees, 2011; Krueger and Johnson, 2008; Nettle, 2006; Plomin and Rende, 1991), the number of researchers who identify as differential psychologists could be much larger still if priority were given to the use of more specific nomenclature.

Differential psychology is the domain of research which occupies the middle road between the strictly unique and universal qualities of a species. By virtue of its focus on the psychological differences, the majority of research in this field focuses on behavioral variance in the human

[1]Further digression might be made to consider the ways in which the terms "trait psychology" and "evolutionary psychology" relate to personality and differential psychology. In brief, "trait psychology" or "trait theory" is very similar to differential psychology in that traits are typically conceptualized as synonymous with individual differences. Use of the term has been confused somewhat by recent study of universal human traits (Buss, 1984; Kappeler et al., 2010; McCrae and Costa, 1997; Pinker, 2002), which are those common among humans but unique to varying degrees relative to other species. Differential psychology is slightly preferred because it emphasizes the key feature of *difference* (within species). Evolutionary psychology is a logical subset of differential psychology on the grounds that variation is a prerequisite for evolution but not all differences (including, possibly, some which are predictive of behavior) can be demonstrated to have an evolutionary effect. There is also some controversy about the relevance of evolutionary pressures for modern humans given the so-called adaptive-lag hypothesis (Courtiol et al., 2012; Laland and Brown, 2006; Smith et al., 2001). In any case, both of these terms overlap considerably with differential psychology; it seems that differential psychology is more specific than trait psychology and more broad than evolutionary psychology.

species. The primary challenges to the development and testing of scientific models in differential psychology relate to the identification and quantitative measurement of variables that represent a balance between generalizability and variance across populations. To borrow the language of Einstein (1941), the aim of differential psychology is the "discovery" of those differences "which permit the association and foretelling of facts" and "reduce the connections discovered to the smallest possible number of mutually independent conceptual elements." This aim will be achieved through the comparative testing of models.

After lengthy digression, it is now proposed that individual differences research *is* a scientific endeavor in Eysenckian terms to the extent that the concepts under study are quantifiable and that explanatory models of these concepts are generalizable and testable. A great deal of research which meets these conditions has been (and is being) conducted under the heading of "personality psychology," though it is proposed here that it would be usefully distinguished from research which does not meet these conditions by the more specific label "differential psychology." Some might take exception to this proposal but it is endorsed here in order to address the question posed earlier regarding the existence of a scientific paradigm in the study of individual differences; it is a non sequitur to evaluate this question for a domain in which the practitioners fail to agree about the qualities of "scientific" output. The question now considered is whether a paradigm exists in *differential psychology*.

1.4 Contributions from the major disciplines

> "[T]here are three absolutely irreducible faculties of the mind, namely, knowledge, feeling, and desire."
>
> *Immanuel Kant (1790)*

In order to emerge as the dominant paradigm in differential psychology, a theory would need to

(mostly) account for the range of the facts which are known regarding individual differences and hold up to extensive testing against predictions of human behavior while simultaneously demonstrating parsimony. Expressed through a range of affects, cognitions and desires, these differences are themselves the manifestations of various genetic and environmental influences which shape individuals over time. This genetic and environmental interplay suggests a dynamic relationship in which the "biological" differences shape and, to some extent, are shaped by the features of our environment, including our interpersonal relations. The universe of individuals differences extends therefore to include the ways in which individuals differ from one another physiologically and circumstantially as well as the ways in which they differentially relate to their environments, including other individuals. Developing a model to account for all of these nuances would be a formidable task.

A pragmatic beginning would be to integrate the distinct disciplines of individual differences research which have evolved since the late 1800s. At the highest level of abstraction, these disciplines map loosely onto the affective, cognitive and conative modes of behavioral expression. The classification of psychology according to these three categories has a long history, originating perhaps with the ancient Greeks (Brett, 1921) but more likely with Immanuel Kant's *Critique of Judgment* (Hilgard, 1980; Kant, 1790). This classification scheme was referenced commonly by psychologists in the late 19th and early 20th centuries (Hilgard, 1980; McDougall, 1923) before trailing off in popularity with the rise of behaviorism. Despite the drop in explicit references, each of these categories seems to have developed a distinct and well-established research tradition. In fact, in contrast to the dated claims of Eysenck and Eysenck (1985) and Loevinger (1987), "dominant" paradigms have emerged over the last few decades in each of these disciplines: the affective domain has produced the Big Five/Five-Factor Model (Costa and McCrae, 1992; Digman, 1990; Goldberg, 1990); the conative domain has developed the RIASEC model of interests (Holland, 1959, 1997); and consensus has begun to

coalesce for two similar models of cognition – the Cattell-Horn-Carroll (Carroll, 1993; McGrew, 2009) and Verbal-Perceptual-Rotation (Johnson and Bouchard, 2005) models.

The affective domain, which seems to be the topic that many researchers have in mind when using the term "personality," has traditionally been referred to as "temperament" (Clark and Watson, 2008; Heineman, 1995). Use of the term personality is once again problematic in this context for the same reason described earlier with regards to individual differences writ large – it is unfortunately vague. It implies the possibility that "non-affective" individual differences are excluded from personality. For this reason, the term temperament is used here (and recommended for use elsewhere in the context of individual differences research) to describe the range of emotional (affective) traits on which individuals differ.

It is also necessary to acknowledge the confusion introduced by developmental researchers (Thomas and Chess, 1977) seeking to distinguish adult "personality" from stable affective patterns in pre-adolescent children, especially infants and toddlers (Heineman, 1995; Shiner and DeYoung, 2013). This point precipitates consideration of several fundamental issues, including the degree to which temperamental differences are dispositional, hereditary (as opposed to environmental), "biological," and stable. These issues will not be reviewed extensively here except to state that temperamental differences are operationally viewed as relatively stable traits which have been found to be associated with various individual differences in neurobiological processes on both the molecular genetic (Krueger and Johnson, 2008; Krueger et al., 2008) and more broadly neuroanatomical levels (Canli et al., 2001; Canli, 2008). Further evidence supports the implication suggested by these biological associations – temperamental differences are innate and, depending on the trait, variability described by genetic and non-shared environmental factors is roughly the same (Clark and Watson, 2008). These data suggest that robust models of temperament should account for empirical claims that various differences are

more or less evident across the lifespan. Indeed, it has been posited that the Big Five traits (and perhaps one additional trait to account for "Activity Level") are well-suited for explaining temperament in infants and children (Shiner and DeYoung, 2013). In any case, the rationale for using the term "temperament" to describe affective differences in infants and children but not adults is not clear.

The cognitive and conative disciplines have traditionally been distinguished from temperament though, strictly speaking, both of these are influenced by affective variance. Research on individual differences in cognition has been a cornerstone of psychological research for well over 100 years (Lubinski, 2004), and also the most frequent source of controversy (Gould, 2006; Herrnstein and Murray, 1994). In fact, with few exceptions (Ackerman, 1997; Duckworth and Seligman, 2005; Lubinski and Humphreys, 1997), social and personality psychologists in the United States had largely abandoned cognitive ability research until a recent resurgence of interest (fortunately, this was not generally the case among personality psychologists elsewhere, particularly in Europe). Today, it is increasingly recognized that individual differences in cognitive ability are predictive of an impressive array of outcomes, including educational attainment, employment status, criminal behavior, marital status, staying healthy, recovery from ill-health and life-expectancy (Gottfredson, 1997; Lubinski and Humphreys, 1997; Deary et al., 2004; Herrnstein and Murray, 1994; Schmidt and Hunter, 2004).

Research on conative individual differences (i.e., differences in desires, motivations, volition and striving) is most frequently conducted through the assessment of interests, especially vocational interests. The dominant interests framework, known as the RIASEC model of vocational interests (Holland, 1959, 1997), organizes both interests and jobs according to six categories (and related scales) – Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. The framework itself allows for hierarchical organization of specific occupations which can be

grouped according to shared "basic interest" categories and these in turn can be grouped at a higher level of six general interest factors (Armstrong et al., 2004). In other words, the basic interests may be seen as equivalent to the facet level of the Big Five in the affective domain. It has also been suggested that the six factor structure can be further simplified to two dimensions which are known as "data/ideas" and "people/things" (Armstrong et al., 2008b; Prediger, 1982).

It should be noted that the assessment of vocational interests as a proxy for conation is practical but inadequate. It does not typically include the assessment of preferences, values, avocational interests or pastimes. More generally, the assessment of conative differences is hampered by the fact that specific activities are often idiosyncratically rooted in previous experience and are generally pursued sequentially, with varying degrees of intensity, in accordance with circumstantial factors. In other words, the use of interests to capture conative differences is problematic because (1) interest in a behavior or activity is often dependent on knowledge about that activity and (2) interest does not reflect the intensity with which an activity is pursued, the enjoyment derived from it, or the circumstantial factors which may impede or demand the pursuit of any given activity (e.g., socioeconomic status, cultural influences, etc.). Related to these issues is the fact that the various aspects of conation are seemingly quite distinct: the assessment of *interests* provides a means of describing one's preferences; *motivation* is generally framed as a measure of intensity (Carver and White, 1994; Gray and McNaughton, 2000); *goals* and *values* are often framed as trait-like heuristics that individuals use to navigate through the stream of choices in life (Higgins et al., 2001; Molden and Higgins, 2005; Peterson and Seligman, 2004).

The stability of these aspects of conation has also received relatively little treatment. The Dynamics of Action model (Atkinson and Birch, 1970) has been proposed for describing state-like variations, including daily and even momentary fluctuations, which directly influence

temporal changes in activity and action tendencies. More recently, this model has been re-parameterized in terms of cues, tendencies and actions (Revelle, 1986; Revelle et al., 2010a) and its efficacy has been simulated for social interactions (Fua et al., 2009, 2010) in an attempt to evaluate even broader models of approach and avoidance motivation, such as Reinforcement Sensitivity Theory (Corr and McNaughton, 2008; Gray and McNaughton, 2000) and control theory (Carver and Scheier, 1982).

1.5 The groundwork for an integrative paradigm

> “[I]t is generally admitted that all mental activity has these three aspects, cognitive, conative, and affective; and when we apply one of these adjectives to any phase of mental process, we mean merely that the aspect named is the most prominent of the three at that moment. Each cycle of activity has this triple aspect; though each tends to pass through these phases in which cognition, conation, and affection are in turn most prominent; as when the naturalist, catching sight of a specimen, recognizes it, captures it, and gloats over its capture.”
>
> *William McDougall (1923)*

1.5.1 Prior work towards integration

Unfortunately, the emergence of cognitive, conative, and affective paradigms has occurred without much regard for the degree to which these frameworks overlap or may be incompatible; these issues are exacerbated by inconsistent framing of related constructs across the various domains. Still, the general lack of cross-domain research of individual differences is by no means universal. There have been several influential efforts to evaluate the relationships across the previously mentioned domains, though the majority of these works have been isolated studies with small samples and have only evaluated two of the domains at a time.

Influential studies involving temperament and interests have typically involved joint

administration of the NEO-PI-R® (Costa and McCrae, 1992) and measures grounded in the RIASEC framework (the Vocational Preference Inventory and the Self-Directed Search®) (Barrick et al., 2003; Costa et al., 1984; Feist, 2012; Gottfredson et al., 1993; McKay and Tokar, 2012), though one recent study did explore the relations between lower-order facets and the RIASEC scales (Armstrong and Anthoney, 2009). The basic findings of this research, to the extent that it is consistent, suggests that significant correlations between the Big Five and RIASEC scales tend to be low and that it is therefore not appropriate to substitute the two measures for one another (Costa et al., 1984; Gottfredson et al., 1993). More recent results suggest that appreciably higher correlations can be found when using more narrow personality measures (Armstrong and Anthoney, 2009). Noteworthy correlations include positive relationships between the Social and Enterprising interests with Extraversion, and positive relationships between the Investigative and Artistic preferences with Openness.

Studies exploring the temperament/cognition relationship have been more varied. In the context of meta-analytic findings regarding the predictive validity of personality generally, Roberts and colleagues (Roberts et al., 2007) evaluated the relative influences of both Big Five traits and IQ (as well as socio-economic status) for educational and occupational attainment and identified significant correlations among several aspects. These and additional studies (Kuncel et al., 2010) specifically suggest that cognitive ability is slightly more predictive than pro-social personality traits for educational and occupational outcomes while personality traits (particularly Conscientiousness) are more predictive for outcomes related to health and longevity. In a theoretical review, DeYoung (2012) argues for the explicit inclusion of intelligence in personality models and suggests that many of the Big Five traits assess "abilities" broadly defined (e.g., Conscientiousness as an expression of the ability to delay gratification). He concludes that intelligence is most reasonably situated as an aspect of Openness and calls for more integrative empirical research.

Research on the overlap between cognitive abilities and interests has tended to evaluate specific educational and occupational outcomes. Several research groups have, for example, recently begun to evaluate abilities and interests related to science, technology, engineering and math ("STEM") outcomes. Data collected from the Study for Mathematically Precocious Youth (Lubinski and Benbow, 2006; Robertson et al., 2010; Wai et al., 2009) have been particularly informative in this regard as has the Spatial Intelligence and Learning Center (Hegarty et al., 2010; Uttal et al., 2013). Spatial ability assessment remains less mature than that of verbal and math ability however, despite growing recognition of the special importance of spatial skills (Lubinski, 2010). Better spatial measures and large scale assessments are needed to inform the ways that spatial interests and abilities interact developmentally, especially across genders (Newcombe and Shipley, 2012; Newcombe et al., 2013). These needs have also been acknowledged more generally for broader studies of interests and cognitive abilities as well (Johnson and Bouchard, 2009).

Attempts to integrate more than two domains of individual differences have been even more limited. Seminal meta-analytic work in this regard has been conducted by Ackerman and colleagues (Ackerman and Heggestad, 1997; Ackerman, 1997; Ackerman and Beier, 2003). Their efforts suggest two prominent lines of findings. First, there are significant commonalities across the domains of interests, cognitive abilities and temperament, as variously described historically. This was particularly true with regards to relationships between temperament and cognitive abilities (positive correlations between Extraversion and Openness with abilities; negative correlations between Neuroticism and abilities). Second, it is possible to identify trait complexes which likely result from various developmental trajectories – indeed, four such trait complexes have been detailed (Ackerman and Heggestad, 1997; Ackerman and Beier, 2003). These complexes are presumably rooted in complementary temperamental and ability dispositions that contribute to interest in specific tasks over time. It should be noted that one of

the proposed rationales for the use of trait complexes is the difficulty of assessment and analysis across the three domains when using traditional data collection methods. Attempts to replicate and extend Ackerman's work are greatly needed, particularly with larger samples.

A second, more theoretical contribution has been proposed by Roberts (Roberts, 2006). This "neo-socioanalytic" model identifies three domains – Traits, Values and Abilities – which are similar in content to the domains described above (temperament, cognitive abilities and interests) with the exception that they explicitly encompass the domain of identity, as assessed with narratives (McAdams, 2001). This model also contains several additional features which extend beyond the more narrow boundaries of differential psychology set forth in Section 1.3; examples include reputational features of personality and person-organization fit. Nevertheless, the organizational framework of individual differences domains (referred to as "units of analysis" in the neo-socioanalytic model) is noteworthy for its similarity to the one used here.

As a conclusion to this review of prior efforts towards integration, it is proposed that the historical lack of communication between disciplines has created an opportunity for substantial advancement of knowledge. Such an advancement would reflect a more nuanced understanding of the manner in which constructs relate to one another across domains and the manner in which criterion variables are differentially (or similarly) predicted by such constructs. The obstacle to integration however has not been a lack of consensus over theory so much as the methodological difficulty of cross-domain assessments which are simultaneously broad and specific.

1.5.2 Challenges to empirically-informed integration and recent innovations

The primary source of difficulty when evaluating across domains is data collection. Clinical psychologists and neuropsychologists overcome this challenge by assessing their patients with

extensive commercial batteries of tests, often including personality measures like the Minnesota Multiphasic Personality Inventory®-2 (Butcher et al., 2003) or the NEO-PI-R® (Costa and McCrae, 1992), cognitive measures such as the Wechsler Adult Intelligence Scale® (Wechsler, 2008), vocational measures like the Strong Interest Inventory® (Harmon et al., 1994), and perhaps a sampling of additional measures evaluating psychopathology, aptitudes or psychophysiological functioning. While this approach can be very effective for cross-domain assessment of a single individual, it is not well-suited for large-scale differential psychology research because it is both expensive and time-consuming. All of the tests described above are commercial measures and each requires an average of 90 to 190 minutes for administration, scoring, and interpretation by a licensed practitioner (Camara et al., 2000).

Prior research in the affective, cognitive, and conative domains has required dramatically different methods. Perhaps the most important methodological distinction stems from the use of samples which are convenient to university-based researchers – college students and, to a lesser extent, community samples. While this allows for assessment across far more individuals than the clinical approach, it still usually suffers from issues of representativeness and insufficient size for detecting small but stable relationships between constructs and evaluating the fit of models with many parameters (Kenny, 2012). These concerns are exacerbated when attempting to detect more complex relationships between multiple, lengthy measures because participants of this type are rarely willing to participate in studies lasting more than a couple of hours. In essence, the qualities of the clinical method (extensive testing of a few participants costing considerable time and money) and the traditional research method (more participants at little cost but relatively few measures) are both lacking when it comes to effective cross-domain assessment.

It seems that, beneath the over-arching difficulty of data collection across multiple domains in

differential psychology, there exist three underlying challenges. The first of these relates to the need for samples which are large and relatively "representative" of the broader population (or at least the population of interest). Second, the use of large samples precipitates the need for measures which can be administered for little or no cost. The last challenge relates to the need to administer a large number of variables across the sample in order to evaluate the structure across domains. Fortunately, innovative solutions for addressing each of these challenges have been developed over the last two decades, as briefly described in the following sections.

1.5.2.1. Telemetric assessment

The number and variety of techniques for collecting data from large samples has increased dramatically since the beginning of the "internet-era," largely because it has become increasingly easy to reach participants outside of the research laboratory (Wilt et al., 2011). Web-based methods have demonstrated improved sample characteristics in terms of both size and breadth with little loss of validity (Fraley, 2004; Gosling et al., 2004; Skitka and Sargis, 2006). Many large-scale, research-driven internet surveys now exist and several of these have collected samples of unprecedented size – hundreds of thousands of participants or more (Condon and Revelle, 2014; Gosling et al., 2004; Kosinski et al., 2013; Peterson et al., 2005; Revelle et al., 2010b; Sandy et al., 2013; Soto et al., 2010; Wang et al., 2012). With few exceptions (Condon and Revelle, 2014; Kosinski et al., 2013; Revelle et al., 2010b), the data collected from these samples has been limited to short questionnaires which assess constructs from only a singular domain. Most of these also make use of traditional website frameworks, though it has become increasingly common to collect data from mobile devices (Wilt et al., 2011). These include both older technologies such personal data assistants and SMS-enabled phones as well as more modern devices such as smartphones and tablet computers.

While the number of research groups collecting very large samples has been relatively limited, many more have benefitted from the use of more modest telemetric techniques. These include third-party tools such as survey software providers (e.g., SurveyMonkey and Qualtrics) and recruitment services (Amazon's Mechanical Turk) as well as the ability to electronically access data collected in very large panel studies (e.g., the Programme for International Student Assessment (Anderson et al., 2007; OECD, 2012) and the General Social Survey (Smith et al., 2011)).

1.5.2.2. The development (and aggregation) of public-domain measures

Concomitant with the need to collect data from large samples, it has been increasingly necessary to utilize measures which are not burdened by the costs which are typically associated with using copyright-restricted scales. In conjunction with his proposal for more rapid advancement in personality, Goldberg (Goldberg, 1999) introduced a large pool of personality items for use in the public domain and which were designed to accommodate assessment needs across a broad range of constructs. Historical dependence on copyright-protected measures, as Goldberg has argued (Goldberg, 1999; Grucza and Goldberg, 2007), reduces progress because the owners of these proprietary measures have little incentive to consistently revise or validate them extensively against other measures. Over time, the number of commercial measures has proliferated while few of the most established have been improved to account for findings from novel research. These problems can be avoided through the use of public-domain measures in that the items will be developed, administered and improved by the research community at large.

The International Personality Item Pool ("IPIP") now contains more than 2,500 items and has come to be used widely within personality research (Goldberg, 2014). These items, when

supplemented by an additional 1600 items from various sources (mainly from shorter scales of more narrow focus), form a database of more than 4000 temperament items in total. This database does not include scales designed to assess the cognitive and conative domains (at least, not to the extent that these are distinct from the affective). Public-domain scales of Interests have only recently been developed; they include the Oregon Vocational (Pozzebon et al., 2010) and Avocational Interest Scales (Goldberg, 2010) and O*NET Interest Profiler (Armstrong et al., 2008a; Rounds et al., 2010) which together number approximately 500 items.

Public-domain items for cognitive ability have not previously been available, in part because this type of assessment is considerably different from items which ask participants about their typical behavior or attitudes. Cognitive ability measures, by contrast, attempt to assess the level at which an individual "maximally" performs (Condon and Revelle, 2014). In these cases, items are not only copyrighted for their commercial value but also for the sake of test security. As such, efforts to develop and validate public-domain items have recently been pursued by the present author and colleagues (Condon and Revelle, 2014; Revelle et al., 2010b; ICAR, 2014). The challenge to item development lies in the fact that the items are to remain in the public-domain while still maintaining adequate validity. This is accomplished with modern item-generation techniques (Arendasy et al., 2006; Dennis et al., 2002) that make use of algorithms which dictate the parameters of new items with predictable difficulty and in many alternate forms. These techniques allow for the creation of item types where the universe of possible items is very large. This, in turn, reduces the threat to validity that arises from item disclosure. These techniques can even be used to enhance test validity under administration paradigms that expose participants to sample items prior to testing and use alternate forms during assessment as this methodology reduces the effects of differential test familiarity across participants.

The first validation (Revelle et al., 2010b) of these efforts was based on the administration of a

preliminary set of 56 items to more than 65,000 participants. These procedures (as well as description and validation of more recently created item types) have since been refined (Condon and Revelle, 2014) and many more item types are now under development as part of an internationally-funded collaboration to develop the "International Cognitive Ability Resource" (ICAR, 2014). At the current time, the resource includes 60 items which are intended to assess four different constructs within cognitive ability: (1) Matrix Reasoning; (2) Verbal Reasoning; (3) Letter and Number Series; and (4) Three-Dimensional Rotation. Validation results (Condon and Revelle, 2014) suggest that correlations between these items and other measures of cognitive ability are promising; correlations range from 0.4 to 0.5 with self-reported achievement test scores and 0.8 with a brief commercial IQ measure, the *Shipley-2* (Shipley et al., 2010), after correcting for restriction of range.

In addition to the ICAR items, the IPIP items assessing temperament, and the vocational and avocational interest items, several scales have recently been developed for assessing mental and physical health outcomes. These include the Patient Reported Outcomes Measurement Information System (PROMIS), organized around the domain-mapping framework of the World Health Organization for physical, mental and social health (Cella et al., 2007), and the Personality Inventory for DSM-5 (the PID-5), which has been designed to assess the first empirically based model of maladaptive personality traits (Krueger and Markon, 2014). When considered together, this growing pool of resources provide a pool of freely available items of unprecedented breadth for individual differences assessment.

1.5.2.3. Synthetic Aperture Personality Assessment ("SAPA") sampling procedures

While telemetric assessment techniques have meaningfully improved the ability to collect larger and more diverse samples, they have not generally been applied to collect data across wider sets

of individual differences variables. In other words, they have been used to increase sample sizes (n) but not to increase the number of variables administered (i). This is because they do not inherently provide a means of assessing participants on a large pool of items without over-burdening individual participants.

Synthetic aperture personality assessment represents a variation on the standard method of web-based assessment and is perhaps best explained by analogy to the technique on which it is based in radio and optical astronomy. An historically problematic issue in these fields stemmed from the fact that the resolution of a telescope is limited by its diameter. This resolution can be functionally increased by combining input from multiple, linked sites into one coherent image. Effectively, a very large telescope is created by synthesizing the input from many smaller ones. A prototypical example of this in radio astronomy is the Very Large Array in Socorro, New Mexico where 27 relatively small (25 meter) radio telescopes are spread out in a Y-shaped configuration to simulate the resolution of a 36 km telescope. In optical astronomy, similar techniques are used in inferometry at the Keck Observatory in Hawaii.

Analogous techniques are available for data collection over the internet. Rather than combining signals from the same source using different telescopes as is done in astronomy, the structure of personality can be studied by combining the responses of many people across more items than any one person is willing to answer. Instead of observing celestial objects beyond the visible range, psychologists can observe the relations between personality constructs which would not otherwise be visible given practical assessment constraints. This can be done by sweeping the assessment "telescope" across a wide range of constructs or by focusing for short periods of time on high-priority topics.

This procedure is not without precedent. Lord (1955) has previously described theoretical procedures for the sampling of items (rather than participants) in the context of testing and

similar sampling techniques have long been used by the Educational Testing Service in order to develop new achievement test items. The latter is done by administration of small, random subsets of items to subsamples of test-takers. At ETS, these items are typically under evaluation for discriminant and concurrent validity among the items in their proprietary set, though this is suggestive of an additional methodological innovation in its own right. That is, the advent of broadly used, public-domain scales of individual differences.

1.5.3 Combining these innovations via SAPA-Project.org

A web-based application at SAPA-Project.org has been developed to make use of synthetic aperture measurement techniques with public-domain measures of individual differences administered over the internet. In practice, the true value of these methodological innovations can only be appreciated when applied to contexts involving many participants. Thanks in large part to web traffic for related websites (mainly personality-project.org and especially personality-project.org/r), the SAPA-project.org website has averaged approximately 140 unique participants each day since May 20, 2013. In exchange for customized feedback about their personality, participants anonymously provide data on more than 25 demographic variables and respond to an average of 155 items assessing temperament, cognitive ability, and interests. In keeping with SAPA procedures, the items are chosen as semi-random subsets of the much larger group of items under concomitant administration (about 800 items total). Very large, synthetic correlation matrices are formed on the basis of these "Massively Missing Completely at Random" responses from many participants over time (approximately 8.8 million data points per year at the current rate).

It is not accurate to say that this sample is necessarily representative of any population other than those individuals who want to take internet-based personality surveys, but it is more

demographically diverse than the samples typically available to university researchers. For example, the 97,000 person sample collected between August 18, 2010 and May 20, 2013 includes participants from 199 countries, 34 of which are represented by more than 100 participants. Approximately 66% of the sample is female (consistent with broader web-traffic trends) and 78% is from the United States. Among Americans, 32.6% represent ethnic minorities. Median and mean ages are 22 and 26 years old respectively (sd = 10.6). Additional categories of data collection include educational and occupational outcomes, parental education and employment information, marital status, height, weight, health data (smoking, exercising, sleep patterns), and self-reported achievement test scores. As may be evident based on the methodological techniques described, these procedures have been refined after several years of online data collection. In total, data have been collected from more than 300,000 participants to date.

1.6 Application of the SAPA-Project to develop a testable integrative model

> “[W]e like to think of breakthrough ideas as sudden accelerations on the timeline, where a genius jumps ahead fifty years and invents something that normal minds, trapped in the present moment, couldn’t possibly have come up with. But the truth is that technological (and scientific) advances rarely break out of the adjacent possible; the history of cultural progress is, almost without exception, a story of one door leading to another door, exploring the palace one room at a time.”
>
> *Steven Johnson (2010a)*

If the primary challenge to the development of a testable and integrated model of individual differences is methodology, paradigm adoption in differential psychology may be a matter of evolving consensus on the heels of incremental technological improvements rather than theoretical “revolution” (to use Kuhn’s terminology). The SAPA Project represents technological improvement in data collection in that it is well-suited to evaluation of the structure of the

multidimensional space that is described by the many public-domain items that have been (or are currently being) developed across the affective, cognitive and conative domains.

The goal of the current project – that which is described in the remaining chapters – is to set forth an empirically-informed, integrated assessment model. It is "empirically-informed" in that it makes use of data which have recently been collected through the SAPA Project based on the administration of modern, widely-used measures in each of these three primary domains of individual differences research. It should be noted that the cross-domain assessment model proposed herein will be preliminary as well as overlapping in the case of some constructs and incomplete with regards to several more. Iterative refinement will no doubt be needed.

Development of such an assessment model requires the identification or reification of consensual models in each domain; these should include the right balance of theoretical breadth and parsimony across a range of predicted outcomes. In the affective domain, this work will draw heavily on the Big Five and Big Six models generated by research in the psycholexical tradition. Given the large number of measures which have grown out of this tradition, the first step in this process (Chapter 2) will be to consider the extent to which the structure of several widely-used public-domain scales matches the proposed structure of the universe of trait descriptors (the Big Five). The data collected to explore this question will also be used to consider revisions to several of these widely-used scales in order to enhance their simple structure. In Chapter 3, the issue of structure in the affective domain is explored more broadly based on the concomitant administration of eight set of broad-bandwidth scales to a large international sample. This data set is then used to propose a hierarchical assessment model for the affective domain which allows for description of personality at various levels of specificity.

The next two chapters address the need to include other domains in the assessment of individual differences. Chapter 4 proposes and describes procedures for developing a

public-domain assessment model of cognitive abilities. Evaluating the degree to which the proposed measures capture the structure of cognitive abilities is complicated by the limited availability of public-domain items, though preliminary evidence suggest that the proposed assessment model is consistent with existing theory. The conative domain is addressed in Chapter 5 where the relationship between two extant public-domain assessment models of vocational interests is considered. An integrated model is then proposed in order to better describe the underlying empirical structure of vocational interests based on joint administration of these measures to a large international sample.

The final chapter summarizes these efforts across domains by describing features of the complete assessment model (226 items across three domains). It also considers some of the ways in which this integrated, cross-domain assessment model should be tested and refined. The most important suggestions for future research include empirical testing of the extent to which an integrated model can predict a wide range of specific "real-world" behaviors (a wider range than existing Big Five models, for example) and make use of efficient administration protocols (e.g., item-response theory-based test reduction and computer-adaptive testing). It is hoped that the proposals suggested here for integrated assessment of the affective, cognitive, and conative domains will eventually lead to the development and widespread use of a collection of brief predictive measures of individual differences writ large and that the field of differential psychology will enter an era of empirically-testable paradigms that is no longer bound by compartmentalized, domain-specific research.

Chapter 2

The affective domain: Item clusters and complexity in personality scales

2.1 Introduction

One of the fundamental challenges when constructing multidimensional sets of scales for use in psychological assessment is the need to achieve both internal consistency within the scales and parsimoniously describe the structure of the phenomena under investigation. We argue that item clusters plays an essential role in both sides of this balance, and that recognition of this circumstance suggests that existing broad-bandwidth measures of personality can be improved by use of measures of complexity.

On one side of the balance, the development of internally consistent scales is typically accomplished through the inclusion of items (either self/informant-report questions about behavior or performance-based measures of ability) which are similar in terms of content and/or format, with some subtleties perhaps for identifying contextual variability in the construct under measurement (Revelle and Condon, ress). Achieving parsimonious description of the underlying latent structure, on the other hand, is considerably more difficult. Henry Kaiser once

said that solving the number-of-factors problem is easy enough to be done every day before breakfast (even in the days before personal computers), but that the difficulty lies in identifying a solution others will regard as *right* (Horn and Engstrom, 1979; Revelle, 2014). This challenge essentially stems from the technical impossibility of identifying true structure in the absence of anything other than artificial data and the resultant impossibility of describing it parsimoniously with a high degree of accuracy (to say nothing of certainty).

In scale terms, clustering relates to internal consistency in that the "tightness" of a given cluster is directly related to the extent to which that cluster can be described by a scale with a single factor. In other words, scales with high item inter-correlations will describe tight clusters much better than spaces with "looser" structure. When sets of scales are administered together, with each measuring a relatively tight cluster, there are often interstitial spaces which are not well described by any single scale in the set. A strict interpretation of this circumstance is that every latent construct other than those which are directly assessed by a scale in the set has some degree of interstitiality. To use a single set of Big Five scales as an example, only specifically-defined versions of Conscientiousness, Agreeableness, Neuroticism, Extraversion, and Openness are well-described by the 8 to 10 items of one of the Big Five Inventory scales (John and Srivastava, 1999), while all other constructs (including the many facets of these traits) require blends.

The typical goal when developing a set of scales is parsimonious description of a broad universe of underlying variables and most of the scales used for personality assessment have been developed through the use of procedures which are similar to those described by Comrey (1961; 1973; 1984), Harman (1976), McDonald (1985), and Goldberg & Velicer (2006) and many others. These can be summarized as follows: (1) identify the constructs to be investigated; (2) create or identify a homogenous item set for each construct; (3) administer the item sets for all constructs

(ideally to a large representative sample) and factor analyze the resulting covariance matrix; (4) drop those items which do not have high loadings on any of the factors or which have high secondary loadings. The set of remaining items will include "factorially-homogenous item dimensions" (Comrey, 1961) that relate to the constructs under investigation and which allow for the use of scale scores with greater reliability than individual items.

In the development of broad-bandwidth measures of personality, none of these steps are inconsequential. Personality researchers have resolved the seemingly intractable difficulty of the first step – identifying the variables under investigation – by invoking the Lexical Hypothesis (Galton, 1884; Goldberg, 1981, 1993b) to posit basic assumptions that allow circumscription of the universe of personality variables. The result of this solution has been considerable advancement for the field of personality research (Goldberg, 1993b; John and Srivastava, 1999; Roberts et al., 2007) though some have acknowledged that this presumption may not precisely reflect the scope of personality in nature (Block, 1995; Uher, 2013).

The second step – identification of item sets for administration – is particularly sensitive to clusters in the underlying structure as it makes it difficult to determine which items should be chosen from the full universe of items available. Several strategies have been used by personality researchers in attempts to address this issue; see Goldberg (p. 28, 1992) for a summary of those used by developers of factor markers for the multidimensional trait descriptor space. In order to circumvent the challenges of administering *all* trait descriptors, various researchers have employed representative (Goldberg, 1990; Peabody, 1987), uniform (Wiggins, 1979), or cluster-based sampling (Goldberg, 1992; John and Srivastava, 1999; Norman, 1963) of the descriptors in order to assess their relative structure. While the method used to sample items for structural analyses has no bearing on the structure of the phenomena under investigation (the trait descriptors, for example), it will have a consequential effect on the resultant factor

solutions by virtue of determining the items to be administered.

Strictly speaking, detection of the presence of clusters in the items *after* they have been chosen in the second step (note that this may be much different from the extent of clustering in the full item universe depending on the sampling method used) is not directly related to the factor analyses of the third step, but it is intimately tied to subsequent rotation of the factors. Despite being frequently confounded in statistical software, factor extraction and factor rotation are importantly distinct in that the former attempts to answer the question of "how many" while the latter determines "how they are oriented." The issue of orientation has no bearing on the degree to which a given factor model fits the underlying data, but it will affect the content of the scales which result from retention of the most highly loaded items for each factor (Step 4 above) (Thurstone, 1947, 1954; Carroll, 1953).

For any given factor analytic model, however rotated, it is possible to identify the orientation of each individual item relative to the axes of each factor and the origin; this is item *complexity*. While multiple methods exist for quantifying complexity (Browne, 2001; Cattell, 1952; Hofmann, 1977, 1978; Kaiser, 1974; Velicer, 1976), most of these were developed for evaluation of complexity at the level of the factor solution. Hofmann's (1977; 1978) index provides an intuitively appealing method for evaluating complexity at the item-level. This method defines item complexity, c_i, as a function of the number of factors, the factor rotation/transformation, and the resulting factor loadings, as follows:

$$c_i = (\sum_{j=1}^{r} a_{ij}^2)^2 / \sum_{j=1}^{r} a_{ij}^4$$

where *r* is the number of factors and *a* is the item loadings (on each factor).

Items with the lowest possible complexity ($c_i = 1.0$) would be located "on" one of the axes on the exterior of the multidimensional space containing all of the items factored. In terms of factor

analytic output, such an item would have a high primary loading (± 1.0) and low secondary loadings (0.0). Items would demonstrate higher complexity to the extent that they are distant from the axes (in interstitial space) and/or closer to the origin; an item's complexity will increase as its primary loading decreases and/or its secondary loadings increase.

The complexity of any specific factor solution can be seen as an index of the complexities for all of the items factored. It should be noted that item-level complexity and rotation are closely related to one another (and to the issue of item clusters) in that the complexity of any given item will change as the factors are rotated. While it is possible to rotate the axes of a factor solution in order to decrease the complexity for any single item, this will simultaneously alter the complexities for all of the remaining items as well as the overall complexity for that factor solution.

If the items are clustered, it is generally the case that the axes of the unrotated factor solution will not go through the clusters (exceptions are possible, especially in cases of over-extraction). This is because the first factor will attempt to maximize the explained covariance of all items and each subsequent factor will be orthogonal to those which have already been extracted. The fact that unrotated factor solutions tend not to go through the clusters makes them difficult to interpret. Most scales are developed based on factor solutions which have been rotated and the primary distinction between rotation methods hinges on the issue of orthogonality.

Orthogonal rotational methods, as the name implies, maintain the characteristic that each axis is at right angles to all other axes; the resulting factors are therefore uncorrelated. Oblique transformational methods[1] are not constrained by this restriction but instead allow the axes to orient such that the factors represent the tightest clusters, potentially resulting in correlated factors. This implies that oblique transformations generally allow for solutions with lower

[1] Strictly speaking, oblique "rotations" are transformations of orthogonal solutions and are therefore referred to as oblique transformations.

complexity in that they result in higher primary loadings and lower secondary loadings of items on average (this is also referred to as "simple structure", Thurstone, 1947, 1954; Carroll, 1953).

The selection of rotational/transformational methods is typically regarded as a matter of preference for scale developers or perhaps one which should be context dependent (Goldberg and Velicer, 2006). As Tucker & MacCallum (Ch.10, 1997) point out, the rationale for orthogonal rotations is not entirely clear as they do not represent the underlying structure of the factored items as well as oblique rotations. In cases where the underlying items *are* orthogonally oriented, oblique transformations will identify them as such. On the other hand, only oblique transformations will allow for detection of an underlying general factor. If "simple structure" is preferred, as most researchers and scale developers seem to agree, oblique transformations should be preferred as well. The most conservative recommendation might be to evaluate both the orthogonal rotation and oblique transformations, though it has been noted that the orthogonal rotation is generally only chosen after such a comparison if the difference between the two is inconsequential (p. 205, Gorsuch, 1983)

Rotation to simple structure may be viewed therefore as a function of item complexity; minimizing mean complexity is a rotational criterion for simple structure. This also means that it is possible for an item to have a low factorial complexity despite subjectively high *conceptual* complexity (or vice versa). The issue of item-level factorial complexity is typically given far less consideration than the scale-level complexity of various factor analytic solutions, though this oversight can mask variability in the underlying structure of a scale.

Assessment models with simple structure would have a mean item complexity of 1.0 in the optimal case as this would imply that items have high loadings on only one factor and negligible loadings on all other factors (Carroll, 1953; Thurstone, 1954). Goldberg (1993a) has noted that achievement of this optimal outcome is not likely for psychological attributes (largely due to the

reality that most items are at least somewhat factorially complex), though this has not discouraged the majority of test developers from selecting items based on Thurstonian principles of simple structure (Thurstone, 1947). The two-dimensional item circumplex approach of the AB5C (Hofstee et al., 1992), for example, was constructed with the explicit goal of developing items that are distributed around each of the 10 possible two-dimensional planes for five factors; this would theoretically reflect an average item complexity of 1.41. The expected values can also be estimated for cases with equally distributed items around a three-dimensional sphere (1.65) as well as four- and five-dimensional hyper-spheres (2.00 and 2.34).

Unfortunately, evaluations of complexity in samples other than those on which the scales were initially developed are rare after the scales have been published (Pettersson and Turkheimer, 2010). Often, this shortcoming results from the proprietary status of the underlying scales because replication studies are costly and the test-owners are not highly motivated to incorporate revisions (Goldberg, 1999). Fortunately, the recently developed International Personality Item Pool (Goldberg, 2014, 1999) is not subject to these constraints. To the contrary, the procedures for developing scales with IPIP items have been transparently described in detail (Goldberg et al., 2006) and suggestions for revision are encouraged.

2.1.1 Goals of the current studies

Two studies of complexity in several IPIP scales are described below. The first of these studies evaluated the relative complexity of several scales on the item level. It should be noted that, while related to the topic of model complexity, item-level complexity is distinct in that it allows for comparative evaluation of items within a given model or across models of the same dimensionality. More specifically, item complexity was evaluated in order to determine (1) the extent to which each of four IPIP-based measures approximates simple structure using various

rotations and (2) the relative complexity of their underlying items.

The second study evaluated the potential for using complexity analyses to improve upon extent scales. This was accomplished by using factor extension procedures whereby the factor structure for a set of scales was extended on the remaining items administered. For example, the five-factor solution empirically derived in Study 1 for the 100 items of the IPIP Big Five Factor Marker scales (the "IPIP100", Goldberg, 1999) was extended into the multidimensional space resulting from administration of all 373 items (the IPIP100 items and the remaining 273 items from the other scales). In other words, this technique was used to evaluate whether the scales which were constructed using the original IPIP data set (the Eugene-Springfield Community Sample) provided the simplest structural representation possible in large, alternate samples.

2.2 Study 1

The primary goal of Study 1 was to evaluate the complexity for each of four widely-used IPIP measures. This included evaluation of the mean complexity for each measure and the relative complexity of the underlying items in each set of scales. Secondary goals included the replication of the prescribed factor structure for each set of scales and examination of the correlations among all scales in large international samples.

2.2.1 Predictions regarding complexity

It should be noted that these measures were each constructed with slightly different goals in mind and that these design procedures were expected to affect the relative item complexities. The IPIP100, for example, was designed to approximate the relationships between the Big Five factor markers, suggesting that the average item complexity would be low (closer to 1.0 than the

expected two-dimensional circumplex complexity of 1.41). The Big Five Aspect Scales ("BFAS", DeYoung et al., 2007), by contrast, were designed as a hierarchical measure which might be expected to reflect both 5 and 10 factors. These items were also chosen with less strict procedures for eliminating items with high secondary loadings across the aspects (see p. 886, DeYoung et al., 2007). As such, the average item complexity for the five factor solution using the BFAS was expected to be higher than that for the IPIP100. Similar predictions were made for two six factor scales; the 48 item Questionnaire Big Six ("QB6", Thalmayer et al., 2011) was expected to be less complex than the 240 item IPIP-HEXACO (Ashton and Lee, 2007) as the latter set of scales was developed on the basis of a hierarchical structure with 24 facets beneath the 6 higher order factors.

2.2.2 Method

2.2.2.1 Participants

Two independent samples were used in Study 1.[2] Sample 1 included 42,272 individuals (54.7% female) from 178 countries who completed an online survey at SAPA-Project.org between January 22, 2013 and December 7, 2013 in exchange for customized feedback about their personalities. Sample 2 included 23,681 individuals (64% female) from 172 countries who completed the survey between December 8, 2013 and July 26, 2014. All data were self-reported. The mean self-reported age was 25.3 years (sd = 9.9, median = 22) with a range from 14 to 90 years in Sample 1 and 26.6 years (sd = 11.1, median = 22) with a range from 14 to 90 years in Sample 2. Educational attainment levels for the participants are given in Table 3.2. The largest group of participants were current university school students, though a wide range of educational attainment levels were represented. Race/ethnicity is presented for U.S. participants

[2]Both of these samples are available for further analysis. Sample 1 is included as part of the Supplementary Materials and Sample 2 can be obtained by contacting the first author.

Table 2.1: Participants by educational attainment in both samples

	Sample 1				Sample 2				U.S.
	% of total	Mean age	Median age	% Female	% of total	Mean age	Median age	% Female	% of total
Less than 12 years	14.3	17.1	17	55.5	13.2	17.3	17	62.3	14.8
High school graduate	8.3	22.4	18	43.4	9.3	22.7	18	57.0	28.5
Currently in college/university	43.3	23.6	21	60.5	41.1	24.3	21	68.5	NA*
Some college/university, but did not graduate	6.8	30.8	27	46.3	6.1	33.6	30	57.2	21.4*
College/university degree	14.7	31.2	28	50.7	15.1	33.8	31	62.6	25.1
Currently in graduate or professional school	5.3	28.6	26	52.6	5.4	29.4	26	65.3	NA
Graduate or professional school degree	7.5	37.4	34	52.6	9.8	39.5	37	60.1	10.3

U.S. data from the 2009 American Community Sample of the U.S. Census Bureau (Bureau, 2012).
* ACS data does not differentiate between those who are active students and those who are no longer enrolled.

in Table 3.3; participants from outside the United States were not prompted for information regarding race/ethnicity.

Figure 3.1 shows the geographic distribution of participants in both samples from the continental United States who provided optional ZIP code information (80.9% of U.S. participants and 55.2% of all participants in Sample 1 and 97.7% of U.S. participants and 60.1% of all participants in Sample 2). The correlation of ZIP code distributions between Sample 1 and Sample 2 was 0.57.

The correlation between ZIP code distributions in both samples combined and the U.S. population based on U.S. Census data (Census, 2011) was 0.78 when using the broader 3 digit ZIP codes (known as the regional prefixes, of which there are approximately 890). Note that these calculations required matching of ZIP codes to the U.S. Census Bureau's Zip Code Tabulation Areas.

2.2.2.2 Measures

Four sets of scales from the International Personality Item Pool were administered: the 100 IPIP items corresponding to the Big Five factor markers (Goldberg, 1999), the 100 items of the Big

Table 2.2: Participants by race/ethnicity in both samples

	Sample 1		Sample 2		
	Count	%	Count	%	% of U.S.*
African-American	2,194	7.6	1,329	9.1	12.2
Asian-American	1,300	4.5	775	5.3	4.4
Hispanic-American	2,399	8.3	1,284	8.8	15.7
Native-American	252	0.90	137	0.90	0.90
White/Caucasian	18,442	63.9	8,291	56.9	64.9
Multi-ethnic	1,677	5.8	809	5.6	1.7
Other	324	1.1	185	1.3	NA
Not specified	2,285	7.9	1,763	12.1	NA

* 2009 U.S. data from the U. S. Census Bureau (2012)

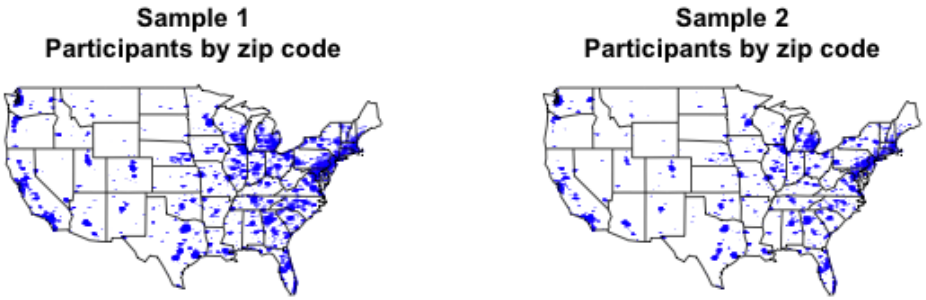


Figure 2.1: Participants by ZIP code for the continental United States

Five Aspect Scales (DeYoung et al., 2007), the 240 items of the IPIP-HEXACO inventory (Ashton et al., 2007), and the 48 items of the Questionnaire Big Six scales (Thalmayer et al., 2011). Administration of these four scales also implies the administration of several other measures which are abbreviations of these scales, including the 24 and 36 item Questionnaire Big Six scales (Thalmayer et al., 2011), the 50 item IPIP scales corresponding to the Big Five factor markers (Goldberg, 2014), and the 20 item "mini-IPIP" scales (Donnellan et al., 2006). None of these shorter scales were directly evaluated in Study 1; such analyses could be conducted using the same procedures described below and the data provided in the Supplementary Materials.

The 488 items from these measures contain 115 duplicates, resulting in a total set of 373 unique items. Of these, 279 items are in only one set of scales, 76 items are included in two sets of scales, 15 items are in three, and 3 items are in all four sets of scales ("Have a rich vocabulary", "Like order", and "Get angry easily"). All of the items were administered with the same six response options ("Very Inaccurate", "Moderately Inaccurate", "Slightly Inaccurate", "Slightly Accurate", "Moderately Accurate", "Very Accurate").

The items were administered using the Synthetic Aperture Personality Assessment ("SAPA") technique (Revelle et al., 2010b), a variant of matrix sampling procedures discussed by Lord (1955). This method produces data which contain "massive missingness" by design (Revelle and Brown, 2013). This missingness qualifies for classification as missing completely at random ("MCAR", Graham, 2009) and it is further described as massively missing because the mean level of missingness by participant was approximately 84% in Sample 1 and 86% in Sample 2. The items were presented to participants in random order, and participants responded to as many items as they wished. The mean number of items to which participants responded was 60.0 (sd = 38.3; median = 48) in Sample 1 and 52.6 (sd = 23.8; median = 48) in Sample 2. The number of items administered to each participant was procedurally independent of participant response

characteristics. The number of administrations for each item varied considerably in both samples (Sample 1: median = 6,318; m = 6,797; sd = 1,111; Sample 2: median = 2,802; m = 3,342; sd = 878) as did the number of pairwise administrations between any two items in the set (Sample 1: median = 1,480; m = 1,525; sd = 251; Sample 2: median = 534; m = 561; sd = 177). The minimum number of pairwise administrations among items (Sample 1: 1,066; Sample 2: 281) provided sufficiently high stability in the covariance matrix for the structural analyses described below (Kenny, 2012).

2.2.2.3 Analyses

The first step in this study entailed exploratory factor analyses ("EFA") for each set of scales in both samples in order to evaluate the fits for factor solutions based on the extraction of 1 to 20 factors. The EFA results reported below were based on the Pearson correlations between scored responses using Ordinary Least Squares ("OLS") regression models with oblique rotation (Revelle, 2014). Variations on these factor analytic methods are demonstrated in the analytic summary included as part of the Supplementary Materials.

Goodness-of-fit was evaluated using the Minimum Average Partial criterion ("MAP", Velicer, 1976), the Root Mean Squared Error of Approximation ("RMSEA", Hu and Bentler, 1999), the Standardized Root Mean Square of the Residual ("SRMR", Hu and Bentler, 1999), and an empirically-derived measure of the Bayesian Information Criterion ("eBIC", Schwarz, 1978; Revelle, 2014). For all of these fit statistics, lower values indicate a superior fit, though the MAP and BIC will often indicate a localized minimum while the RMSEA and SRMR values will decrease as more factors are extracted. In the latter cases, good fits are typically indicated by RMSEA values of 0.05 and SRMR values of 0.08 (Kenny, 2012). Measures of item-level complexity are based on the Hofmann (1977, 1978) complexity index and the complexity

reported for a given factor solution reflects the mean of the item-level complexities.

For each set of scales, factor congruences across the two samples were calculated based on solutions using the expected number of factors (e.g., factor congruences were calculated based on the five factor solutions for the IPIP100). Intercorrelations between the scales were also calculated based on the prescribed scoring procedures for each set (note that these are independent of the factor solutions). These correlations between scales were corrected for item overlap in order to account for the spurious effects of shared items (Bashaw and Anderson, 1967; Hsu, 1992, 1994).

2.2.3 Results

Results of the exploratory factor analyses for each of the sets of scales is shown in Tables 2.3 to 2.6. The MAP and eBIC fit statistics for the IPIP100 (Table 2.3) suggested that 10 factors provided the best fit in Sample 1 and 8 and 9 factors, respectively, in Sample 2. SRMR suggested a good fit when more than 3 or 4 factors were extracted in Samples 1 and 2, respectively. RMSEA did not indicate that any of the solutions were good in either sample. The MAP and eBIC fit statistics for the BFAS (Table 2.4) suggested that 9 and 14 factors provided the best fit in Sample 1 and 7 and 10 factors, respectively, in Sample 2. RMSEA indicate a good fit at 12 factors in Sample 1 and was inconclusive in Sample 2. SRMR suggested that more than 2 factors provided a good fit in Sample 1 and more than 3 factors in Sample 2.

For the IPIP-HEXACO (Table 2.5), MAP and eBIC suggested 19 and 20 factors, respectively, in Sample 1 and 7 or 17 factors in Sample 2. RMSEA did not indicate good fits for any of the solutions; SRMR suggested more than 2 factors in Sample 1 and 3 factors in Sample 2. MAP and eBIC for the QB6 (Table 2.6) suggested 6 and 10 factors in Sample 1 and 5 and 9 factors in Sample 2. RMSEAs were good for 7 or more factors in Sample 1 but did not indicate good fit for

Table 2.3: IPIP100 – fit statistics for extraction of 1 to 20 factors

	Sample 1					Sample 2				
Factors	MAP	RMSEA	SRMR	eBIC	complexity	MAP	RMSEA	SRMR	eBIC	complexity
1	0.0256	0.13	0.13	228,531	1.00	0.0262	0.29	0.13	67,152	1.00
2	0.0189	0.12	0.11	130,304	1.29	0.0214	0.29	0.11	37,297	1.29
3	0.0154	0.12	0.08	72,864	1.48	0.0178	0.29	0.09	15395	1.51
4	0.0109	0.11	0.06	27,545	1.41	0.0139	0.29	0.07	-511	1.45
5	0.0062	0.11	0.04	-4,476	1.33	0.0099	0.29	0.05	-11,708	1.32
6	0.0056	0.11	0.04	-8,512	1.50	0.0094	0.30	0.05	-13,556	1.49
7	0.0055	0.11	0.04	-11,158	1.62	0.0094	0.30	0.05	-14,217	1.65
8	0.0053	0.11	0.03	-12,694	1.75	**0.0093**	0.30	0.05	-14,824	1.82
9	0.0052	0.11	0.03	-13,561	1.81	0.0095	0.30	0.04	**-14,851**	2.00
10	**0.0051**	0.11	0.03	**-14,018**	2.04	0.0097	0.31	0.04	-14,775	2.00
11	0.0052	0.11	0.03	-14,000	2.19	0.0100	0.31	0.04	-14,671	2.10
12	0.0053	0.11	0.03	-13,982	2.31	0.0102	0.31	0.04	-14,500	2.16
13	0.0053	0.11	0.03	-13,953	2.21	0.0105	0.32	0.04	-14,283	2.27
14	0.0055	0.11	0.03	-13,747	2.33	0.0108	0.32	0.04	-14,134	2.32
15	0.0056	0.12	0.03	-13,551	2.39	0.0111	0.32	0.04	-13,932	2.45
16	0.0057	0.12	0.03	-13,265	2.43	0.0115	0.33	0.04	-13,705	2.50
17	0.0059	0.12	0.03	-12,900	2.54	0.0118	0.33	0.04	-13,508	2.62
18	0.0060	0.12	0.03	-12,620	2.61	0.0122	0.34	0.04	-13,262	2.70
19	0.0062	0.12	0.03	-12,352	2.55	0.0125	0.34	0.04	-13,022	2.87
20	0.0063	0.12	0.03	-12,043	2.59	0.0129	0.35	0.04	-12,761	2.81

Fits are based on the "minres" factoring method and "oblimin" rotation procedures using the 'nfactors' function in the *psych* package in R. Minimum values are bolded for eBIC and MAP.

Table 2.4: BFAS – fit statistics for extraction of 1 to 20 factors

	Sample 1					Sample 2				
Factors	MAP	RMSEA	SRMR	eBIC	complexity	MAP	RMSEA	SRMR	eBIC	complexity
1	0.0197	0.09	0.12	180,446	1.00	0.0206	0.27	0.12	51,508	1.00
2	0.0149	0.08	0.10	102,115	1.32	0.0167	0.27	0.10	26,359	1.19
3	0.0105	0.07	0.08	53,827	1.55	0.0132	0.27	0.09	10,166	1.58
4	0.0075	0.07	0.06	14,421	1.72	0.0105	0.27	0.07	-3,687	1.69
5	0.0052	0.06	0.04	-7,749	1.76	0.0086	0.27	0.06	-11,214	1.71
6	0.0046	0.06	0.04	-12,391	2.04	0.0082	0.27	0.05	-13,038	1.96
7	0.0044	0.06	0.03	-14,797	2.04	**0.0081**	0.27	0.05	-13,742	1.98
8	0.0043	0.06	0.03	-16,039	2.12	0.0082	0.28	0.05	-13,868	2.20
9	**0.0042**	0.06	0.03	-17,249	2.11	0.0083	0.28	0.05	-13,967	2.21
10	0.0042	0.06	0.03	-17,575	2.29	0.0084	0.28	0.05	**-13,994**	2.39
11	0.0043	0.06	0.03	-17,635	2.36	0.0086	0.29	0.04	-13,913	2.48
12	0.0043	0.05	0.03	-17,940	2.31	0.0088	0.29	0.04	-13,829	2.47
13	0.0044	0.05	0.02	-18,035	2.34	0.0090	0.29	0.04	-13,678	2.51
14	0.0045	0.05	0.02	**-18,040**	2.43	0.0092	0.29	0.04	-13,502	2.60
15	0.0046	0.05	0.02	-17,944	2.44	0.0094	0.30	0.04	-13,291	2.61
16	0.0047	0.05	0.02	-17,777	2.55	0.0097	0.30	0.04	-13,065	2.76
17	0.0048	0.05	0.02	-17,709	2.52	0.0100	0.31	0.04	-12,831	2.84
18	0.0050	0.05	0.02	-17,491	2.54	0.0102	0.31	0.04	-12,600	3.04
19	0.0051	0.05	0.02	-17,268	2.51	0.0105	0.31	0.04	-12,352	3.18
20	0.0053	0.05	0.02	-16,974	2.78	0.0109	0.32	0.04	-12,099	3.16

Fits are based on the "minres" factoring method and "oblimin" rotation procedures using the 'nfactors' function in the *psych* package in R. Minimum values are bolded for eBIC and MAP.

Table 2.5: IPIP-HEXACO – fit statistics for extraction of 1 to 20 factors

	Sample 1					Sample 2				
Factors	MAP	RMSEA	SRMR	eBIC	complexity	MAP	RMSEA	SRMR	eBIC	complexity
1	0.0174	0.19	0.12	1014,576	1.00	0.0182	0.19	0.12	289,425	1.00
2	0.0129	0.15	0.09	581,281	1.35	0.0145	0.16	0.10	141,362	1.35
3	0.0095	0.16	0.07	286,946	1.76	0.0119	0.13	0.08	53,054	1.71
4	0.0071	0.16	0.06	121,761	1.93	0.0096	0.11	0.07	-14,862	1.80
5	0.0053	0.16	0.05	3,468	1.85	0.0081	0.10	0.06	-56,448	1.87
6	0.0044	0.16	0.04	-42,356	2.06	0.0074	0.09	0.05	-75,630	2.11
7	0.0039	0.16	0.04	-68,909	2.18	**0.0070**	0.05	0.05	-86,391	2.23
8	0.0038	0.16	0.04	-77,669	2.40	0.0070	0.08	0.05	-88,829	2.40
9	0.0037	0.16	0.04	-84,373	2.56	0.0070	0.08	0.05	-91,187	2.53
10	0.0036	0.16	0.03	-89,957	2.64	0.0071	0.08	0.05	-92,714	2.70
11	0.0035	0.16	0.03	-94,774	2.69	0.0071	0.08	0.05	-93,848	2.87
12	0.0035	0.16	0.03	-98,161	2.76	0.0072	0.08	0.05	-94,690	3.00
13	0.0034	0.16	0.03	-100,897	2.73	0.0072	0.08	0.05	-95,467	3.18
14	0.0034	0.16	0.03	-103,675	2.84	0.0073	0.08	0.04	-96,024	3.13
15	0.0034	0.16	0.03	-105,660	2.77	0.0074	0.07	0.04	-96,351	3.30
16	0.0033	0.16	0.03	-107,091	2.90	0.0075	0.07	0.04	-96,551	3.35
17	0.0033	0.17	0.03	-108,496	3.08	0.0076	0.07	0.04	**-96,575**	3.48
18	0.0033	0.17	0.03	-109,368	3.16	0.0077	0.07	0.04	-96,534	3.51
19	**0.0033**	0.17	0.03	-110,152	3.16	0.0079	0.07	0.04	-96,371	3.53
20	0.0033	0.17	0.03	**-110,587**	3.17	0.0080	0.07	0.04	-96,062	3.64

Fits are based on the "minres" factoring method and "oblimin" rotation procedures using the 'nfactors' function in the *psych* package in R. Minimum values are bolded for eBIC and MAP.

any of the solutions in Sample 2. SRMR fits were good for 3 or more factors in both samples.

Mean item complexities are also included for each item set in Tables 2.3 to 2.6 and are depicted graphically in Figure 2.2. As expected, the mean complexities steadily increased for all item sets as more factors were extracted. In general, the trajectory of complexities was similar across samples for all of the item sets, though some divergence was seen at higher levels of extraction, particularly for the QB6. The most consistent pattern across item sets was the tendency for the slope of the complexity lines to be negative or near zero at five factors for both samples.

The five factor solution for the IPIP100 items in Sample 1 is shown in Table 2.7. With only one exception, all of the items had primary loadings on the factor which included items from the same scale (i.e., all of the Conscientiousness items had primary loadings on the same factor) and all of the items had primary loadings above ± 0.4, with two exceptions ("Bottle up my feelings"

Table 2.6: QB6 – fit statistics for extraction of 1 to 20 factors

	Sample 1					Sample 2				
Factors	MAP	RMSEA	SRMR	eBIC	complexity	MAP	RMSEA	SRMR	eBIC	complexity
1	0.0181	0.10	0.12	38,077	1.00	0.0200	0.12	0.12	11,793	1.00
2	0.0148	0.09	0.10	25,763	1.27	0.0165	0.11	0.10	5,732	1.26
3	0.0115	0.08	0.07	10,427	1.67	0.0143	0.11	0.08	1,958	1.49
4	0.0087	0.07	0.05	3,376	1.59	0.0113	0.10	0.06	-853	1.60
5	0.0072	0.06	0.04	-856	1.63	**0.0104**	0.10	0.05	-2,161	1.61
6	**0.0070**	0.06	0.04	-2,002	1.74	0.0106	0.10	0.05	-2,461	1.75
7	0.0072	0.05	0.03	-2,457	1.74	0.0109	0.09	0.05	-2,492	1.87
8	0.0076	0.05	0.03	-2,794	1.84	0.0115	0.09	0.04	-2,459	1.92
9	0.0081	0.05	0.03	-2,885	1.76	0.0123	0.09	0.04	**-2,522**	2.04
10	0.0086	0.05	0.02	**-2,999**	1.87	0.0129	0.10	0.04	-2,465	2.16
11	0.0093	0.05	0.02	-2,927	1.88	0.0137	0.10	0.04	-2,400	2.23
12	0.0101	0.05	0.02	-2,902	1.92	0.0148	0.09	0.04	-2,342	2.35
13	0.0110	0.04	0.02	-2,841	2.03	0.0158	0.09	0.03	-2,247	2.46
14	0.0118	0.04	0.02	-2,757	2.15	0.0169	0.09	0.03	-2,183	2.85
15	0.0129	0.04	0.02	-2,670	2.21	0.0180	0.10	0.03	-2,063	2.84
16	0.0139	0.04	0.02	-2,603	2.24	0.0191	0.10	0.03	-1,961	3.09
17	0.0149	0.04	0.02	-2,483	2.32	0.0202	0.09	0.03	-1,827	3.09
18	0.0162	0.04	0.01	-2,358	2.44	0.0215	0.10	0.03	-1,729	2.94
19	0.0175	0.04	0.01	-2,228	2.58	0.0228	0.10	0.03	-1,626	3.20
20	0.0190	0.03	0.01	-2,088	2.57	0.0242	0.10	0.02	-1,516	3.11

Fits are based on the "minres" factoring method and "oblimin" rotation procedures using the 'nfactors' function in the *psych* package in R. Minimum values are bolded for eBIC and MAP.

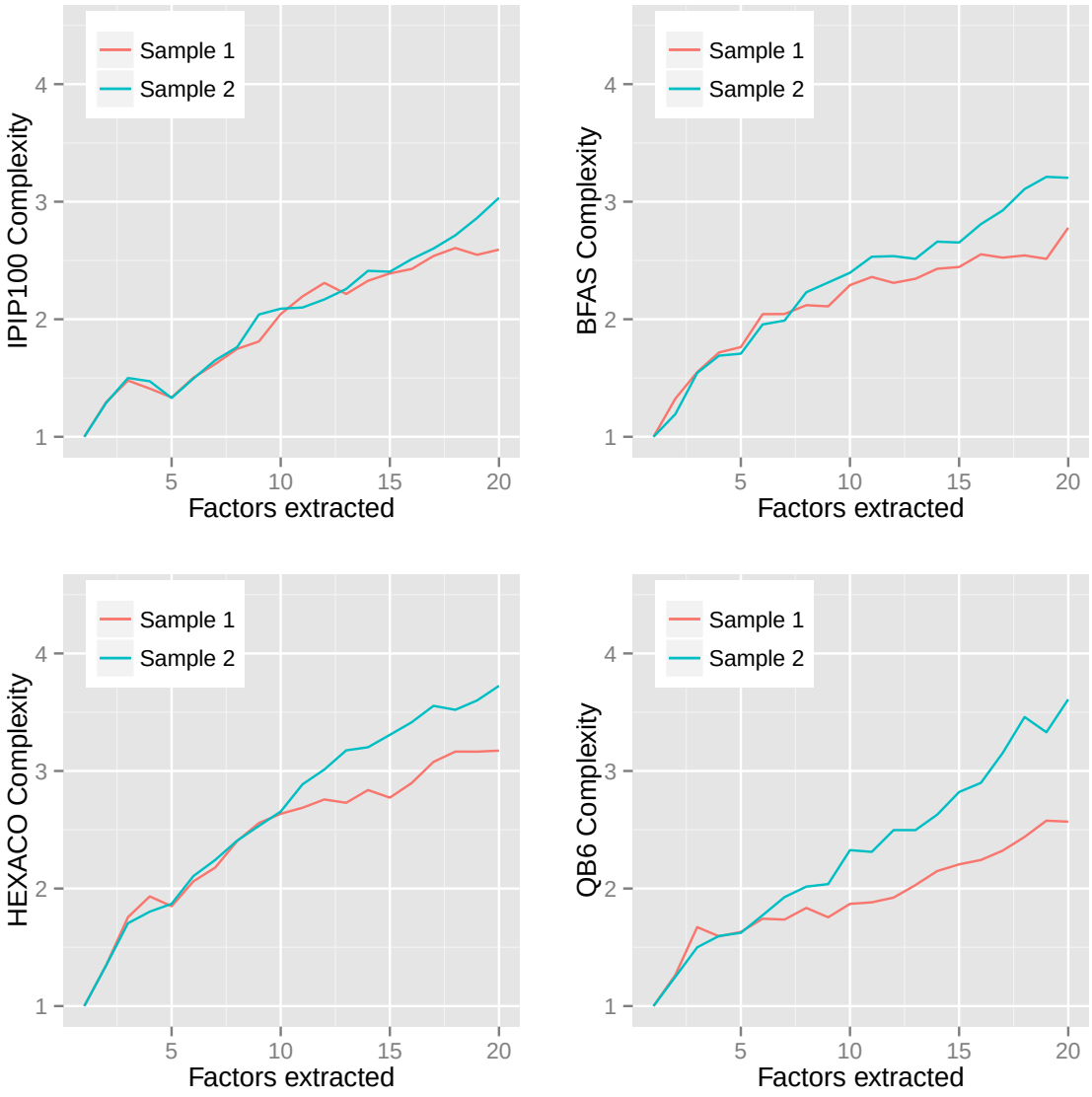


Figure 2.2: Complexities based on extraction of 1 to 20 factors for 373 items in two samples

as an Extraversion item and "Am on good terms with nearly everyone" as an Agreeableness item). The exceptional item was a poorly performing item from the Agreeableness scale ("Am hard to get to know") which loaded high on the Extraversion factor (-0.51) and relatively low on the Agreeableness factor (-0.20). This item performed similarly in the five factor solution for Sample 2. Several items were notable for their complexity (complexities greater than 2.0 are bolded), though 94 of the 100 items had complexities less than 2.0.

Table 2.7: Five factor solution for the IPIP100 in Sample 1

Item	Extraversion	Emotional Stability	Conscientiousness	Agreeableness	Intellect	complexity
Keep in the background. (IPIP100:E)	**-0.80**	-0.04	0.00	0.05	0.02	1.01
Am the life of the party. (IPIP100:E)	**0.77**	0.05	-0.07	-0.01	0.06	1.04
Am quiet around strangers. (IPIP100:E)	**-0.76**	0.02	0.02	0.03	-0.01	1.01
Talk to a lot of different people at parties. (IPIP100:E)	**0.76**	0.01	-0.04	0.07	0.01	1.02
Don't talk a lot. (IPIP100:E)	**-0.75**	-0.15	0.05	-0.08	0.02	1.12
Start conversations. (IPIP100:E)	**0.75**	0.05	0.05	0.12	0.05	1.08
Find it difficult to approach others. (IPIP100:E)	**-0.74**	0.12	-0.08	-0.01	0.04	1.08
Don't mind being the center of attention. (IPIP100:E)	**0.71**	0.07	-0.12	-0.11	0.11	1.18
Feel comfortable around people. (IPIP100:E)	**0.71**	-0.15	0.03	0.20	-0.03	1.27
Feel at ease with people. (IPIP100:E)	**0.69**	-0.18	0.03	0.19	-0.04	1.31
Make friends easily. (IPIP100:E)	**0.68**	-0.04	0.05	0.19	-0.07	1.19
Often feel uncomfortable around others. (IPIP100:E)	**-0.68**	0.22	-0.05	-0.03	0.08	1.26
Don't like to draw attention to myself. (IPIP100:E)	**-0.67**	-0.09	0.17	0.11	-0.01	1.21
Am skilled in handling social situations. (IPIP100:E)	**0.67**	-0.06	0.10	0.18	0.09	1.25
Am a very private person. (IPIP100:E)	**-0.61**	-0.01	0.11	-0.10	0.17	1.30
Have little to say. (IPIP100:E)	**-0.58**	-0.12	0.02	-0.08	-0.23	1.45
Take charge. (IPIP100:E)	**0.55**	0.07	0.26	-0.08	0.21	1.83
Know how to captivate people. (IPIP100:E)	**0.55**	0.03	0.04	0.04	0.29	1.54
Am hard to get to know. (IPIP100:A)	**-0.51**	0.05	-0.04	-0.20	0.23	1.79
Wait for others to lead the way. (IPIP100:E)	**-0.50**	0.06	-0.21	0.18	-0.25	**2.25**
Bottle up my feelings. (IPIP100:E)	**-0.37**	0.13	-0.06	-0.08	0.05	1.46
Get upset easily. (IPIP100:ES)	0.03	**0.79**	0.02	0.03	-0.09	1.03
Get irritated easily. (IPIP100:ES)	0.11	**0.76**	0.06	-0.24	0.01	1.25
Get stressed out easily. (IPIP100:ES)	-0.11	**0.75**	0.02	0.16	-0.04	1.15
Get angry easily. (IPIP100:ES)	0.17	**0.75**	0.09	-0.26	-0.06	1.40
Have frequent mood swings. (IPIP100:ES)	0.01	**0.73**	-0.12	0.03	0.03	1.06
Am not easily bothered by things. (IPIP100:ES)	-0.03	**-0.72**	-0.06	-0.03	0.07	1.04
Rarely get irritated. (IPIP100:ES)	-0.10	**-0.71**	-0.08	0.19	0.01	1.21
Change my mood a lot. (IPIP100:ES)	0.00	**0.67**	-0.18	0.04	0.05	1.16
Take offense easily. (IPIP100:ES)	0.00	**0.65**	0.05	0.04	-0.10	1.07
Panic easily. (IPIP100:ES)	-0.14	**0.65**	-0.05	0.19	-0.11	1.35
Get overwhelmed by emotions. (IPIP100:ES)	-0.05	**0.64**	-0.10	**0.33**	-0.06	1.59
Seldom get mad. (IPIP100:ES)	-0.18	**-0.62**	-0.06	0.19	0.04	1.39
Worry about things. (IPIP100:ES)	-0.19	**0.62**	0.06	0.22	0.02	1.47

continued on next page

Table 2.7 – Five factor solution for the IPIP100 in Sample 1 (continued from previous page)

Item	Extraversion	Emotional Stability	Conscientiousness	Agreeableness	Intellect	complexity
Am relaxed most of the time. (IPIP100:ES)	0.09	**-0.61**	-0.07	-0.01	0.00	1.07
Get caught up in my problems. (IPIP100:ES)	-0.17	**0.59**	-0.14	0.14	0.08	1.46
Am easily disturbed. (IPIP100:ES)	-0.01	**0.57**	0.01	0.07	-0.07	1.06
Feel threatened easily. (IPIP100:ES)	-0.26	**0.54**	-0.08	0.07	-0.03	1.53
Grumble about things. (IPIP100:ES)	-0.01	**0.53**	-0.09	-0.16	0.06	1.27
Often feel blue. (IPIP100:ES)	**-0.34**	**0.51**	-0.22	0.06	0.15	**2.43**
Seldom feel blue. (IPIP100:ES)	0.25	**-0.43**	0.13	-0.08	-0.07	1.95
Love order and regularity. (IPIP100:C)	-0.15	0.20	**0.67**	0.04	-0.09	1.34
Do things according to a plan. (IPIP100:C)	-0.07	0.11	**0.66**	0.04	0.00	1.08
Follow a schedule. (IPIP100:C)	0.00	0.12	**0.65**	0.07	-0.07	1.12
Neglect my duties. (IPIP100:C)	-0.05	0.15	**-0.65**	-0.06	0.04	1.15
Make plans and stick to them. (IPIP100:C)	0.03	0.03	**0.64**	0.04	0.03	1.02
Like order. (IPIP100:C)	-0.08	0.20	**0.64**	0.02	0.00	1.23
Get chores done right away. (IPIP100:C)	0.05	-0.01	**0.64**	0.04	-0.07	1.05
Am always prepared. (IPIP100:C)	0.01	-0.01	**0.64**	0.00	0.13	1.08
Like to tidy up. (IPIP100:C)	0.01	0.12	**0.63**	0.10	-0.06	1.14
Leave a mess in my room. (IPIP100:C)	-0.02	0.08	**-0.62**	0.03	0.12	1.11
Often forget to put things back in their proper place. (IPIP100:C)	0.04	0.09	**-0.62**	0.05	0.06	1.09
Do things in a half-way manner. (IPIP100:C)	-0.04	0.10	**-0.62**	-0.07	-0.12	1.17
Make a mess of things. (IPIP100:C)	-0.01	0.23	**-0.59**	0.04	-0.02	1.31
Shirk my duties. (IPIP100:C)	0.03	0.11	**-0.58**	-0.12	-0.09	1.21
Find it difficult to get down to work. (IPIP100:C)	-0.09	0.15	**-0.57**	0.02	0.01	1.20
Leave my belongings around. (IPIP100:C)	0.04	0.06	**-0.57**	0.12	0.12	1.21
Waste my time. (IPIP100:C)	-0.15	0.16	**-0.55**	-0.01	0.04	1.34
Continue until everything is perfect. (IPIP100:C)	-0.06	0.20	**0.53**	0.04	0.23	1.71
Am exacting in my work. (IPIP100:C)	-0.08	0.04	**0.49**	0.05	0.29	1.74
Pay attention to details. (IPIP100:C)	-0.08	0.07	**0.45**	0.07	0.30	1.92
Sympathize with others' feelings. (IPIP100:A)	0.01	0.07	-0.01	**0.80**	0.01	1.02
Have a soft heart. (IPIP100:A)	-0.03	0.10	0.04	**0.73**	-0.08	1.07
Feel others' emotions. (IPIP100:A)	0.06	0.13	0.01	**0.71**	0.07	1.11
Am indifferent to the feelings of others. (IPIP100:A)	0.02	0.01	0.02	**-0.70**	-0.06	1.02
Love to help others. (IPIP100:A)	0.09	-0.01	0.08	**0.69**	0.01	1.07
Think of others first. (IPIP100:A)	-0.05	-0.06	0.13	**0.68**	-0.06	1.12
Inquire about others' well-being. (IPIP100:A)	0.15	0.07	0.04	**0.66**	0.07	1.17
Take time out for others. (IPIP100:A)	0.11	-0.06	0.07	**0.65**	0.02	1.10
Feel little concern for others. (IPIP100:A)	0.03	0.02	0.01	**-0.63**	-0.04	1.01

continued on next page

Table 2.7 – Five factor solution for the IPIP100 in Sample 1 (continued from previous page)

Item	Extraversion	Emotional Stability	Conscientiousness	Agreeableness	Intellect	complexity
Am not interested in other people's problems. (IPIP100:A)	-0.06	0.01	0.07	**-0.61**	-0.01	1.05
Am not really interested in others. (IPIP100:A)	-0.27	0.00	0.02	**-0.56**	0.05	1.48
Know how to comfort others. (IPIP100:A)	0.30	0.02	0.07	**0.55**	0.06	1.62
Have a good word for everyone. (IPIP100:A)	0.05	-0.18	0.06	**0.52**	0.03	1.29
Am interested in people. (IPIP100:A)	**0.31**	0.00	-0.06	**0.51**	0.08	1.72
Make people feel at ease. (IPIP100:A)	0.28	-0.12	0.05	**0.50**	0.02	1.73
Show my gratitude. (IPIP100:A)	0.13	-0.03	0.18	**0.45**	0.08	1.58
Insult people. (IPIP100:A)	0.21	**0.33**	-0.18	**-0.43**	0.06	**2.84**
Love children. (IPIP100:A)	0.20	-0.02	0.14	**0.41**	-0.10	1.86
Am on good terms with nearly everyone. (IPIP100:A)	0.07	-0.28	0.09	**0.39**	0.02	**2.02**
Have a rich vocabulary. (IPIP100:I)	-0.04	0.01	-0.02	-0.02	**0.67**	1.01
Love to read challenging material. (IPIP100:I)	-0.16	-0.05	-0.04	0.00	**0.66**	1.14
Use difficult words. (IPIP100:I)	-0.06	0.05	-0.06	-0.03	**0.65**	1.05
Am full of ideas. (IPIP100:I)	0.16	0.05	0.03	0.04	**0.62**	1.16
Have difficulty understanding abstract ideas. (IPIP100:I)	0.06	0.18	0.10	0.04	**-0.61**	1.25
Have excellent ideas. (IPIP100:I)	0.20	0.02	0.10	-0.03	**0.60**	1.29
Avoid difficult reading material. (IPIP100:I)	0.14	0.09	-0.03	0.02	**-0.59**	1.16
Am quick to understand things. (IPIP100:I)	0.05	-0.12	0.12	-0.05	**0.59**	1.20
Can handle a lot of information. (IPIP100:I)	0.02	-0.13	0.19	-0.06	**0.58**	1.35
Carry the conversation to a higher level. (IPIP100:I)	0.25	0.05	0.00	0.03	**0.57**	1.39
Have a vivid imagination. (IPIP100:I)	0.05	0.08	-0.17	0.15	**0.54**	1.42
Am not interested in abstract ideas. (IPIP100:I)	0.11	0.10	0.17	-0.07	**-0.52**	1.43
Love to think up new ways of doing things. (IPIP100:I)	0.08	-0.05	0.00	0.04	**0.50**	1.08
Will not probe deeply into a subject. (IPIP100:I)	0.05	0.04	0.01	-0.08	**-0.49**	1.09
Catch on to things quickly. (IPIP100:I)	0.09	-0.10	0.17	-0.02	**0.49**	1.40
Have difficulty imagining things. (IPIP100:I)	-0.04	0.05	0.08	-0.11	**-0.49**	1.20
Am good at many things. (IPIP100:I)	0.26	-0.08	0.22	0.00	**0.46**	**2.19**
Try to avoid complex people. (IPIP100:I)	-0.07	0.11	0.14	-0.05	**-0.45**	1.43
Do not have a good imagination. (IPIP100:I)	-0.04	0.00	0.08	-0.11	**-0.45**	1.22
Spend time reflecting on things. (IPIP100:I)	-0.21	0.11	-0.07	0.27	**0.45**	**2.37**
SS loadings	11.31	9.64	8.32	8.26	7.14	

The ten factor solution for the BFAS items in Sample 1 is shown in Table 2.8. In this case, the factor solution was inconsistent with the prescribed organization of items by scale. The first factor had primary loadings for all of the Compassion items and one Politeness item. The second factor was a mixture of Assertiveness and Politeness items, with one additional Withdrawal item. The third factor was comprised mainly of Volatility items, the fourth Intellect, the fifth Orderliness, the sixth Enthusiasm, the seventh Industriousness with two items from both Withdrawal and Politeness, the eighth Openness, the ninth was a combination of the remaining Volatility and Withdrawal items and the tenth factor was undefined. 43 of the 100 items had complexities greater than 2.0 in this solution.

Table 2.8: Ten factor solution for the BFAS in Sample 1

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	complexity
Can't be bothered with other's needs. (BFAS:A:C)	**-0.73**	0.00	-0.03	-0.02	0.07	0.00	0.08	0.01	0.08	0.12	1.13
Sympathize with others' feelings. (BFAS:A:C)	**0.72**	-0.04	-0.01	-0.02	0.02	-0.05	0.00	0.05	0.07	0.05	1.05
Like to do things for others. (BFAS:A:C)	**0.67**	0.09	-0.10	-0.03	0.05	0.04	-0.01	0.01	0.03	0.12	1.18
Am not interested in other people's problems. (BFAS:A:C)	**-0.67**	-0.04	0.00	0.01	0.06	0.05	-0.06	0.01	0.07	0.11	1.14
Take an interest in other people's lives. (BFAS:A:C)	**0.66**	0.12	-0.02	0.03	0.00	-0.08	0.12	-0.02	0.07	0.03	1.20
Am indifferent to the feelings of others. (BFAS:A:C)	**-0.66**	0.10	0.02	-0.04	0.01	0.09	0.05	-0.07	0.01	0.13	1.22
Inquire about others' well-being. (BFAS:A:C)	**0.66**	0.11	0.00	-0.01	0.08	-0.03	0.07	0.05	-0.03	0.08	1.17
Feel others' emotions. (BFAS:A:C)	**0.64**	0.08	-0.01	-0.04	0.02	0.00	0.00	0.11	0.11	0.08	1.19
Take no time for others. (BFAS:A:C)	**-0.58**	-0.04	0.04	-0.01	0.03	0.09	0.08	0.02	0.02	0.01	1.12
Don't have a soft side. (BFAS:A:C)	**-0.38**	0.11	-0.04	0.03	-0.01	0.17	-0.02	-0.13	-0.03	0.08	**2.01**
Am out for my own personal gain. (BFAS:A:P)	**-0.38**	0.26	0.03	-0.01	0.09	0.00	0.28	-0.06	0.00	0.04	**2.95**
Take charge. (BFAS:E:A)	0.16	**0.64**	0.03	0.09	0.07	0.01	-0.10	-0.01	-0.01	0.12	1.31
See myself as a good leader. (BFAS:E:A)	0.13	**0.60**	-0.07	0.10	0.16	-0.05	-0.01	-0.01	-0.02	0.11	1.43
Wait for others to lead the way. (BFAS:E:A)	0.00	**-0.57**	0.06	-0.09	0.05	0.05	0.30	-0.08	-0.02	0.09	1.73
Do not have an assertive personality. (BFAS:E:A)	-0.04	**-0.56**	-0.06	0.01	0.02	0.13	0.09	-0.03	0.05	0.10	1.30
Am the first to act. (BFAS:E:A)	0.11	**0.54**	-0.04	0.08	0.02	-0.02	-0.11	0.05	0.06	0.14	1.43
Have a strong personality. (BFAS:E:A)	0.07	**0.53**	0.09	0.11	-0.02	-0.12	-0.06	0.08	0.02	0.11	1.48
Lack the talent for influencing people. (BFAS:E:A)	-0.10	**-0.49**	0.09	-0.15	-0.03	0.13	0.04	-0.06	-0.01	0.03	1.60
Can talk others into doing things. (BFAS:E:A)	0.06	**0.48**	-0.03	0.16	0.04	-0.07	0.13	0.09	0.00	0.08	1.65
Know how to captivate people. (BFAS:E:A)	0.01	**0.47**	-0.14	0.11	0.08	-0.23	0.08	0.16	0.12	0.10	**2.53**
Love a good fight. (BFAS:A:P)	-0.09	**0.46**	0.16	-0.07	-0.09	0.09	0.10	-0.01	-0.06	0.07	1.79
Seek conflict. (BFAS:A:P)	-0.09	**0.43**	0.19	0.01	-0.12	0.08	0.23	-0.15	0.01	-0.07	**2.81**
Avoid imposing my will on others. (BFAS:A:P)	0.14	**-0.42**	-0.16	0.02	0.00	0.13	-0.08	0.09	0.05	0.18	**2.43**
Rarely put people under pressure. (BFAS:A:P)	0.05	**-0.41**	-0.19	-0.05	-0.03	0.02	-0.06	0.13	0.00	0.22	**2.38**
Hate to seem pushy. (BFAS:A:P)	0.20	**-0.40**	-0.09	0.01	0.10	0.04	0.11	0.07	0.05	0.18	**2.57**
Hold back my opinions. (BFAS:E:A)	0.02	**-0.38**	-0.15	-0.09	0.06	0.27	0.08	-0.06	0.06	0.14	**3.03**
Am not embarrassed easily. (BFAS:N:W)	-0.11	**0.35**	-0.22	-0.01	-0.16	-0.09	-0.15	0.06	-0.09	0.13	**3.67**
Get angry easily. (BFAS:N:V)	-0.08	0.04	**0.79**	0.01	-0.02	0.03	-0.08	-0.06	0.06	0.09	1.10

continued on next page

Table 2.8 – Ten factor solution for the BFAS in Sample 1 (continued from previous page)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	complexity
Rarely get irritated. (BFAS:N:V)	-0.01	-0.07	**-0.72**	0.05	-0.05	-0.04	0.03	-0.01	0.01	0.13	1.12
Am not easily annoyed. (BFAS:N:V)	0.07	-0.02	**-0.72**	0.03	-0.04	-0.01	0.03	-0.02	0.04	0.09	1.08
Can be stirred up easily. (BFAS:N:V)	0.11	-0.01	**0.71**	0.01	0.01	-0.08	0.08	0.01	0.09	0.06	1.15
Get easily agitated. (BFAS:N:V)	-0.07	0.00	**0.70**	0.00	0.04	0.03	0.02	0.04	0.16	0.08	1.17
Get upset easily. (BFAS:N:V)	0.02	-0.11	**0.62**	-0.04	0.07	-0.06	0.00	0.06	0.25	0.05	1.47
Rarely lose my composure. (BFAS:N:V)	0.04	-0.02	**-0.43**	0.15	0.07	0.18	0.10	-0.01	-0.27	0.17	**3.10**
Insult people. (BFAS:A:P)	-0.15	**0.33**	**0.35**	-0.02	-0.08	0.11	0.30	-0.13	-0.08	0.00	**4.16**
Worry about things. (BFAS:N:W)	0.18	-0.15	0.27	0.01	0.25	0.09	0.18	0.00	0.24	-0.01	**5.49**
Feel threatened easily. (BFAS:N:W)	0.07	-0.17	0.26	-0.02	0.11	0.15	0.23	-0.04	0.21	-0.01	**5.11**
Am quick to understand things. (BFAS:I:I)	0.01	-0.02	-0.03	**0.78**	-0.01	0.02	0.02	-0.01	0.05	0.07	1.03
Can handle a lot of information. (BFAS:I:I)	0.00	-0.01	-0.02	**0.75**	0.02	0.02	-0.05	-0.03	0.05	0.02	1.02
Learn things slowly. (BFAS:I:I)	0.02	0.06	-0.02	**-0.72**	0.06	0.14	0.10	0.07	-0.02	0.12	1.23
Think quickly. (BFAS:I:I)	-0.04	0.17	0.00	**0.61**	0.02	-0.07	-0.02	0.00	0.08	0.14	1.35
Like to solve complex problems. (BFAS:I:I)	0.00	0.00	-0.05	**0.53**	0.02	0.16	0.14	0.08	-0.09	-0.02	1.49
Formulate ideas clearly. (BFAS:I:I)	0.00	0.11	-0.04	**0.49**	0.11	0.02	-0.02	0.13	-0.05	0.06	1.45
Have a rich vocabulary. (BFAS:I:I)	0.03	0.07	0.00	**0.45**	0.04	0.09	0.16	0.24	-0.03	-0.12	**2.20**
Avoid difficult reading material. (BFAS:I:I)	-0.06	0.03	-0.04	**-0.45**	0.04	-0.11	-0.01	-0.17	0.11	0.20	**2.12**
Have difficulty understanding abstract ideas. (BFAS:I:I)	0.04	0.02	0.03	**-0.45**	0.13	-0.01	-0.05	-0.27	0.12	0.20	**2.57**
Get things done quickly. (BFAS:C:I)	0.01	0.14	0.09	0.30	0.20	0.02	-0.26	-0.08	-0.04	0.22	**4.68**
Always know what I am doing. (BFAS:C:I)	-0.06	0.15	0.01	0.22	0.19	0.07	-0.21	-0.02	-0.18	0.15	**5.98**
Like order. (BFAS:C:O)	-0.02	-0.01	0.01	0.01	**0.75**	-0.01	0.06	-0.01	-0.07	0.00	1.04
Want everything to be "just right." (BFAS:C:O)	-0.01	0.00	0.11	0.04	**0.62**	0.03	0.19	0.03	0.03	0.06	1.31
Want every detail taken care of. (BFAS:C:O)	-0.02	0.05	0.08	0.09	**0.61**	0.05	0.04	0.05	0.03	0.05	1.16
Keep things tidy. (BFAS:C:O)	-0.07	0.06	-0.12	-0.04	**0.59**	-0.07	-0.25	0.01	0.07	-0.03	1.58
Am not bothered by disorder. (BFAS:C:O)	0.06	0.04	-0.01	0.06	**-0.52**	0.11	0.07	-0.03	-0.06	0.23	1.65
Follow a schedule. (BFAS:C:O)	0.08	0.00	0.11	0.01	**0.49**	0.01	-0.21	-0.03	-0.12	0.08	1.76
See that rules are observed. (BFAS:C:O)	0.16	-0.07	0.07	-0.01	**0.48**	0.04	-0.01	-0.11	-0.10	0.12	1.68
Am not bothered by messy people. (BFAS:C:O)	0.08	-0.07	-0.02	0.09	**-0.48**	0.07	0.12	-0.01	-0.03	0.20	1.75
Dislike routine. (BFAS:C:O)	-0.11	0.09	-0.05	0.04	**-0.45**	0.00	0.05	0.16	0.10	0.10	1.81
Leave my belongings around. (BFAS:C:O)	0.16	-0.07	0.21	0.09	**-0.39**	0.05	**0.31**	0.00	-0.08	0.16	**3.74**
Respect authority. (BFAS:A:P)	0.25	-0.16	0.02	-0.10	**0.32**	-0.07	-0.18	-0.14	-0.01	0.19	**4.80**
Am hard to get to know. (BFAS:E:E)	-0.01	0.06	-0.03	0.03	-0.02	**0.72**	0.05	0.06	0.08	0.02	1.08
Keep others at a distance. (BFAS:E:E)	-0.06	0.02	0.09	-0.01	-0.02	**0.71**	0.05	0.06	0.04	0.09	1.12
Reveal little about myself. (BFAS:E:E)	-0.06	0.00	-0.14	0.01	0.03	**0.64**	-0.04	0.02	0.06	0.13	1.24
Warm up quickly to others. (BFAS:E:E)	0.18	0.05	-0.09	0.10	0.00	**-0.55**	0.11	-0.04	0.08	0.19	1.79

continued on next page

Table 2.8 – Ten factor solution for the BFAS in Sample 1 (continued from previous page)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	complexity
Make friends easily. (BFAS:E:E)	0.08	0.25	-0.04	-0.01	0.03	**-0.48**	0.01	0.03	-0.01	0.25	**2.17**
Show my feelings when I'm happy. (BFAS:E:E)	0.20	0.06	0.08	-0.04	0.08	**-0.41**	0.01	0.13	0.04	0.23	**2.71**
Rarely get caught up in the excitement. (BFAS:E:E)	-0.13	-0.08	-0.12	0.04	0.03	**0.41**	-0.09	-0.06	-0.06	-0.06	1.80
Am not a very enthusiastic person. (BFAS:E:E)	-0.19	-0.21	0.03	-0.01	-0.11	**0.39**	0.00	-0.12	0.04	-0.04	**2.62**
Laugh a lot. (BFAS:E:E)	0.12	0.08	0.06	0.03	0.02	**-0.37**	0.13	0.12	-0.12	0.30	**3.14**
Waste my time. (BFAS:C:I)	0.00	-0.09	0.02	-0.03	-0.11	0.06	**0.63**	-0.01	0.01	-0.05	1.15
Find it difficult to get down to work. (BFAS:C:I)	-0.04	-0.06	-0.06	-0.09	-0.14	0.01	**0.54**	0.03	0.15	-0.01	1.43
Postpone decisions. (BFAS:C:I)	0.00	-0.21	-0.04	-0.02	-0.08	0.04	**0.50**	0.01	0.15	0.07	1.69
Take advantage of others. (BFAS:A:P)	-0.30	**0.33**	0.06	-0.04	0.03	0.06	**0.40**	-0.06	-0.01	-0.05	**3.05**
Finish what I start. (BFAS:C:I)	0.14	0.16	0.02	0.10	0.23	0.06	**-0.39**	-0.10	-0.10	0.14	**3.38**
Mess things up. (BFAS:C:I)	0.02	-0.04	0.10	-0.15	-0.18	0.04	**0.37**	0.00	0.18	0.02	**2.65**
Am filled with doubts about things. (BFAS:N:W)	0.03	-0.14	0.15	-0.03	0.13	0.18	**0.36**	0.05	0.26	-0.04	**3.68**
Don't put my mind on the task at hand. (BFAS:C:I)	-0.08	0.01	-0.05	-0.22	-0.22	0.00	**0.36**	0.02	0.16	0.04	**3.15**
Am easily distracted. (BFAS:C:I)	0.04	0.01	0.08	-0.17	-0.17	-0.03	**0.35**	0.09	0.20	0.17	**3.69**
Believe that I am better than others. (BFAS:A:P)	-0.29	0.28	-0.03	0.20	0.09	-0.01	**0.32**	-0.03	0.04	-0.06	**4.02**
Am easily discouraged. (BFAS:N:W)	0.02	-0.30	0.12	-0.16	0.08	0.04	**0.31**	0.02	0.27	0.03	**4.05**
Carry out my plans. (BFAS:C:I)	0.13	0.24	0.05	0.12	0.23	0.10	-0.30	0.01	-0.15	0.21	**5.62**
Am afraid of many things. (BFAS:N:W)	0.05	-0.22	0.22	-0.17	0.16	0.07	0.22	0.03	0.18	0.01	**6.02**
Believe in the importance of art. (BFAS:I:O)	0.03	-0.06	0.08	0.02	-0.02	0.05	0.00	**0.70**	-0.10	0.00	1.10
Need a creative outlet. (BFAS:I:O)	-0.05	0.03	0.06	-0.04	-0.03	0.04	0.02	**0.59**	0.05	0.06	1.10
See beauty in things that others might not notice. (BFAS:I:O)	0.00	0.05	-0.09	0.01	0.05	0.02	-0.09	**0.57**	0.11	0.15	1.36
Seldom notice the emotional aspects of paintings and pictures. (BFAS:I:O)	-0.14	-0.04	0.00	0.03	0.00	0.04	0.08	**-0.53**	0.03	0.22	1.58
Do not like poetry. (BFAS:I:O)	-0.08	-0.07	0.03	0.02	0.02	-0.02	0.05	**-0.52**	-0.04	0.07	1.18
Get deeply immersed in music. (BFAS:I:O)	-0.09	0.01	0.03	0.01	-0.01	-0.04	0.05	**0.43**	0.11	0.10	1.40
Enjoy the beauty of nature. (BFAS:I:O)	0.05	-0.07	0.01	0.01	0.05	-0.02	-0.08	**0.42**	-0.07	0.18	1.67
Love to reflect on things. (BFAS:I:O)	0.06	-0.05	-0.05	0.18	0.07	0.07	0.09	**0.38**	0.01	0.01	1.89
Avoid philosophical discussions. (BFAS:I:I)	-0.01	-0.13	0.03	-0.18	0.06	-0.01	-0.14	**-0.33**	0.10	0.23	**3.74**
Seldom daydream. (BFAS:I:O)	-0.04	0.08	-0.13	-0.09	0.08	-0.02	-0.20	-0.25	0.07	0.16	**4.48**
Change my mood a lot. (BFAS:N:V)	-0.03	0.06	0.12	0.04	-0.05	-0.01	0.04	-0.02	**0.78**	0.05	1.10
Am a person whose moods go up and down easily. (BFAS:N:V)	0.02	0.02	0.19	0.00	-0.03	0.08	-0.04	0.00	**0.77**	0.01	1.16

continued on next page

Table 2.8 – Ten factor solution for the BFAS in Sample 1 (continued from previous page)

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	complexity
Seldom feel blue. (BFAS:N:W)	-0.10	0.00	0.00	0.03	-0.05	-0.18	-0.06	-0.05	**-0.44**	**0.34**	**2.50**
Rarely feel depressed. (BFAS:N:W)	-0.05	0.01	-0.04	0.03	-0.05	-0.21	-0.12	-0.12	**-0.44**	**0.33**	**2.80**
Keep my emotions under control. (BFAS:N:V)	-0.09	0.02	**-0.34**	0.10	0.06	0.20	0.02	-0.01	**-0.34**	0.16	**3.47**
Become overwhelmed by events. (BFAS:N:W)	0.03	-0.22	0.17	-0.14	0.15	0.00	0.16	0.09	**0.32**	0.03	**4.39**
Have a lot of fun. (BFAS:E:E)	0.03	0.15	-0.03	0.01	0.00	**-0.37**	0.02	0.09	-0.10	**0.42**	**2.52**
Feel comfortable with myself. (BFAS:N:W)	-0.07	0.21	-0.16	0.06	-0.01	-0.15	-0.16	0.07	-0.25	0.29	**5.10**
Seldom get lost in thought. (BFAS:I:O)	-0.09	0.08	-0.02	-0.12	0.06	-0.09	-0.22	-0.19	0.04	0.27	**4.36**
SS loadings	5.72	5.91	5.11	4.59	4.13	4.19	4.27	3.29	3.93	2.25	

The six factor solution for the IPIP-HEXACO items in Sample 1 is shown in Table 2.9. The six HEXACO factors were recognizable in this solution as the majority of items in each factor with primary loadings above ± 0.4 were from the same scale. Only eight items had primary loadings above ± 0.4 on factors which differed from their prescribed scale. These included: "Like to attract attention", an Honesty/Humility item with a primary loading on Extraversion (0.57); "Get upset easily", an Agreeableness item with a primary loading on Emotionality (0.49) and Agreeableness (-0.49); "React strongly to criticism", an Agreeableness item with a primary loading on Emotionality (0.41); "Try to follow the rules", an Honesty/Humility item with a primary loading on Conscientiousness (0.48); "Like to be viewed as proper and conventional", an Openness item with a primary loading on Conscientiousness (0.42); and three similar items relating to "seeing oneself as average" which were prescribed as Honesty/Humility items but had primary loadings on Openness (-0.44, -0.47, and -0.48). Each factor also had several items with secondary loadings above ± 0.3 or primary loadings below ± 0.2. Of the 240 items, 111 had complexities greater than 2.0 in this solution.

Table 2.9: Six factor solution for the IPIP-HEXACO in Sample 1

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Don't talk a lot. (H:X:E)	**-0.77**	-0.09	0.18	0.03	-0.02	0.08	1.16
Seem to derive less enjoyment from interacting with people than others do. (H:X:S)	**-0.77**	-0.07	-0.10	0.00	0.15	0.07	1.14
Say little. (H:X:E)	**-0.76**	-0.04	0.19	0.04	-0.09	0.09	1.20
Keep in the background. (H:X:SB)	**-0.75**	0.11	0.15	0.00	-0.03	0.06	1.15
Feel comfortable around people. (H:X:SB)	**0.74**	-0.06	0.16	0.08	-0.02	-0.02	1.13
Talk a lot. (H:X:E)	**0.71**	0.19	-0.15	-0.04	0.02	0.09	1.26
Talk to a lot of different people at parties. (H:X:S)	**0.71**	-0.03	0.06	0.01	0.04	0.12	1.08
Am the life of the party. (H:X:E)	**0.70**	-0.05	0.00	-0.03	0.07	0.16	1.15
Make friends easily. (H:X:S)	**0.69**	0.00	0.12	0.08	-0.04	0.03	1.10
Usually like to spend my free time with people. (H:X:S)	**0.68**	0.10	0.12	0.00	-0.16	0.08	1.25
Find it difficult to approach others. (H:X:SB)	**-0.68**	0.20	0.00	-0.07	0.01	0.11	1.26
Love to chat. (H:X:S)	**0.67**	0.21	0.03	0.04	-0.04	0.10	1.26
Hate being the center of attention. (H:X:SB)	**-0.64**	0.05	0.06	0.08	-0.09	-0.21	1.33
Don't mind being the center of attention. (H:X:SB)	**0.64**	-0.06	-0.03	-0.04	0.13	0.28	1.49
Keep others at a distance. (H:X:S)	**-0.61**	-0.01	-0.10	0.02	0.13	0.18	1.33
Don't like to draw attention to myself. (H:X:E)	**-0.60**	0.05	0.10	0.13	-0.07	-0.15	1.33
Would not enjoy a job that involves a lot of social interaction. (H:X:S)	**-0.60**	-0.03	-0.05	-0.01	0.05	0.04	1.04
Have little to say. (H:X:SB)	**-0.59**	-0.05	0.19	-0.03	-0.28	0.15	1.81
Rarely enjoy being with people. (H:X:S)	**-0.59**	-0.09	-0.16	-0.01	0.04	0.07	1.24
Am hard to get to know. (H:X:S)	**-0.59**	-0.06	-0.05	-0.01	0.20	0.17	1.44
Enjoy being part of a group. (H:X:S)	**0.58**	0.14	0.22	0.02	-0.11	0.19	1.76
Like to attract attention. (H:H:M)	**0.57**	0.09	-0.05	-0.12	0.10	**0.39**	**2.04**
Have a lot of fun. (H:X:L)	**0.57**	-0.02	0.22	0.07	0.01	0.09	1.41
Am usually active and full of energy. (H:X:L)	**0.56**	-0.10	0.18	0.22	0.01	0.11	1.71
Speak softly. (H:X:E)	**-0.55**	0.11	**0.31**	0.07	0.00	0.07	1.74
Smile a lot. (H:X:L)	**0.54**	0.15	0.28	0.09	-0.01	0.03	1.78
Have a strong personality. (H:X:SB)	**0.50**	-0.10	-0.19	0.15	0.29	0.02	**2.29**
Laugh a lot. (H:X:L)	**0.50**	0.13	0.18	0.03	0.05	0.06	1.50
Maintain high energy throughout the day. (H:X:L)	**0.47**	-0.20	0.19	0.26	-0.02	0.10	**2.49**
Have leadership abilities. (H:X:SB)	**0.46**	-0.18	-0.01	**0.31**	0.22	0.07	**2.72**
Bottle up my feelings. (H:X:E)	**-0.45**	0.06	0.03	0.01	-0.02	0.27	1.71
Feel healthy and vibrant most of the time. (H:X:L)	**0.41**	-0.22	0.25	0.26	-0.04	0.03	**3.12**

continued on next page

Table 2.9 – Six factor solution for the IPIP-HEXACO in Sample 1 (continued from previous page)

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Can't do without the company of others. (H:E:D)	**0.40**	0.26	0.01	-0.08	-0.10	0.15	**2.34**
Would be afraid to give a speech in public. (H:X:SB)	**-0.40**	0.27	0.04	-0.02	-0.22	-0.01	**2.39**
Tell people about it when I'm irritated. (H:X:E)	**0.40**	0.13	-0.25	0.05	0.05	-0.03	**2.04**
Am never at a loss for words. (H:X:E)	**0.38**	-0.19	-0.11	0.09	0.24	0.03	**2.58**
Would not enjoy being a famous celebrity. (H:H:GA)	**-0.38**	0.02	0.10	0.04	0.06	-0.18	1.67
Am good at making impromptu speeches. (H:X:SB)	**0.37**	-0.14	0.02	0.01	**0.33**	0.17	**2.75**
Have an intense, boisterous laugh. (H:X:E)	**0.36**	0.12	0.00	-0.07	0.15	0.13	**2.03**
Am willing to take risks. (H:E:F)	**0.34**	-0.20	0.09	-0.03	0.27	0.16	**3.28**
Do not like concerts. (H:O:A)	**-0.32**	-0.14	-0.02	0.03	-0.08	0.11	1.84
Feel that I have a lot of inner strength. (H:X:L)	0.29	-0.23	0.14	0.29	0.26	0.00	**4.35**
Don't strive for elegance in my appearance. (H:H:GA)	-0.29	-0.05	0.05	-0.22	0.00	-0.10	**2.27**
Have trouble guessing how others will react. (H:O:C)	-0.27	0.12	-0.06	-0.10	-0.19	0.12	**3.13**
Panic easily. (H:E:A)	-0.12	**0.63**	-0.26	-0.04	-0.09	0.04	1.48
Need reassurance. (H:E:D)	-0.01	**0.60**	-0.05	-0.04	-0.05	0.21	1.29
Immediately feel sad when hearing of an unhappy event. (H:E:S)	0.20	**0.60**	0.05	0.10	-0.03	-0.18	1.50
Get stressed out easily. (H:E:A)	-0.12	**0.58**	**-0.35**	0.02	-0.02	0.05	1.76
Worry about things. (H:E:A)	-0.16	**0.57**	-0.26	0.11	0.01	0.05	1.66
Need the approval of others. (H:E:D)	0.02	**0.57**	0.03	-0.04	-0.08	**0.33**	1.69
Rarely worry. (H:E:A)	0.12	**-0.56**	0.28	-0.06	0.01	0.07	1.64
Begin to panic when there is danger. (H:E:F)	-0.05	**0.55**	-0.16	0.02	-0.28	0.04	1.70
Seldom get emotional. (H:E:S)	-0.21	**-0.55**	0.26	0.06	-0.10	0.24	**2.34**
Get upset by unpleasant thoughts that come into my mind. (H:E:A)	-0.08	**0.55**	-0.16	-0.03	0.05	0.08	1.28
Need protection. (H:E:D)	-0.08	**0.54**	-0.03	0.04	-0.09	0.14	1.27
Often worry about things that turn out to be unimportant. (H:E:A)	-0.17	**0.53**	-0.17	0.04	0.02	0.18	1.73
Am deeply moved by others' misfortunes. (H:E:S)	0.15	**0.52**	0.16	0.06	0.19	-0.20	**2.05**
Don't understand people who get emotional. (H:E:S)	-0.27	**-0.51**	-0.02	0.05	-0.10	**0.34**	**2.47**
Cry during movies. (H:E:S)	0.19	**0.51**	-0.05	0.02	0.06	-0.19	1.63
Tremble in dangerous situations. (H:E:F)	-0.08	**0.50**	-0.07	0.00	-0.17	0.02	1.32
Get upset easily. (H:A:P)	0.01	**0.49**	**-0.49**	-0.02	-0.06	0.04	**2.05**
Am sensitive to the needs of others. (H:E:S)	0.23	**0.49**	0.30	0.15	0.14	-0.15	**2.85**
Feel others' emotions. (H:E:S)	0.27	**0.49**	0.19	0.10	0.17	-0.15	**2.61**
Want to be liked. (H:E:D)	0.16	**0.48**	0.14	0.01	-0.07	**0.32**	**2.27**
Am not easily disturbed by events. (H:E:A)	-0.08	**-0.48**	0.27	0.03	0.05	0.18	**2.02**

continued on next page

Table 2.9 – Six factor solution for the IPIP-HEXACO in Sample 1 (continued from previous page)

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Show my sadness. (H:E:D)	0.20	**0.48**	-0.15	0.00	0.02	-0.07	1.61
Rarely cry during sad movies. (H:E:S)	-0.20	**-0.44**	0.08	0.00	-0.01	0.25	**2.13**
Remain calm under pressure. (H:E:A)	0.09	**-0.44**	0.27	0.17	0.18	0.08	**2.61**
Suspect that my facial expressions give me away when I feel sad. (H:E:D)	0.15	**0.44**	-0.10	0.03	-0.02	-0.08	1.42
Seek support. (H:E:D)	**0.32**	**0.43**	0.02	0.10	-0.06	-0.04	**2.05**
Often need help. (H:E:D)	0.02	**0.43**	0.00	-0.19	-0.20	0.17	**2.21**
Don't worry about things that have already happened. (H:E:A)	0.12	**-0.42**	**0.31**	0.01	-0.01	0.06	**2.07**
React strongly to criticism. (H:A:FL)	0.00	**0.41**	-0.28	0.05	-0.03	0.17	**2.24**
Often feel blue. (H:X:L)	**-0.40**	**0.40**	-0.20	-0.20	0.19	0.08	**3.52**
Face danger confidently. (H:E:F)	0.18	**-0.40**	0.14	0.09	**0.31**	0.17	**3.24**
Let myself be influenced by others. (H:E:D)	-0.04	**0.40**	0.16	-0.12	-0.15	**0.33**	**2.84**
Am seldom bothered ... suffering of strangers. (H:E:S)	-0.15	**-0.38**	-0.07	-0.01	-0.21	**0.33**	**2.97**
Tire out quickly. (H:X:L)	-0.30	**0.37**	-0.12	-0.21	-0.06	0.07	**2.96**
Would feel very badly for a long time if I were to steal from someone. (H:H:F)	0.03	**0.36**	0.15	0.25	-0.02	-0.27	**3.12**
Rarely feel depressed. (H:E:A)	**0.31**	**-0.35**	0.25	0.17	-0.19	0.00	**3.92**
Let people push me around to help them feel important. (H:H:S)	-0.21	**0.35**	**0.31**	-0.11	-0.06	0.29	**3.92**
Would fear walking in a high-crime part of a city. (H:E:F)	-0.10	**0.34**	-0.08	0.15	-0.18	-0.01	**2.27**
Seldom feel weepy while reading the sad part of a story. (H:E:S)	-0.12	**-0.34**	0.06	0.01	-0.11	0.21	**2.33**
Am a physical coward. (H:E:F)	-0.30	**0.33**	0.00	-0.16	-0.08	0.12	**2.87**
Don't know why I do some of the things I do. (H:C:PR)	-0.10	**0.31**	-0.06	-0.25	-0.07	0.27	**3.38**
Do not enjoy watching dance performances. (H:O:A)	-0.20	-0.24	-0.08	-0.05	-0.09	0.14	**3.36**
Would be good at rescuing people from a burning building. (H:E:F)	0.14	-0.23	0.15	0.10	0.19	0.15	**4.73**
Am hard to convince. (H:A:FL)	-0.15	-0.22	-0.19	0.16	0.18	0.09	**5.03**
Find that it takes a lot to make me feel angry at someone. (H:A:P)	-0.07	-0.09	**0.72**	0.01	0.08	0.08	1.10
Rarely feel angry with people. (H:A:P)	-0.04	-0.07	**0.70**	-0.01	-0.03	0.05	1.04
Rarely get irritated. (H:A:P)	-0.02	-0.21	**0.70**	-0.04	0.01	0.09	1.22
Get angry easily. (H:A:P)	0.06	0.17	**-0.68**	0.02	-0.10	0.08	1.21
Am easily annoyed. (H:A:P)	-0.07	0.17	**-0.65**	0.05	-0.01	0.13	1.27

continued on next page

Table 2.9 – Six factor solution for the IPIP-HEXACO in Sample 1 (continued from previous page)

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Get irritated easily. (H:A:P)	0.00	0.23	**-0.64**	0.02	-0.05	0.13	1.36
Seldom get mad. (H:A:P)	-0.11	-0.16	**0.63**	0.02	0.02	0.09	1.24
Am inclined to forgive others. (H:A:FO)	0.13	0.27	**0.60**	-0.02	0.05	-0.01	1.50
Lose my temper. (H:A:P)	0.11	0.24	**-0.60**	-0.03	0.00	0.07	1.43
Am usually a patient person. (H:A:P)	-0.16	-0.01	**0.60**	0.12	0.07	0.01	1.26
Try to forgive and forget. (H:A:FO)	0.11	0.18	**0.59**	0.04	0.03	0.02	1.28
Find it hard to forgive others. (H:A:FO)	-0.12	-0.06	**-0.56**	0.10	-0.01	0.14	1.33
Hold a grudge. (H:A:FO)	-0.07	0.07	**-0.53**	0.04	-0.03	0.18	1.33
Have a good word for everyone. (H:A:G)	0.19	0.24	**0.50**	0.13	0.08	0.04	**2.01**
Rarely complain. (H:A:G)	-0.11	-0.24	**0.49**	0.12	0.01	-0.03	1.72
Am nice to people I should be angry at. (H:A:FO)	-0.05	**0.31**	**0.48**	0.06	0.00	0.14	1.98
Speak ill of others. (H:A:G)	0.02	0.04	**-0.46**	-0.11	-0.01	0.25	1.74
Love my enemies. (H:A:FO)	0.06	0.08	**0.45**	-0.04	0.13	0.04	1.30
Become frustrated and angry with people when they don't live up to my expectations. (H:A:G)	-0.01	0.12	**-0.45**	0.18	0.07	0.26	**2.26**
Take things as they come. (H:A:G)	0.06	-0.16	**0.44**	-0.08	0.03	0.14	1.66
Am annoyed by others' mistakes. (H:A:FL)	-0.12	-0.01	**-0.44**	0.21	0.09	**0.31**	**2.57**
Accept people as they are. (H:A:G)	0.07	0.13	**0.43**	0.01	0.02	-0.11	1.40
Criticize others' shortcomings. (H:A:G)	-0.01	-0.05	**-0.43**	0.02	0.08	**0.31**	1.96
Get back at people who insult me. (H:A:FO)	0.08	-0.15	**-0.42**	-0.04	-0.01	**0.35**	**2.31**
Am quick to judge others. (H:A:G)	-0.04	0.01	**-0.41**	0.06	0.03	**0.39**	**2.07**
Am hard to satisfy. (H:A:FL)	-0.18	-0.03	**-0.39**	0.04	0.18	0.24	**2.69**
When interacting with a group of people, am often bothered by at least one of them. (H:A:FL)	-0.23	0.03	**-0.39**	0.04	0.08	0.24	**2.48**
Get even with others. (H:A:FO)	0.07	-0.17	**-0.38**	-0.03	-0.06	**0.36**	**2.55**
Am hard to reason with. (H:A:FL)	-0.01	0.01	**-0.37**	-0.05	-0.08	0.23	1.80
Adjust easily. (H:A:FL)	0.30	-0.23	**0.37**	0.07	0.10	0.09	**3.12**
Feel that most people can't be trusted. (H:A:FO)	-0.29	-0.07	**-0.36**	0.09	0.02	0.25	**3.02**
Have a sharp tongue. (H:A:G)	0.16	-0.11	**-0.35**	0.01	0.25	0.18	**3.11**
Am good at taking advice. (H:A:FL)	0.18	0.05	**0.35**	0.24	0.00	-0.03	**2.44**
Distrust people. (H:A:FO)	**-0.33**	-0.06	**-0.34**	0.04	0.08	0.22	**2.94**
Prefer to just let things happen. (H:C:PE)	-0.07	0.01	**0.33**	**-0.31**	-0.10	0.21	**2.98**
Get upset if others change the way that I have arranged things. (H:A:FL)	-0.14	0.17	**-0.31**	0.29	0.06	0.21	**3.93**
Can't stand being contradicted. (H:A:FL)	0.03	0.13	-0.30	0.14	-0.08	0.28	**3.00**
Do things according to a plan. (H:C:PR)	-0.03	0.05	-0.01	**0.64**	-0.14	0.06	1.14

continued on next page

Table 2.9 – Six factor solution for the IPIP-HEXACO in Sample 1 (continued from previous page)

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Want every detail taken care of. (H:C:PE)	-0.10	0.18	-0.09	**0.63**	0.06	0.21	1.52
Continue until everything is perfect. (H:C:PE)	-0.09	0.13	-0.01	**0.62**	0.14	0.13	1.32
Like order. (H:C:O)	-0.07	0.12	-0.05	**0.62**	-0.11	0.08	1.22
Keep things tidy. (H:C:O)	0.04	-0.04	-0.01	**0.62**	-0.14	-0.03	1.13
Make plans and stick to them. (H:C:PR)	0.09	-0.04	-0.01	**0.59**	-0.06	-0.06	1.10
Am exacting in my work. (H:C:D)	-0.07	0.02	0.04	**0.59**	0.19	0.03	1.25
Push myself very hard to succeed. (H:C:D)	0.22	0.05	0.00	**0.59**	0.05	0.00	1.31
Work hard. (H:C:D)	0.17	0.01	0.02	**0.59**	0.06	-0.10	1.25
Complete tasks successfully. (H:C:D)	0.13	-0.11	0.01	**0.57**	0.13	-0.03	1.30
Pay attention to details. (H:C:PE)	-0.10	0.04	0.04	**0.55**	0.21	0.04	1.40
Don't finish the things that I start. (H:C:O)	-0.21	0.20	0.03	**-0.54**	0.14	0.19	**2.10**
Like to tidy up. (H:C:O)	0.06	0.10	0.03	**0.54**	-0.09	-0.01	1.16
Get started quickly on doing a job. (H:C:D)	0.15	-0.05	0.02	**0.53**	0.00	-0.02	1.17
Get chores done right away. (H:C:O)	0.09	-0.05	0.04	**0.53**	-0.10	-0.08	1.22
Want everything to be "just right." (H:C:O)	-0.10	0.28	-0.09	**0.51**	-0.04	0.27	**2.34**
Do too little work. (H:C:D)	-0.20	0.08	0.02	**-0.51**	0.01	0.26	1.89
Want everything to add up perfectly. (H:C:PE)	-0.16	0.19	-0.08	**0.50**	-0.02	0.30	**2.30**
Make careless mistakes. (H:C:PR)	0.00	0.20	0.03	**-0.50**	-0.05	0.30	**2.04**
Do just enough work to get by. (H:C:D)	-0.13	0.01	0.03	**-0.49**	-0.09	**0.33**	**2.01**
Dislike imperfect work. (H:C:PE)	-0.17	0.11	-0.08	**0.49**	0.16	0.19	**2.01**
Pay too little attention to details. (H:C:PE)	0.09	0.01	0.04	**-0.48**	-0.18	0.11	1.47
Often forget to put things back in their proper place. (H:C:O)	0.00	0.17	0.06	**-0.48**	0.11	0.17	1.68
Try to follow the rules. (H:H:F)	-0.04	0.29	0.15	**0.48**	-0.29	-0.07	**2.70**
Demand quality. (H:C:PE)	-0.06	-0.02	-0.12	**0.48**	0.27	0.17	**2.05**
Leave a mess in my room. (H:C:O)	-0.05	0.16	0.02	**-0.47**	0.15	0.16	1.77
Have an eye for detail. (H:C:PE)	-0.09	0.08	0.06	**0.47**	**0.32**	0.09	**2.05**
Avoid mistakes. (H:C:PR)	-0.17	-0.01	0.04	**0.46**	0.09	0.07	1.43
Quickly lose interest in the tasks I start. (H:C:D)	-0.13	0.20	-0.04	**-0.45**	0.00	0.28	**2.35**
Leave my belongings around. (H:C:O)	0.03	0.20	0.08	**-0.44**	0.17	0.15	**2.13**
Like to be viewed as proper and conventional. (H:O:U)	0.01	0.08	0.14	**0.42**	**-0.37**	0.25	**2.96**
Am not bothered by disorder. (H:C:O)	0.05	-0.11	0.21	**-0.41**	0.10	0.09	1.96
Do things without thinking of the consequences. (H:C:PR)	0.19	0.05	-0.03	**-0.41**	-0.07	**0.31**	**2.46**
Stop when work becomes too difficult. (H:C:D)	-0.18	0.27	-0.02	**-0.38**	-0.11	0.23	**3.22**
Hang around doing nothing. (H:C:D)	-0.21	0.17	0.06	**-0.36**	0.00	0.25	**3.13**

continued on next page

Table 2.9 – Six factor solution for the IPIP-HEXACO in Sample 1 (continued from previous page)

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Jump into things without thinking. (H:C:PR)	0.28	0.17	-0.01	**-0.35**	-0.04	0.27	**3.36**
Make a fool of myself. (H:C:PR)	0.05	0.26	0.11	**-0.35**	0.06	0.27	**3.14**
Make rash decisions. (H:C:PR)	0.18	0.15	-0.09	**-0.34**	0.02	**0.31**	**3.18**
Am more capable than most others. (H:H:M)	0.11	-0.21	-0.06	**0.33**	**0.32**	0.23	**3.83**
Have great stamina. (H:X:L)	0.30	-0.21	0.13	0.30	0.09	0.06	**3.49**
Would never take things that aren't mine. (H:H:F)	-0.04	0.16	0.18	0.29	-0.02	-0.27	**3.34**
Would never cheat on my taxes. (H:H:F)	-0.01	0.19	0.11	0.26	-0.10	-0.18	**3.52**
Return extra change when a cashier makes a mistake. (H:H:F)	0.03	0.16	0.19	0.22	0.15	-0.19	**4.71**
Love to read challenging material. (H:O:I)	-0.18	0.03	0.10	0.08	**0.60**	0.00	1.28
Am full of ideas. (H:O:C)	0.13	0.03	0.07	0.16	**0.58**	0.09	1.36
Have a vivid imagination. (H:O:C)	0.06	0.16	0.07	-0.05	**0.57**	0.02	1.22
Have a rich vocabulary. (H:O:I)	-0.08	-0.01	0.02	0.09	**0.56**	0.04	1.11
Believe in the importance of art. (H:O:A)	-0.05	0.28	0.13	-0.04	**0.54**	-0.11	1.79
Come up with something new. (H:O:C)	0.11	-0.02	0.11	0.11	**0.53**	0.12	1.40
Carry the conversation to a higher level. (H:O:C)	0.20	0.00	0.02	0.13	**0.53**	0.09	1.49
Enjoy being thought of as a normal "mainstream" person. (H:O:U)	0.12	0.09	0.11	0.26	**-0.53**	0.15	1.93
Enjoy intellectual games. (H:O:I)	-0.15	-0.02	0.15	0.18	**0.52**	0.10	1.70
Avoid difficult reading material. (H:O:I)	0.13	0.08	0.01	-0.09	**-0.51**	0.17	1.47
Have excellent ideas. (H:O:C)	0.17	-0.07	0.02	0.22	**0.51**	0.10	1.76
Know that my ideas sometimes surprise people. (H:O:U)	0.01	0.00	0.06	0.03	**0.51**	0.17	1.27
Swim against the current. (H:O:U)	0.00	-0.11	-0.02	-0.08	**0.51**	0.08	1.21
Like to be thought of as a normal kind of person. (H:O:U)	0.04	0.17	0.15	0.30	**-0.50**	0.08	**2.20**
See beauty in things that others might not notice. (H:O:A)	-0.01	0.25	0.20	0.10	**0.50**	-0.03	1.96
Have read the great literary classics. (H:O:A)	-0.09	0.08	0.03	0.05	**0.50**	-0.04	1.16
Will not probe deeply into a subject. (H:O:I)	0.06	0.00	0.01	-0.08	**-0.49**	0.14	1.25
Love to think up new ways of doing things. (H:O:C)	0.04	-0.01	0.17	0.13	**0.49**	0.12	1.52
Am considered to be kind of eccentric. (H:O:U)	-0.03	0.07	-0.01	-0.20	**0.49**	0.15	1.59
Try to avoid complex people. (H:O:U)	-0.07	0.06	-0.04	0.07	**-0.48**	0.09	1.21
See myself as an average person. (H:H:M)	-0.05	0.19	0.20	0.10	**-0.48**	0.04	1.82
Have difficulty imagining things. (H:O:C)	-0.09	-0.04	-0.05	-0.01	**-0.48**	0.11	1.21
Consider myself an average person. (H:H:M)	-0.05	0.21	0.17	0.09	**-0.47**	0.01	1.83

continued on next page

Table 2.9 – Six factor solution for the IPIP-HEXACO in Sample 1 (continued from previous page)

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Do not like art. (H:O:A)	0.01	-0.26	-0.05	0.04	**-0.45**	0.18	**2.01**
Seldom experience sudden intuitive insights. (H:O:C)	-0.01	-0.01	0.03	-0.07	**-0.45**	0.13	1.23
Would hate to be considered odd or strange. (H:O:U)	0.06	0.18	0.01	0.22	**-0.44**	0.16	**2.19**
Am just an ordinary person. (H:H:M)	-0.09	0.19	0.19	0.10	**-0.44**	-0.01	**2.00**
Seldom notice the emotional aspects of paintings and pictures. (H:O:A)	-0.08	-0.23	-0.03	0.02	**-0.44**	0.25	**2.26**
Do not have a good imagination. (H:O:C)	-0.08	-0.08	-0.05	0.01	**-0.43**	0.08	1.24
Do not like poetry. (H:O:A)	-0.05	-0.23	-0.09	0.04	**-0.43**	0.14	1.94
Do things that others find strange. (H:O:U)	-0.20	0.08	0.00	-0.17	**0.43**	0.19	**2.30**
Rebel against authority. (H:O:U)	0.00	-0.13	-0.14	-0.29	**0.42**	0.17	**2.67**
Find political discussions interesting. (H:O:I)	-0.01	-0.05	0.01	0.00	**0.38**	0.08	1.12
Am interested in science. (H:O:I)	-0.24	-0.07	0.15	0.06	**0.37**	0.12	**2.55**
Don't know much about history. (H:O:I)	0.03	0.12	0.01	-0.04	**-0.36**	0.02	1.26
Would love to explore strange places. (H:O:I)	0.10	0.05	0.14	-0.04	**0.36**	0.09	1.68
Detect mistakes. (H:C:PE)	-0.14	-0.03	-0.05	**0.35**	**0.35**	0.12	**2.63**
Get deeply immersed in music. (H:O:A)	-0.02	0.20	0.07	-0.07	**0.35**	0.04	1.82
Don't bother worrying about political and social problems. (H:O:I)	-0.04	-0.07	0.08	-0.02	**-0.33**	0.14	1.62
Enjoy feeling "close to the earth." (H:O:A)	-0.01	0.18	0.22	0.05	0.27	-0.02	**2.85**
Would never go riding down a stretch of rapids in a canoe. (H:E:F)	-0.16	0.18	-0.10	0.10	-0.20	-0.07	**4.26**
Play a role in order to impress people. (H:H:S)	0.03	0.18	0.01	-0.05	0.06	**0.60**	1.23
Tell other people what they want to hear so that they will do what I want them to do. (H:H:S)	-0.03	-0.04	-0.05	-0.03	-0.02	**0.60**	1.03
Put on a show to impress people. (H:H:S)	0.20	0.13	0.02	-0.13	0.15	**0.59**	1.60
Would like to have more power than other people. (H:H:M)	0.11	-0.15	-0.21	0.12	0.07	**0.56**	1.68
Try to impress others. (H:H:GA)	0.23	0.29	0.00	0.01	0.02	**0.54**	1.95
Seek status. (H:H:GA)	0.24	0.05	-0.10	0.14	-0.02	**0.52**	1.70
Have a strong need for power. (H:H:GA)	0.23	-0.13	-0.24	0.15	0.05	**0.51**	**2.25**
Am likely to show off if I get the chance. (H:H:M)	**0.32**	0.04	-0.10	-0.11	0.11	**0.51**	1.99
Am out for my own personal gain. (H:H:GA)	0.00	-0.25	-0.18	-0.03	-0.05	**0.51**	1.77
Pretend to be concerned for others. (H:H:S)	-0.22	-0.23	-0.14	-0.06	-0.08	**0.48**	**2.20**
Use flattery to get ahead. (H:H:S)	0.26	0.03	-0.07	-0.04	0.02	**0.48**	1.63
Cheat to get ahead. (H:H:F)	0.02	-0.18	-0.12	-0.26	-0.07	**0.48**	**2.07**
Act like different people in different situations. (H:H:S)	-0.17	0.13	0.02	-0.13	0.13	**0.48**	1.75

continued on next page

Table 2.9 – Six factor solution for the IPIP-HEXACO in Sample 1 (continued from previous page)

Item	Extraversion	Emotionality	Agreeableness	Conscientiousness	Openness	Honesty/Humility	complexity
Am mainly interested in money. (H:H:GA)	0.09	-0.14	-0.14	0.14	**-0.32**	**0.46**	**2.54**
Admire a really clever scam. (H:H:F)	-0.10	-0.19	0.01	-0.17	0.20	**0.46**	**2.20**
Switch my loyalties when I feel like it. (H:H:S)	-0.03	-0.04	-0.05	-0.17	-0.07	**0.43**	1.42
Believe that I am better than others. (H:H:M)	0.00	-0.24	-0.20	0.03	0.22	**0.43**	**2.63**
Would not regret my behavior if I were to take advantage of someone impulsively. (H:H:F)	0.00	-0.28	-0.16	-0.11	-0.02	**0.41**	**2.31**
Boast about my virtues. (H:H:M)	0.24	0.00	-0.14	0.00	-0.01	**0.40**	1.94
Find it necessary to please the people who have power. (H:H:S)	0.10	**0.31**	0.07	0.16	-0.19	**0.39**	**3.04**
Love luxury. (H:H:GA)	0.21	0.08	-0.10	0.18	-0.09	**0.37**	**2.56**
Find fault with everything. (H:A:G)	-0.28	0.06	**-0.33**	0.00	0.08	**0.36**	**3.06**
Don't pretend to be more than I am. (H:H:S)	-0.05	-0.04	0.14	0.19	-0.03	**-0.36**	**2.01**
Steal things. (H:H:F)	-0.02	-0.11	-0.08	-0.26	0.06	**0.33**	**2.42**
Love dangerous situations. (H:E:F)	0.16	-0.27	0.04	-0.14	0.23	**0.32**	**3.82**
Don't think that I'm better than other people. (H:H:M)	-0.07	0.24	0.30	0.00	-0.17	**-0.31**	**3.61**
Cheat on people who have trusted me. (H:H:F)	0.04	-0.06	-0.13	-0.18	-0.04	0.28	**2.39**
Act impulsively when something is bothering me. (H:C:PR)	0.11	0.25	-0.22	-0.20	0.07	0.28	**4.28**
Like to do frightening things. (H:E:F)	0.15	-0.20	0.10	-0.12	0.17	0.27	**4.15**
Wish to stay young forever. (H:H:GA)	0.03	0.05	0.00	0.02	-0.06	0.26	1.21
Prefer to eat at expensive restaurants. (H:H:GA)	0.18	0.01	-0.08	0.10	0.08	0.25	**2.68**
SS loadings	18.3	13.84	14.05	13.34	12.28	12.50	

Table 2.10: IPIP100 – Factor congruence between samples

		Sample 2				
		Extraversion	Stability	Conscientiousness	Agreeableness	Intellect
Sample 1	Extraversion	1.00	-0.09	0.04	0.14	0.06
	Stability	-0.06	0.99	-0.04	-0.02	-0.07
	Conscientiousness	0.05	-0.08	0.99	0.08	0.01
	Agreeableness	0.13	-0.01	0.07	0.99	0.06
	Intellect	0.05	-0.05	0.05	0.03	0.99

Table 2.11: BFAS – Factor congruence between samples

		Sample 2				
		Stability	Extraversion	Agreeableness	Conscientiousness	Openness
Sample 1	Stability	0.99	-0.08	-0.04	-0.06	-0.06
	Extraversion	-0.15	0.99	0.07	0.11	0.08
	Agreeableness	0.01	0.07	0.98	0.04	0.16
	Conscientiousness	-0.13	0.12	0.06	0.98	-0.10
	Openness	-0.05	0.16	-0.04	0.07	0.97

The six factor solution for the QB6 items in Sample 1 is shown in Table **??**. Similar to the HEXACO loadings, most of the highly loaded items on the six QB6 factors were from the same scales, though there were several items with secondary loadings above ± 0.3 or without a primary loading above ± 0.3. The two items which loaded highly on inappropriate factors were "Waste my time", a Conscientiousness item which loaded on Resiliency (reverse coded) (0.41); and "Get back at people who insult me", an Agreeableness item which loaded highly on Honesty/Propriety (0.48). 14 of the 48 QB6 items had complexities above 2.0 in this solution.

Tables 2.10, 2.11, 2.12, and 2.13 illustrate the congruence of factors for each set of scales across Sample 1 and Sample 2. In all cases, correlations between factors representing the same construct ranged between 0.96 and 1.00. In most cases, the correlations among factors for different constructs were low (less than ± 0.15). The exception to this was found in the solution for the six factor scales – the IPIP-HEXACO and QB6 – where the correlations were slightly higher between the Agreeableness, Honesty/Humility, and Emotionality/Resiliency factors.

Table 2.12: IPIP-HEXACO – Factor congruence between samples

		Sample 2					
		Extraversion	Conscientiousness	Agreeableness	Emotionality	Honesty/Humility	Openness
Sample 1	Extraversion	0.99	0.08	-0.06	0.08	0.05	0.08
	Conscientiousness	0.06	0.99	-0.04	0.04	-0.11	0.04
	Agreeableness	0.09	0.03	0.99	-0.13	-0.21	0.06
	Emotionality	0.08	0.03	-0.15	0.99	-0.22	0.06
	Honesty/Humility	0.04	-0.11	0.14	-0.23	0.97	0.03
	Openness	0.06	0.01	-0.12	0.00	0.03	0.99

Table 2.13: QB6 – Factor congruence between samples

		Sample 2					
		Extraversion	Conscientiousness	Resiliency	Honesty/Propriety	Agreeableness	Originality
Sample 1	Extraversion	0.98	0.03	-0.10	-0.02	-0.03	-0.09
	Conscientiousness	0.00	0.97	-0.06	-0.16	-0.05	-0.03
	Resiliency	-0.10	-0.14	0.97	-0.05	-0.13	0.04
	Honesty/Propriety	-0.09	-0.22	0.01	0.96	-0.24	-0.05
	Agreeableness	-0.04	-0.05	-0.13	-0.19	0.96	-0.01
	Originality	-0.05	-0.02	0.11	-0.06	-0.06	0.97

Table 2.14: Sample 1 scale intercorrelations corrected for item overlap and attenuation

	IPIP					HEXACO						QB6					
	Agree	Consc	Extra	Intel	Stabl	Agree	Consc	Extra	Open	Emoti	HonHu	Agree	Consc	Extra	Origi	Resil	HonPr
IPIP100 Agreeableness	*0.92*	0.30	0.51	0.20	0.13	0.51	0.28	0.55	0.17	-0.28	0.36	0.41	0.37	0.70	0.25	0.18	0.49
IPIP100 Conscientiousness	0.28	*0.91*	0.24	0.13	0.23	0.18	0.98	0.27	-0.04	0.12	0.27	0.13	1.05	0.15	0.14	0.32	0.52
IPIP100 Extraversion	0.47	0.22	*0.94*	0.28	0.28	0.22	0.23	0.97	0.24	0.18	-0.21	0.00	0.30	0.89	0.34	0.46	0.01
IPIP100 Intellect	0.18	0.12	0.26	*0.89*	0.21	0.14	0.26	0.28	0.92	0.27	-0.17	0.10	0.17	0.16	1.01	0.23	-0.10
IPIP100 Stability	0.12	0.21	0.27	0.19	*0.93*	0.76	0.23	0.31	0.12	0.76	0.17	0.71	0.22	0.18	0.23	0.94	0.15
IPIP-HEXACO Agreeableness	0.48	0.17	0.20	0.13	0.71	*0.94*	0.15	0.25	0.10	0.31	0.48	0.99	0.17	0.29	0.16	0.61	0.43
IPIP-HEXACO Conscientiousness	0.26	0.90	0.21	0.24	0.21	0.14	*0.93*	0.25	0.09	0.12	0.23	0.12	1.02	0.11	0.27	0.30	0.51
IPIP-HEXACO Extraversion	0.51	0.25	0.92	0.26	0.29	0.24	0.23	*0.95*	0.22	0.17	-0.18	0.02	0.34	0.93	0.33	0.50	0.06
IPIP-HEXACO Openness	0.16	-0.03	0.22	0.82	0.11	0.09	0.08	0.21	*0.90*	0.19	-0.19	0.08	-0.02	0.16	0.97	0.11	-0.20
IPIP-HEXACO Emotionality	-0.26	0.11	0.16	0.24	0.70	0.29	0.11	0.16	0.17	*0.91*	-0.13	0.25	0.09	-0.08	0.25	0.83	-0.19
IPIP-HEXACO Honesty Humility	0.33	0.25	-0.19	-0.16	0.16	0.44	0.21	-0.16	-0.17	-0.12	*0.89*	0.52	0.27	-0.05	-0.22	0.11	0.88
QB6 Agreeableness	0.34	0.11	0.00	0.08	0.60	0.84	0.10	0.02	0.07	0.21	0.43	*0.76*	0.09	0.08	0.11	0.46	0.40
QB6 Conscientiousness	0.31	0.88	0.26	0.14	0.19	0.15	0.86	0.29	-0.02	0.07	0.22	0.07	*0.78*	0.20	0.17	0.34	0.51
QB6 Extraversion	0.60	0.13	0.77	0.13	0.16	0.25	0.09	0.82	0.14	-0.07	-0.04	0.06	0.16	*0.80*	0.23	0.31	0.15
QB6 Originality Intellect	0.20	0.11	0.28	0.79	0.18	0.13	0.21	0.27	0.76	0.20	-0.17	0.08	0.13	0.17	*0.68*	0.27	-0.09
QB6 Resiliency	0.15	0.27	0.39	0.19	0.81	0.53	0.25	0.44	0.09	0.70	0.10	0.36	0.26	0.25	0.20	*0.79*	0.15
QB6 Honesty Propriety	0.37	0.39	0.01	-0.07	0.12	0.34	0.39	0.05	-0.15	-0.15	0.66	0.28	0.36	0.11	-0.06	0.11	*0.64*

All correlations have been corrected for item overlap. The alpha for each scale is reported on the diagonal. Uncorrelated correlations are shown below the diagonal and the alpha corrected correlations are reported above.

Tables 2.14 and 2.15 contain the correlations between scales within and across sets at the five and six factor levels. These tables include the alphas for each scale along the diagonal, the correlations corrected for item overlap below the diagonal (this is only relevant for correlations across item sets), and the correlations corrected for item overlap and attenuation above the diagonal.

Table 2.15: Sample 2 scale intercorrelations corrected for item overlap and attenuation

	IPIP					HEXACO						QB6					
	Agree	Consc	Extra	Intel	Stabl	Agree	Consc	Extra	Open	Emoti	HonHu	Agree	Consc	Extra	Origi	Resil	HonPr
IPIP100 Agreeableness	*0.91*	0.30	0.47	0.16	0.16	0.55	0.27	0.51	0.17	-0.27	0.45	0.46	0.35	0.66	0.23	0.22	0.53
IPIP100 Conscientiousness	0.27	*0.92*	0.24	0.15	0.27	0.20	0.98	0.26	-0.03	0.15	0.29	0.17	1.04	0.12	0.10	0.31	0.58
IPIP100 Extraversion	0.43	0.22	*0.94*	0.28	0.31	0.21	0.20	0.98	0.25	0.22	-0.19	-0.02	0.26	0.87	0.34	0.48	0.02
IPIP100 Intellect	0.15	0.14	0.26	*0.89*	0.26	0.17	0.25	0.28	0.91	0.29	-0.10	0.11	0.18	0.16	1.00	0.26	-0.02
IPIP100 Stability	0.15	0.25	0.29	0.23	*0.93*	0.75	0.27	0.34	0.13	0.74	0.15	0.67	0.26	0.21	0.27	0.96	0.17
IPIP-HEXACO Agreeableness	0.51	0.19	0.20	0.15	0.70	*0.93*	0.17	0.25	0.14	0.27	0.48	0.97	0.20	0.27	0.22	0.63	0.45
IPIP-HEXACO Conscientiousness	0.25	0.91	0.19	0.23	0.25	0.16	*0.93*	0.22	0.05	0.16	0.27	0.17	1.02	0.07	0.22	0.28	0.58
IPIP-HEXACO Extraversion	0.47	0.24	0.92	0.26	0.32	0.24	0.21	*0.95*	0.24	0.22	-0.16	0.01	0.30	0.92	0.35	0.57	0.07
IPIP-HEXACO Openness	0.16	-0.02	0.23	0.81	0.12	0.13	0.05	0.22	*0.90*	0.21	-0.11	0.12	0.01	0.18	1.00	0.16	-0.13
IPIP-HEXACO Emotionality	-0.24	0.13	0.20	0.26	0.68	0.25	0.14	0.20	0.19	*0.90*	-0.17	0.17	0.11	-0.07	0.27	0.84	-0.18
IPIP-HEXACO Honesty Humility	0.41	0.26	-0.17	-0.09	0.14	0.44	0.24	-0.15	-0.10	-0.15	*0.89*	0.56	0.29	-0.02	-0.13	0.10	0.84
QB6 Agreeableness	0.38	0.14	-0.02	0.09	0.56	0.82	0.14	0.01	0.10	0.14	0.46	*0.75*	0.14	0.06	0.16	0.42	0.44
QB6 Conscientiousness	0.30	0.89	0.22	0.15	0.22	0.17	0.88	0.26	0.01	0.09	0.24	0.11	*0.80*	0.16	0.15	0.30	0.60
QB6 Extraversion	0.56	0.11	0.76	0.14	0.18	0.23	0.06	0.80	0.16	-0.06	-0.01	0.05	0.13	*0.81*	0.23	0.36	0.10
QB6 Originality Intellect	0.18	0.08	0.27	0.78	0.21	0.18	0.18	0.28	0.78	0.21	-0.10	0.12	0.11	0.17	*0.68*	0.34	-0.01
QB6 Resiliency	0.18	0.26	0.41	0.22	0.81	0.53	0.23	0.48	0.14	0.70	0.09	0.32	0.23	0.28	0.24	*0.77*	0.13
QB6 Honesty Propriety	0.40	0.43	0.02	-0.02	0.13	0.34	0.44	0.05	-0.10	-0.14	0.63	0.30	0.42	0.07	-0.01	0.09	*0.62*

All correlations have been corrected for item overlap. The alpha for each scale is reported on the diagonal. Uncorrelated correlations are shown below the diagonal and the alpha corrected correlations are reported above.

2.2.4 Discussion

The fit statistics for factor analyses with extraction of 1 to 20 factors were generally inconclusive for all sets of scales. With regards to complexity in particular, the fact that the five factor solutions demonstrated decreases or only small increases in complexity at 5 factors of extraction suggests that this number of factors provides a relatively simple solution across the item sets, even for the six factor sets of scales. Despite this (rather weak) suggestion regarding the superiority of five factors, the content of the prescribed factors was generally recovered for each set of scales at the appropriate level of extraction. The only exception to this was the BFAS, where the tenth factor was unidentifiable (the Politeness items were distributed among several of the other nine factors).

In terms of the predictions regarding complexity, data from both samples confirmed that the IPIP100 scales have lower mean item complexity than the BFAS scales at nearly all levels of factor extraction. Mean complexity for the IPIP100 at 5 and 10 factors was 1.3 and 2.0, respectively, while mean complexity at the same levels of extraction for the BFAS was approximately 1.7 and 2.3, respectively. It was also the case, as predicted, that the QB6 was had lower mean item complexity than the IPIP-HEXACO. The results of both of these comparisons are presumed to reflect the fact that the BFAS and the IPIP-HEXACO were designed to allow for hierarchical assessment. Note that it was not the case that the 5 factor sets of scales were consistently less complex than the 6 factor sets; the QB6 generally demonstrated lower mean complexity than the BFAS and, at higher levels of extraction, similar complexity to the IPIP100.

There are several prominent findings suggested by Tables 2.14 and 2.15. First, the correlations are highly consistent across the samples. Only 4 of the 992 correlations differ in magnitude by 0.10 or more (the maximum difference is between the attenuation-corrected correlations for BFAS Politeness and Compassion).

Second, the tables provide empirical evidence of the degree of orthogonality between scales within each of the four sets. Correlations between the IPIP100 scales after correcting for attenuation were mainly between 0.20 and 0.3, with slightly lower correlations between Intellect

and Conscientiousness (0.13) and Agreeableness and Stability (0.13). The highest correlation among these scales was between Extraversion and Agreeableness, which were strongly correlated at 0.51 in Sample 1. (Note that the values of the following correlations are based on the attenuation-corrected correlations in Table 2.14; highly similar values are reported for Sample 2 in Table 2.15). The magnitude of the correlations were similar for the BFAS at the five factor level but there were larger correlations among the 10 BFAS aspects. Several of these merely reflected the high correlations between aspects of the same factor (as expected): Politeness and Compassion (0.51), Industriousness and Orderliness (0.55), Enthusiasm and Assertiveness (0.59), Intellect and Openness (0.44), and Volatility and Withdrawal (0.72). Several large correlations among aspects from different factors were also evident: Assertiveness and Industriousness (0.52), Enthusiasm and Compassion (0.58), Intellect and Assertiveness (0.50), Withdrawal (reverse-coded) and Industriousness (0.63), and Withdrawal (reverse-coded) and Assertiveness (0.55). For both of the six-dimensional sets of scales, the loadings were generally around ± 0.3 or less except for the constellation of factors including Agreeableness, Emotionality (Resiliency in the QB6), and Honesty/Humility (Honesty/Propriety in the QB6).

The third result evidenced by Tables 2.14 and 2.15 stems from the suggestion that the correlations between similarly framed constructs across sets of scales varied considerably. For example, IPIP100 Extraversion correlated 0.88 with BFAS factor level Extraversion, 0.92 with IPIP-HEXACO Extraversion and 0.77 with QB6 Extraversion, but IPIP100 Agreeableness, by contrast, correlated 0.82 with BFAS factor level Agreeableness, 0.48 with IPIP-HEXACO Agreeableness and 0.34 with QB6 Agreeableness. In fact, among the QB6 factors, IPIP100 Agreeableness was almost as highly correlated with Conscientiousness (0.31) and more highly correlated with Extraversion (0.60) and Honesty/Propriety (0.37).

2.3 Study 2

The goal of study 2 was to evaluate the potential for improving the scales administered in Study 1 by means of reducing their complexity. More specifically, the intent was to consider whether the scales could be made less complex (e.g., more simple) by extending the factor structure onto the remainder of items administered through the use factor extension procedures. This study was limited to the sets of scales which were not hierarchically organized (the IPIP100 and the QB6) as those which were designed with specific hierarchical structures in mind are less well-suited to revision at any particular level based solely on complexity. This is because, as explored in Study 1, hierarchical scales which make use of all items at multiple levels are designed to include more complexity than would be necessary at any individual level. To the extent that the data in these large international samples differ from that which was used to develop these IPIP-based sets of scales originally (the Eugene-Springfield Community Sample), it was expected that the complexity of the IPIP100 and QB6 could be reduced by the inclusion of alternate items.

2.3.1 Methods

The samples and measures used in Study 2 were identical to those used in Study 1. The factor extension procedures used are based on those available in the *psych* package in R (Revelle, 2014) and are essentially based on the extension of exploratory factor analysis to new variables which were not part of the original set of items factored. More specifically, the five factor solution based on the 100 items of the IPIP100 was extended onto the remaining 273 items administered using the underlying correlations among all items (this was done independently for both samples). The top 20 items for each of the five factors were then compared across the samples. Items from the IPIP100 scales which were missing from the top 20 items in *both* samples were dropped and items which were not in the original IPIP100 scales but which were present among the top 20 items in *both* samples were added. It should be noted that no particular emphasis was placed on the need to retain exactly 20 items in each of the five scales. In order to compare the revised scales to the factor analytic output presented in Study 1, the last step was to extract 5 factors from the items identified as the new revised set by themselves (rather than with all 373 items together). These same procedures were used for the QB6 scales with allowances for the different number of items and factors.

2.3.2 Results

The factor extension procedures reduced the mean item complexity from 1.33 for the original 100 items of the IPIP100 scales to a mean item complexity of 1.18 for the 99 items of the revised scales where the minimum possible value for mean item complexity is 1.0. As shown in Table 2.16, this resulted in the addition of 20 new items (and the removal of 21 of the original items). The full set of items and their respective loadings on each factor is given in Table 2.19. Table 2.16 also suggests that the psychometric properties of the revised scales were improved across

Table 2.16: Comparison of scale properties: IPIP100 and revised IPIP

	Agreeableness	Conscientiousness	Extraversion	Intellect	Stability
(Standardized) Alpha:					
Sample 1 IPIP100	0.92	0.91	0.94	0.89	0.93
Sample 2 IPIP100	0.91	0.92	0.94	0.89	0.93
Sample 1 IPIP Revised	0.94	0.92	0.95	0.90	0.95
Sample 2 IPIP Revised	0.93	0.93	0.95	0.90	0.94
Average item correlation:					
Sample 1 IPIP100	0.36	0.34	0.46	0.29	0.41
Sample 2 IPIP100	0.34	0.35	0.44	0.29	0.39
Sample 1 IPIP Revised	0.44	0.39	0.52	0.31	0.46
Sample 2 IPIP Revised	0.40	0.40	0.50	0.32	0.43
Signal to Noise ratio based upon average r and n:					
Sample 1 IPIP100	11.2	10.3	16.8	8.0	13.9
Sample 2 IPIP100	10.1	10.9	15.8	8.3	12.7
Sample 1 IPIP Revised	15.5	11.9	20.2	9.0	18.1
Sample 2 IPIP Revised	13.5	12.9	18.8	9.4	16.1
Number of items:					
IPIP100	20	20	20	20	20
IPIP Revised	20 (5 new)	19 (3 new)	19 (4 new)	20 (3 new)	21 (5 new)

both samples for all five scales. Table 2.17 shows the scale intercorrelations for both the original and revised scales after correcting for item overlap (below the diagonal) and after correcting for item overlap and attenuation (above the diagonal). The correlations between constructs across the original and revised scales is uniformly high, ranging from 0.89 to 0.94 before correcting for attenuation. Table 2.18 shows the factor congruences for the revised scales between Sample 1 and Sample 2. The congruences suggested that the scales function identically in both samples (congruences were 0.99 to 1.0).

Table 2.17: IPIP100 and IPIP Revised scale intercorrelations in sample 1

	IPIP100					IPIP Revised				
	Agree	Consc	Extra	Intel	Stabl	Agree	Consc	Extra	Intel	Stabl
IPIP100 Agreeableness	*0.92*	0.30	0.51	0.20	0.13	0.99	0.30	0.51	0.16	0.12
IPIP100 Conscientiousness	0.28	*0.91*	0.24	0.13	0.23	0.25	1.00	0.18	0.13	0.19
IPIP100 Extraversion	0.47	0.22	*0.94*	0.28	0.28	0.43	0.24	0.99	0.25	0.21
IPIP100 Intellect	0.18	0.12	0.26	*0.89*	0.21	0.21	0.07	0.20	1.00	0.20
IPIP100 Stability	0.12	0.21	0.27	0.19	*0.93*	0.04	0.23	0.23	0.22	0.99
IPIP R Agreeableness	0.92	0.23	0.41	0.19	0.03	*0.94*	0.24	0.44	0.17	0.03
IPIP R Conscientiousness	0.27	0.91	0.22	0.06	0.22	0.22	*0.92*	0.19	0.06	0.19
IPIP R Extraversion	0.48	0.17	0.94	0.19	0.22	0.41	0.18	*0.95*	0.18	0.16
IPIP R Intellect	0.14	0.11	0.23	0.89	0.20	0.15	0.06	0.16	*0.90*	0.21
IPIP R Stability	0.11	0.18	0.20	0.19	0.93	0.03	0.18	0.15	0.20	*0.95*

All correlations have been corrected for item overlap. The alpha for each scale is reported on the diagonal and the alpha corrected correlations are reported above the diagonal.

Table 2.18: Revised IPIP – Factor congruence between samples

		Sample 2				
		Agreeableness	Conscientiousness	Extraversion	Intellect	Stability
Sample 1	Agree	0.99	0.04	0.09	0.05	0.01
	Conscientiousness	0.05	0.99	0.03	0.01	-0.06
	Extraversion	0.11	0.02	1.00	0.04	-0.05
	Intellect	0.03	0.03	0.02	0.99	-0.05
	Stability	0.00	-0.02	-0.03	-0.05	1.00

Table 2.19: Five factor solution for the revised IPIP in Sample 1

Item	Stability	Extraversion	Agreeableness	Conscientiousness	Intellect	complexity	new item
Get easily agitated.	**0.82**	0.01	-0.12	0.03	0.02	1.05	Y
Get irritated easily.	**0.80**	0.07	-0.19	0.06	0.03	1.15	N
Get upset easily.	**0.80**	-0.01	0.06	0.00	-0.07	1.03	N
Get angry easily.	**0.78**	0.12	-0.22	0.09	-0.04	1.24	N
Can be stirred up easily.	**0.76**	0.11	0.07	-0.01	-0.01	1.06	Y
Am easily annoyed.	**0.74**	0.01	-0.23	0.06	0.05	1.21	Y
Lose my temper.	**0.74**	0.12	-0.12	0.01	0.02	1.11	Y
Rarely get irritated.	**-0.73**	-0.05	0.13	-0.08	0.00	1.10	N
Am a person whose moods go up and down easily.	**0.72**	-0.09	0.10	-0.14	0.00	1.15	Y
Get stressed out easily.	**0.72**	-0.15	0.19	-0.01	-0.05	1.24	N
Have frequent mood swings.	**0.71**	-0.03	0.08	-0.14	0.01	1.11	N
Am not easily bothered by things.	**-0.69**	0.01	-0.08	-0.04	0.08	1.06	N
Change my mood a lot.	**0.66**	-0.03	0.08	-0.18	0.03	1.19	N
Seldom get mad.	**-0.64**	-0.12	0.14	-0.06	0.01	1.19	N
Take offense easily.	**0.63**	-0.04	0.09	0.02	-0.11	1.11	N
Get overwhelmed by emotions.	**0.61**	-0.09	**0.35**	-0.11	-0.07	1.77	N
Panic easily.	**0.60**	-0.17	0.22	-0.08	-0.12	1.56	N
Am relaxed most of the time.	**-0.58**	0.12	-0.06	-0.05	0.00	1.12	N
Worry about things.	**0.58**	-0.20	0.23	0.04	0.00	1.58	N
Am easily disturbed.	**0.55**	-0.04	0.09	-0.01	-0.07	1.10	N
Get caught up in my problems.	**0.55**	-0.18	0.16	-0.16	0.05	1.62	N
Keep in the background.	-0.02	**-0.79**	0.06	-0.03	0.01	1.02	N
Am the life of the party.	0.04	**0.77**	-0.04	-0.03	0.05	1.02	N
Say little.	-0.13	**-0.77**	-0.04	0.03	-0.05	1.08	Y
Don't talk a lot.	-0.14	**-0.77**	-0.08	0.03	0.02	1.09	N
Talk to a lot of different people at parties.	-0.01	**0.76**	0.04	0.01	0.03	1.01	N
Am quiet around strangers.	0.04	**-0.75**	0.03	-0.01	-0.03	1.01	N
Start conversations.	0.02	**0.74**	0.11	0.08	0.05	1.08	N
Don't mind being the center of attention.	0.06	**0.73**	-0.13	-0.10	0.12	1.17	N
Find it difficult to approach others.	0.15	**-0.72**	0.02	-0.10	0.03	1.13	N
Feel comfortable around people.	-0.17	**0.72**	0.16	0.07	-0.02	1.23	N
Talk a lot.	0.26	**0.72**	0.09	-0.06	0.01	1.32	Y
Hate being the center of attention.	-0.07	**-0.72**	0.09	0.11	-0.06	1.12	Y
Feel at ease with people.	-0.19	**0.70**	0.14	0.07	-0.02	1.25	N
Often feel uncomfortable around others.	0.23	**-0.69**	0.01	-0.09	0.06	1.27	N
Don't like to draw attention to myself.	-0.07	**-0.69**	0.13	0.14	-0.02	1.18	N
Make friends easily.	-0.05	**0.68**	0.15	0.09	-0.05	1.15	N

continued on next page

Table 2.19 – Five factor solution for the revised IPIP in Sample 1(continued from previous page)

Item	Stability	Extraversion	Agreeableness	Conscientiousness	Intellect	complexity	new item
Seem to derive less enjoyment from interacting with people than others do.	0.09	**-0.66**	-0.24	-0.03	0.16	1.43	Y
Am skilled in handling social situations.	-0.08	**0.66**	0.16	0.13	0.10	1.28	N
Am a very private person.	0.02	**-0.61**	-0.08	0.09	0.16	1.21	N
Sympathize with others' feelings.	0.05	0.00	**0.80**	-0.01	0.00	1.01	N
Am sensitive to the needs of others.	0.03	-0.02	**0.79**	0.06	0.03	1.02	Y
Have a soft heart.	0.05	-0.04	**0.72**	0.03	-0.10	1.06	N
Love to help others.	-0.04	0.08	**0.72**	0.07	0.00	1.05	N
Feel others' emotions.	0.10	0.05	**0.71**	0.01	0.06	1.07	N
Am indifferent to the feelings of others.	0.04	0.04	**-0.71**	0.03	-0.05	1.02	N
Can't be bothered with other's needs.	0.07	0.02	**-0.69**	-0.01	-0.02	1.02	Y
Like to do things for others.	-0.07	0.04	**0.69**	0.07	0.03	1.05	Y
Think of others first.	-0.09	-0.06	**0.69**	0.13	-0.06	1.14	N
Take an interest in other people's lives.	0.06	0.15	**0.67**	-0.05	0.06	1.14	Y
Am deeply moved by others' misfortunes.	0.10	-0.08	**0.67**	-0.02	0.05	1.09	Y
Inquire about others' well-being.	0.04	0.15	**0.66**	0.05	0.06	1.14	N
Take time out for others.	-0.08	0.10	**0.66**	0.08	0.00	1.11	N
Feel little concern for others.	0.04	0.03	**-0.64**	0.02	-0.02	1.02	N
Am not interested in other people's problems.	0.06	-0.07	**-0.62**	0.07	0.01	1.07	N
Am not really interested in others.	0.04	-0.28	**-0.56**	0.01	0.06	1.51	N
Know how to comfort others.	-0.01	0.26	**0.54**	0.08	0.05	1.53	N
Am interested in people.	-0.03	**0.32**	**0.51**	-0.05	0.06	1.74	N
Have a good word for everyone.	-0.19	0.08	**0.46**	0.06	0.02	1.45	N
Show my gratitude.	-0.04	0.13	**0.43**	0.19	0.08	1.70	N
Keep things tidy.	0.08	0.01	-0.02	**0.76**	-0.03	1.03	Y
Complete my duties as soon as possible.	0.05	0.02	0.05	**0.70**	0.00	1.02	Y
Get chores done right away.	0.01	0.04	0.04	**0.68**	-0.03	1.02	N
Like to tidy up.	0.13	0.01	0.08	**0.66**	-0.03	1.11	N
Leave a mess in my room.	0.07	-0.01	0.04	**-0.66**	0.10	1.08	N
Follow a schedule.	0.12	-0.01	0.07	**0.66**	-0.04	1.09	N
Neglect my duties.	0.14	-0.05	-0.07	**-0.65**	0.01	1.13	N
Love order and regularity.	0.21	-0.16	0.03	**0.65**	-0.06	1.36	N
Do things according to a plan.	0.10	-0.07	0.03	**0.65**	0.02	1.08	N
Make plans and stick to them.	0.02	0.03	0.05	**0.65**	0.05	1.03	N
Often forget to put things back in their proper place.	0.08	0.04	0.06	**-0.64**	0.04	1.06	N
Like order.	0.20	-0.09	0.02	**0.63**	0.02	1.24	N
Am always prepared.	0.00	0.00	0.01	**0.63**	0.15	1.11	N
Do things in a half-way manner.	0.10	-0.04	-0.07	**-0.60**	-0.13	1.18	N
Finish what I start.	-0.09	0.11	0.04	**0.60**	0.05	1.14	Y

continued on next page

Table 2.19 – Five factor solution for the revised IPIP in Sample 1(continued from previous page)

Item	Stability	Extraversion	Agreeableness	Conscientiousness	Intellect	complexity	new item
Leave my belongings around.	0.06	0.05	0.13	**-0.59**	0.09	1.18	N
Make a mess of things.	0.21	-0.03	0.06	**-0.59**	-0.05	1.29	N
Shirk my duties.	0.08	0.03	-0.12	**-0.58**	-0.12	1.22	N
Find it difficult to get down to work.	0.13	-0.08	0.02	**-0.58**	-0.02	1.16	N
Love to read challenging material.	-0.01	-0.17	0.03	-0.04	**0.68**	1.14	N
Have a rich vocabulary.	0.04	-0.05	0.01	-0.01	**0.66**	1.02	N
Use difficult words.	0.09	-0.06	-0.01	-0.06	**0.65**	1.07	N
Enjoy intellectual games.	-0.06	-0.12	0.02	0.02	**0.64**	1.09	Y
Am full of ideas.	0.08	0.16	0.04	0.02	**0.64**	1.16	N
Like to solve complex problems.	-0.11	-0.13	-0.05	0.03	**0.63**	1.16	Y
Have excellent ideas.	0.05	0.19	-0.03	0.10	**0.61**	1.26	N
Have difficulty understanding abstract ideas.	0.14	0.06	0.02	0.10	**-0.61**	1.18	N
Avoid difficult reading material.	0.05	0.14	-0.01	-0.02	**-0.60**	1.13	N
Am quick to understand things.	-0.08	0.04	-0.04	0.11	**0.60**	1.13	N
Can handle a lot of information.	-0.09	0.02	-0.05	0.18	**0.59**	1.25	N
Come up with something new.	0.01	0.16	0.01	-0.02	**0.59**	1.14	Y
Carry the conversation to a higher level.	0.06	0.25	0.04	0.01	**0.57**	1.40	N
Love to think up new ways of doing things.	-0.02	0.07	0.04	0.00	**0.54**	1.05	N
Have a vivid imagination.	0.10	0.06	0.14	-0.14	**0.53**	1.40	N
Am not interested in abstract ideas.	0.07	0.09	-0.08	0.17	**-0.51**	1.38	N
Catch on to things quickly.	-0.06	0.10	-0.02	0.16	**0.50**	1.31	N
Will not probe deeply into a subject.	0.03	0.07	-0.10	0.02	**-0.49**	1.15	N
Have difficulty imagining things.	0.02	-0.05	-0.10	0.06	**-0.48**	1.15	N
Do not have a good imagination.	-0.03	-0.05	-0.10	0.06	**-0.45**	1.17	N
SS loadings	10.94	11.16	9.86	8.47	7.17		

The factor extension procedures reduced the mean item complexity from 1.75 for the original 48 items of the QB6 scales to a mean item complexity of 1.14 for the 44 items of the revised scales. This resulted in the addition of 30 new items (and the removal of 34 of the original items).

Note that two items from the original set of scales were dropped despite the fact that they were only present in the "drop" list for Sample 1. This means that these two items from the original QB6 scales were among the top 48 items in Sample 2 but not Sample 1. These items were removed because they demonstrated high complexity in both Sample 1 and Sample 2. The items were "Like to do frightening things" and "Take risks that could cause trouble for me". When these items were retained in the scales, mean item complexity was 1.25 in Sample 1 and 1.19 in Sample 2.

The full set of items and their respective loadings on each factor is given in Table 2.23. Table 2.20 also suggests that the psychometric properties of the revised scales were improved across both samples for all six scales. Table 2.21 shows the scale intercorrelations for both the original and revised scales after correcting for item overlap (below the diagonal) and after correcting for item overlap and attenuation (above the diagonal). The correlations between constructs across the original and revised scales ranged from 0.71 to 0.81 before correcting for attenuation and 0.93 to 1.00 after correcting for attenuation. Table 2.22 shows the factor congruences for the revised scales between Sample 1 and Sample 2. The congruences suggested that the revised scales function nearly identically in both samples (congruences were 0.98 to 0.99).

Table 2.20: Comparison of scale properties: QB6 and revised QB6

	Agreeable	Conscien	Extraversion	Originality	Resiliency	Honesty/ Propriety
(Standardized) Alpha:						
Sample 1 QB6	0.76	0.78	0.80	0.68	0.79	0.64
Sample 2 QB6	0.75	0.80	0.81	0.68	0.77	0.62
Sample 1 QB6 Revised	0.87	0.84	0.88	0.78	0.85	0.84
Sample 2 QB6 Revised	0.85	0.87	0.89	0.78	0.84	0.83
Average item correlation:						
Sample 1 QB6	0.29	0.30	0.34	0.21	0.32	0.18
Sample 2 QB6	0.28	0.34	0.34	0.21	0.29	0.17
Sample 1 QB6 Revised	0.49	0.44	0.49	0.28	0.48	0.42
Sample 2 QB6 Revised	0.45	0.48	0.49	0.28	0.46	0.41
Signal to Noise ratio based upon average r and n:						
Sample 1 QB6	3.3	3.5	4.0	2.2	3.7	1.7
Sample 2 QB6	3.1	4.1	4.2	2.1	3.3	1.6
Sample 1 QB6 Revised	6.6	5.4	7.6	3.6	5.5	5.1
Sample 2 QB6 Revised	5.8	6.5	7.8	3.5	5.2	5.0
Number of items:						
QB6	8	8	8	8	8	8
QB6 Revised	7 (5 new)	7 (3 new)	8 (7 new)	9 (4 new)	6 (4 new)	7 (7 new)

Table 2.21: QB6 and revised QB6 scale intercorrelations in sample 1

	QB6						QB6 Revised					
	Agree	Conscien	Extraversion	Original	Resil	Honesty/ Propriety	Agree	Conscien	Extraversion	Original	Resil	Honesty/ Propriety
QB6 Agreeableness	*0.76*	0.09	0.08	0.11	0.46	0.40	0.94	-0.03	0.08	0.11	0.36	0.43
QB6 Conscientiousness	0.07	*0.78*	0.20	0.17	0.34	0.51	0.09	0.93	0.25	0.02	0.26	0.41
QB6 Extraversion	0.07	0.16	*0.80*	0.23	0.31	0.15	0.02	0.15	0.97	0.07	0.23	0.15
QB6 Originality	0.08	0.13	0.17	*0.69*	0.27	-0.09	0.19	0.02	0.17	0.97	0.27	0.02
QB6 Resiliency	0.36	0.26	0.25	0.20	*0.79*	0.16	0.48	0.11	0.31	0.11	0.99	0.18
QB6 Honesty/Propriety	0.28	0.36	0.11	-0.06	0.11	*0.64*	0.26	0.46	0.16	-0.14	0.04	1.00
QB6 Revised Agreeableness	0.77	0.07	0.02	0.15	0.40	0.19	*0.87*	0.00	0.02	0.17	0.42	0.26
QB6 Revised Conscientiousness	-0.02	0.75	0.12	0.02	0.09	0.34	0.00	*0.85*	0.20	-0.08	0.04	0.27
QB6 Revised Extraversion	0.06	0.21	0.81	0.13	0.26	0.12	0.02	0.18	*0.88*	-0.01	0.21	0.15
QB6 Revised Originality	0.09	0.02	0.06	0.71	0.09	-0.10	0.14	-0.06	-0.01	*0.78*	0.12	-0.01
QB6 Revised Resiliency	0.29	0.21	0.19	0.21	0.81	0.03	0.36	0.03	0.19	0.10	*0.85*	0.08
QB6 Revised Honesty/Propriety	0.34	0.33	0.13	0.01	0.14	0.73	0.22	0.22	0.13	-0.01	0.07	*0.84*

All correlations have been corrected for item overlap. The alpha for each scale is reported on the diagonal and the alpha corrected correlations are reported above the diagonal.

Table 2.22: QB6 – Factor congruence between samples

		Sample 2					
		Agreeable	Conscien	Extraversion	Honesty/ Propriety	Originality	Resilience
Sample 1	Agreeableness	0.99	-0.01	-0.02	-0.05	0.03	-0.05
	Conscientiousness	0.02	0.99	0.01	-0.01	-0.07	-0.06
	Extraversion	0.00	0.05	0.99	-0.04	0.02	-0.02
	Honesty/Propriety	-0.04	-0.03	-0.02	0.99	-0.02	-0.01
	Originality	0.02	-0.01	0.02	-0.04	0.98	0.01
	Resiliency	-0.04	-0.06	-0.02	0.01	-0.03	0.99

Table 2.23: Six factor solution for the revised QB6 in Sample 1

Item	Extraversion	Agreeable	Conscientious	Honesty/Prop	Resiliency	Originality	complexity	new item
Usually enjoy being with people.	**0.86**	0.05	0.02	0.00	0.01	0.00	1.01	Y
Usually like to spend my free time with people.	**0.82**	0.02	-0.03	0.03	0.03	-0.09	1.03	Y
Seem to derive less enjoyment from interacting with people than others do.	**-0.81**	-0.01	0.03	0.02	0.06	0.06	1.03	Y
Love to chat.	**0.75**	-0.04	0.02	0.07	0.09	0.01	1.05	Y
Talk to a lot of different people at parties.	**0.67**	-0.06	0.00	0.09	-0.11	0.10	1.16	N
Don't think it's important to socialize with others.	**-0.62**	-0.06	0.02	0.08	-0.04	0.01	1.07	Y
Show my feelings when I'm happy.	**0.54**	-0.10	0.06	-0.13	0.04	0.05	1.26	Y
Laugh a lot.	**0.53**	0.03	0.02	-0.03	0.01	0.09	1.07	Y
Get angry easily.	-0.06	**-0.81**	0.05	0.04	0.06	-0.04	1.04	N
Lose my temper.	0.02	**-0.78**	-0.02	0.04	0.06	0.02	1.02	Y
Seldom get mad.	0.03	**0.77**	0.01	0.04	-0.03	-0.02	1.01	N
Rarely show my anger.	-0.10	**0.76**	0.00	0.06	0.09	-0.03	1.08	Y
Find that it takes a lot to make me feel angry at someone.	0.07	**0.71**	0.00	-0.03	-0.02	0.03	1.03	Y
Am usually a patient person.	0.01	**0.58**	0.08	-0.06	0.04	0.05	1.09	Y
Hate waiting for anything.	0.07	**-0.41**	0.03	0.09	0.05	-0.02	1.20	Y
Love order and regularity.	-0.12	-0.01	**0.76**	0.00	0.07	-0.09	1.09	N
Do things according to a plan.	0.00	0.07	**0.75**	0.04	0.01	0.00	1.02	N
Like order.	-0.02	0.00	**0.69**	0.03	0.05	0.01	1.02	N
Follow a schedule.	0.06	-0.04	**0.68**	-0.06	-0.05	-0.02	1.05	N
Like to plan ahead.	0.01	0.00	**0.66**	0.01	0.01	0.07	1.02	Y
Am a goal-oriented person.	0.23	-0.03	**0.53**	-0.02	-0.13	0.11	1.61	Y
Complete my duties as soon as possible.	0.05	-0.06	**0.52**	-0.14	-0.15	-0.02	1.37	Y
Use others for my own ends.	-0.02	-0.02	0.05	**0.78**	-0.05	0.03	1.02	Y
Take advantage of others.	0.01	-0.01	0.03	**0.77**	0.01	0.01	1.00	Y
Cheat to get ahead.	0.05	0.01	-0.09	**0.69**	-0.01	-0.07	1.07	Y
Tell other people what they want to hear so that they will do what I want them to do.	0.03	0.02	0.12	**0.65**	0.06	0.02	1.09	Y
Steal things.	-0.01	-0.01	-0.09	**0.53**	0.01	0.02	1.07	Y
Cannot imagine lying or cheating.	0.02	0.06	0.16	**-0.50**	0.03	-0.05	1.27	Y
Would never take things that aren't mine.	0.00	0.08	0.13	**-0.50**	0.06	0.00	1.21	Y
Worry about things.	0.07	-0.05	0.07	-0.04	**0.80**	0.05	1.05	N
Rarely worry.	-0.05	0.06	-0.04	0.09	**-0.74**	-0.01	1.06	Y
Am often worried by things that I said or did.	-0.02	0.08	-0.01	0.05	**0.69**	0.07	1.07	Y
Panic easily.	0.03	-0.12	-0.02	-0.02	**0.67**	-0.08	1.10	N
Am afraid of many things.	-0.04	0.00	-0.03	0.06	**0.63**	-0.13	1.12	Y
Feel a sense of worthlessness or hopelessness.	-0.23	0.01	-0.17	0.13	**0.56**	0.02	1.66	Y
Have a rich vocabulary.	0.01	0.01	0.06	0.06	0.03	**0.75**	1.03	N
Use difficult words.	-0.06	-0.02	0.03	0.04	0.03	**0.74**	1.03	N
Love to read challenging material.	-0.07	0.04	0.03	0.00	0.00	**0.63**	1.04	N

continued on next page

Table 2.23 – Six factor solution for the Revised QB6 in Sample 1 (continued from previous page)

Item	Extraversion	Agreeable	Conscientious	Honesty/Prop	Resiliency	Originality	complexity	new item
Have difficulty understanding abstract ideas.	0.03	-0.04	0.13	0.03	0.12	**-0.56**	1.22	N
Am not interested in abstract ideas.	0.02	-0.04	0.16	0.09	-0.01	**-0.50**	1.30	N
Am not all that curious about the world.	-0.09	-0.04	0.00	0.04	-0.03	**-0.44**	1.13	Y
Believe in the importance of art.	0.07	0.01	-0.09	-0.11	0.10	**0.39**	1.49	Y
Seldom experience sudden intuitive insights.	-0.04	0.03	0.08	0.08	0.05	**-0.39**	1.24	Y
Don't pride myself on being original.	-0.10	0.08	0.09	0.04	0.11	**-0.35**	1.72	Y
SS loadings	4.26	3.60	3.32	3.12	3.05	2.80		

2.3.3 Discussion

Evidence from Study 2 suggests that the use of factor extension procedures to reduce mean item complexity is an effective method for improving the psychometric properties of extant scales. In the large international samples reported here, both sets of revised scales are considerably less complex than the original versions; mean item complexity was reduced from 1.33 to 1.18 for the IPIP100 and from 1.75 to 1.14 for the QB6. These suggested revisions also offered considerable benefit in terms of more traditional psychometrics. Internal consistency measures were improved and the total number of items was decreased in both cases.

It should also be noted that the revised scales reported here are empirical with the exception of the two QB6 items previously described. This suggests that no efforts were made to balance the number of negative or positively keyed items. The number of negatively worded items in the IPIP100 revision increased as follows: 1 fewer negatively worded item for Agreeableness (6 of 20 in the original scale and 5 of 20 in the revised scale) and Conscientiousness (9 of 20 in the original scale and 8 of 19 in the revised scale); 2 fewer negatively worded items for Intellect (7 of 20 in the original scale and 5 of 20 in the revised scale); 2 additional negatively worded items for Stability (15 of 20 in the original scale and 17 of 20 in the revised scale); the number of negatively worded items for Extraversion was unchanged. Changes to the number of negatively worded items for the QB6 were as follows: Agreeableness had 5 of 8 in the original scale and 3 of 7 in the revision; Conscientiousness had 4 of 8 in the original scale and 0.00 of 7 in the revision; Extraversion had 4 of 8 in the original scale and 2 of 8 in the revision; Originality had 4 of 8 in the original scale and 5 of 9 in the revision; Resiliency had 5 of 8 in the original scale and 5 of 6 in the revision; Honesty/Propriety had 3 of 8 in the original scale and 5 of 7 in the revision.

Of course, the results of Study 2 should not be misinterpreted to imply that it is uniformly desirable to reduce complexity. To the contrary, the results of Study 1 suggested that the empirical organization of the 373 items administered in these samples is more complex than might be inferred from results which stem from the administration of any single set of scales in isolation. It is worth reiterating that complexity is only a function of rotation, and rotations

with low complexity do not fit the underlying structure of the items “better” than those with high complexity. Instead, lower complexity rotations might be thought of as representing more “purely” framed factors relative to higher complexity solutions at the same level of factor analytic extraction.

2.4 Conclusion

Even when analyses are limited to the use of personality items which have been intentionally selected for the assessment of seemingly consensual constructs such as the Big Five, evidence suggests that the personality space is highly complex. This is supported by the underlying correlational structure of the full set of 373 items administered here as well as the relatively low correlations among scales across sets which are presumed to be similarly framed measures.

This suggestion of inherent complexity does not likely come as a surprise to those who research personality structure nor perhaps to laypersons who reflect upon the myriad individual differences which are readily observed in daily interpersonal interactions. It should serve as a critical reminder however to those researchers who claim to evaluate “personality" based solely on the assessment of five or six narrow scales.

To the extent that personality assessment is limited to only 5 or 6 dimensions with presumed simple structure due to practical considerations (i.e., limited time and/or resources), it would seem advisable to make use of scales which are minimally complex based on large international samples. The revisions proposed herein provide utility in this regard, but perhaps more importantly describe methods which could be applied to improve other widely-used scales.

Of course, the 373 items administered in these two samples represent less than 15% of the extant IPIP items, which in turn represent only a subset of the personality items which are regularly

administered by personality researchers. As such, the primary advantage of the data reported here is, in some sense, not good enough. Additional research using broader sets of items is needed before preliminary conclusions can be drawn about the structure of the phrased item universe. To this end, the authors hope that the methods described herein might precipitate a collaborative and holistic attempt to evaluate the cross-sectional structure of phrased personality items in a large international sample. Such a collaboration has previously been suggested by Goldberg and colleagues (Goldberg et al., 2006) when advocating for broader use of the IPIP. In addition to echoing this suggestion, we believe that the incorporation of SAPA sampling procedures – and concerted efforts among personality researchers to pool their data collection resources for participant recruitment – would hasten the realization of this ambition.

Chapter 3

The affective domain: Structural evaluation and development of the SAPA Personality Inventory

3.1 Introduction

There now exist a surfeit of options for those in search of measures which provide well-validated scores across (slightly distinct variations of) the Big Five dimensions. Each of these has its own raison d'être and loyal advocates. An incomplete list of the most prominent include the unipolar Big Five factor markers (Goldberg, 1992), the NEO-PI-R™(Costa and McCrae, 1992), the Abridged Big-Five Dimensional Circumplex ("AB5C", Hofstee et al., 1992), the Big Five Inventory (John and Srivastava, 1999), the Five- and Ten-Item Personality Inventories (Gosling et al., 2003), and the Mini-IPIP scales (Donnellan et al., 2006). Some of these are available in both the original and slightly altered public-domain formats comprised of items from the International Personality Item Pool (e.g., the "IPIP" items corresponding to the Big Five factor markers, the NEO-PI-R™, and the AB5C). If the list is expanded to include those measures which are often interpreted in five factor terminology (such as the Hogan Personality Inventory,

Hogan, 1992) or include the additional dimension of Honesty/Humility (e.g., the IPIP-HEXACO (Lee and Ashton, 2004) and the 48-item Questionnaire Big Six scales ("QB6", Thalmayer et al., 2011)) or are derivative of the five dimensions (e.g., the Big Five Aspect Scales ("BFAS", DeYoung et al., 2007)) or have been translated into other languages (see Ashton et al., 2004; Saucier, 2009b; De Raad et al., 2010), the seemingly subtle distinctions between measures borders on esoterica.

The flourishing growth of measures illustrates the breadth of support for the (approximately) five factor structure and this consensus has, in turn, precipitated broader recognition of the utility of personality for predicting important life outcomes (Barrick and Mount, 1991; Hogan et al., 1996; John and Srivastava, 1999; Paunonen, 2003; Roberts et al., 2007). Research on the utility of the Big Five scales is based upon the administration of these measures to several million participants over the last two decades (Obschonka et al., 2013; Soto et al., 2010); over the same time frame, the nine measures described in the prior paragraph have been cited more than 17,000 times.

Yet there is one feature of these developments which is often overlooked: the majority of Big Five measures are at least one step removed from the factor-analytic procedures used to evaluate the multidimensional structure of the trait lexicon. This is readily apparent from the format of the measures. All but a few – notably, the Big Five unipolar and bipolar factor markers (Goldberg, 1992) and the rationally sorted factors of the Adjective Check List (John and Srivastava, 1999) – use phrases or sentences ("phrased items") instead of single word adjectives and type-nouns. In other words, the scales are not empirically derived, per se, from factor analyses of the finite trait lexicon.

In some cases, the scales were derived after only one additional step. For example, the IPIP items corresponding to the Big Five factor markers ("IPIP100") are the phrased items which correlate most highly with the unipolar factor markers based on administration to the Eugene-Springfield

Community Sample (Goldberg, 1999). In many more cases, the relationship between the resultant scales and unbiased factor-analytic output is less clear. The NEO-PI-R™, for example, is the result of several iterations of rational scale development with roots in re-analysis of Cattell's structural studies (Cattell et al., 1970) of the trait-descriptors, the two primary factors of Eysenck's P-E-N model (Eysenck, 1981, 1991), and later drawing upon the five factor solutions advocated by Digman (Digman and Takemoto-Chock, 1981) and Goldberg (Costa and McCrae, 1976, 1992; Goldberg, 1981, 1993b; John and Srivastava, 1999; McCrae and Costa, 1985, 1987).

That most personality measures have deviated from assessment by adjective checklist is, in many ways, an innovation rather than a liability. The benefits of using phrased items instead of single word adjectives and type-nouns are widely acknowledged (Briggs, 1992; Gosling et al., 2003; Goldberg, 1999; Hendriks et al., 1999; John and Srivastava, 1999). These include the observation that slightly longer phrases result in more clear, precise, and consistent interpretations of items by laypersons, particularly in relation to low frequency descriptors. Not only do phrased items allow for improved contextualization of trait descriptions but they are also unrestricted by the requirement of being classified as a recognizable dictionary entry.

This last point is particularly consequential. Unlike the universe of trait descriptive words, the number of items which can be formed with phrases and sentences is effectively infinite. Even the limited list of *published* phrased items is several times larger than the list of trait-descriptive words which are recognizable among the general public. For example, Norman's list of trait-descriptor words (Norman, 1967) – substantially derived from the now infamous lists by Allport and Odbert (1936) – contains a total of 2,800 words. Fewer than 1,300 of these are among the 50,000 most frequently used words in the contemporary American-English corpus (Davies, 2008). By comparison, the authors maintain a database of public-domain phrased items which, at the time of publication, contains more than 4,300 items (approximately 2,500 of these are from

the International Personality Item Pool). This list is by no means complete: proprietary items are purposefully excluded and additional items from the public-domain are added only as they are encountered by happenstance in the literature.[1] In addition to a listing of the items, it also denotes the scales with which items are affiliated. Instead of proposing a numeric estimate of published items, it seems reasonable to state that the size of this database is merely the tip of the iceberg.

What, if anything, is to be made of this difference between the finite quality of the trait lexicon and the infinite variety of phrased items? Based on the assessment methodologies advocated by most personality researchers, not much. The consensus seems to be that the benefits of phrased items outweigh the consequences of this distinction in form. But strictly speaking, the structure of the theoretical space occupied by all possible phrased items can only be *inferred* from the structure of the finite space of trait descriptors. Evaluation of the structure of the universe of phrased items within any sample is limited to the space occupied by those items which are administered.

This point extends beyond the observation that many purported structural analyses of "personality" merely constitute reification of the measures given in a new sample (Block, 1995; Cramer et al., 2012; Uher, 2013; Vassend and Skrondal, 1995). Far more important is the need to acknowledge that the structure of the phrased item universe is technically *undetermined* as long as their scope remains unmeasured. The essence of the "Lexical Hypothesis" (Galton, 1884; Goldberg, 1981, 1993b), at least insofar as it relates to personality assessment via questionnaire, is that the multidimensional space of phrased items is *structurally* similar to that of the well-studied and finite space of the trait lexicon. Presumption that the lexical hypothesis is correct allows for the implication that the larger pool of phrased items is well-represented by the smaller list of trait-descriptors. Unfortunately, the evidence for this is less consistent in

[1] This database can be found at https://sapa-project.org/data/MasterItemList.csv

terms of the number and content of the factors than that regarding the structural properties of the trait-descriptors (Ashton et al., 2004; De Raad et al., 2010; Digman and Inouye, 1986; Digman, 1990; Eysenck, 1991, 1994; Hendriks et al., 1999, 2003; Lee et al., 2005; Pettersson and Turkheimer, 2010; Yarkoni, 2013).

Consider the difference between the unipolar Big Five Factor marker "talkative" (Goldberg, 1992) and the phrased items in Table 3.1. These items represent those IPIP items which include the word "talk" and are related to talking in the first-person (items which reference being "talked to" or "talked about" by others are omitted). Note that none of the IPIP items contain the word "talkative" and that this list does not include items which are conceptually related to "talkative" but do not include the word "talk" (e.g., "Am quiet around strangers.").

Preliminary inspection of the items in Table 3.1 might suggest a surprising variety of ways to assess talkativeness, but this is not the primary issue. A more relevant concern is the degree of similarity between (1) the relationship between the items in Table 1 (both individually and as a set) and all other phrased items and (2) the relationship between "talkative" and all other trait descriptors. This second relationship has been the primary focus of psycholexical personality research and several sampling methodologies have been employed in the development of "factor markers" for the multidimensional trait descriptor space. These are well-summarized by Goldberg (1992, p. 28). In order to circumvent the challenges of administering *all* trait descriptors, various researchers have employed representative, uniform, or cluster-based sampling of the descriptors in order to assess their relative structure. While the sampling method has no bearing on the "true" structure of the trait descriptors, it is of course likely to have a consequential effect when evaluating factor solutions for parsimony. Several researchers (Allport and Odbert, 1936; Norman, 1963, 1967; Digman and Takemoto-Chock, 1981; Saucier, 2003) have noted that administration of the full universe of trait descriptors is not an ideal

Table 3.1: Items in the International Personality Item Pool which contain the word talk, exclusive of those which reference being "talked to" or "talked about"

Am willing to talk about myself.
Avoid small talk.
Can talk my way out of anything.
Can talk others into doing things.*
Dislike talking about myself.
Do most of the talking.
Don't call people just to talk.
Don't talk a lot.*
Don't talk badly to outsiders about my own group.
Have a colorful and dramatic way of talking about things.
Like to hear myself talk.
Like to talk about my future plans.
Like to talk about myself.
Never stop talking.
Only talk about my own interests.
Prefer talking to people about their daily activities rather than their feelings.
Rarely talk about sex.
Talk a lot.*
Talk about my worries.
Talk during movies.
Talk even when I know I shouldn't.
Talk for no reason.
Talk mainly about myself.
Talk nonsense.
Talk out loud to myself.
Talk to a lot of different people at parties.*
Talk to myself a lot.
Talk too much.
Talk without thinking.
Tend to talk sarcastically.
Usually like to talk a lot.
Waste time talking.
When with a group, have difficulties selecting a good topic to talk about.

* Included among the items administered in these studies.

solution to this sampling problem because many of the trait descriptors are vague and/or not widely recognized. So, while the degree of similarity between the two relationships is germane (they are presumed to be highly similar by acceptance of the Lexical Hypothesis), it is difficult to evaluate the degree of similarity between these relationships given the challenges inherent to complete data collection, not only for the phrased item universe (of vast, undetermined size) but also for the finite universe of trait descriptors.

Despite this, it remains possible (and perhaps paramount) to consider the first of these relationships independently: what is the structure of the universe of phrased items? This question is functionally the same as that which has stimulated decades of psycholexical research on the structure of the universe of trait descriptors (Allport and Odbert, 1936; Cattell, 1943b, 1947; Tupes and Christal, 1961; Digman and Takemoto-Chock, 1981) and would seem to be equally germane given that most modern personality assessment methods use phrased items to the exclusion of trait descriptors. The challenges of evaluating its structure are similar as well. While the universe of phrased items is far less circumscribed than the universe of trait descriptors, it is equally amenable to the aforementioned strategies for item sampling and, in both cases, evaluations of structure are dependent upon the administration of very large item sets to large representative samples. Before proceeding in an attempt to evaluate the structure of the phrased item universe, more information about the resources and strategies for addressing these challenges are described below.

3.1.1 The International Personality Item Pool

A tremendous asset to the field of personality has been the development of the International Personality Item Pool (Goldberg, 1999; Goldberg et al., 2006; Hendriks et al., 1999). The utility of this resource has already been described: a large and growing number of personality measures

have been reformulated into a common framework of items based on joint administration of the IPIP items themselves and the original measures (Goldberg, 1999). Since its inception, more than 250 personality scales have been "created" by use of this procedure (Goldberg, 2014).

The format for the IPIP items has its origins in a set of approximately 900 items constructed by a team of researchers (Hofstee, De Raad, and Hendriks) at the University of Groningen in the 1990s (Goldberg, 1999, 2009; Hendriks, 1997; Hendriks et al., 1999, 2003), and is summarized as follows:

> "Explicit guidelines (Hofstee, 1991) were followed to ensure creating items for an instrument that can be used for a broad range of educational levels, avoids discrimination of certain people or groups of people, and elicits ratings as objectively as possible. These guidelines address the items' format (phrase items in the third person singular and in observable terms) and wording (avoid idiom, difficult words and expressions, suggestive formulations, etc)." (Hendriks et al., 2003, p. 350)

The first-person English translation of these Groningen items, along with another 500 items generated by Goldberg, served as the initial basis for the IPIP item set. The total item count now stands at roughly 2500, each of which has been administered to various (overlapping) subsets of the Eugene-Springfield Community Sample (for more details, see Goldberg, 1999; Goldberg et al., 2006; Goldberg, 2009).

The history of these items is relevant because the theoretical perspective which guided their development was the Abridged Big Five Dimensional Circumplex (Hofstee et al., 1992, the "AB5C"). This development was rooted in psychometric techniques which attempted to integrate the circumplex and simple structure approaches to personality assessment by explicitly evaluating the 10 two-dimensional circumplexes which can be formed between each of the Big Five factors. This goal suggests the inclusion of items in the AB5C scales which occupy the interstitial spaces between the dimensional axes; such items are described as having

greater complexity than those which are located on or near the dimensional axes (Hofmann, 1978). The term complexity is also frequently used to indicate the degree to which factor rotations result in so-called simple structure, whereby items have high primary loadings and small secondary loadings (Pettersson and Turkheimer, 2010; Velicer, 1976).

Further discussion of the IPIP item format and the topic of complexity lies beyond the scope of the current discussion, with one exception. This is to emphasize that the inclusion and maintenance of a broad range of items in the IPIP is a seldom noted and invaluable benefit to the study of personality structure writ large (and the specific studies described below). This is because evaluation of the complexity of personality structure precludes the administration of only those items which have survived analyses designed to produce scales with simple structure.

In terms of the relationship between phrased items and single-word trait descriptors, the IPIP items in Table 3.1 suggest that the content of many phrased items is often *conceptually* complex. This is not universally true of course. Conceptually simple phrased items ("Don't talk a lot") and complex trait descriptors ("aloof") do exist; both the universe of phrased items and the universe of trait descriptors contain items with a range of conceptual complexity. Phrased items have an advantage over trait descriptors in this regard because many conceptually complex trait descriptors are unfamiliar to a large proportion of the population. It should also be noted that *conceptual* complexity is not necessarily related to *psychometric* complexity (in fact, it is a nonsequitur to compare these two properties for any single item).

In sum, unbiased evaluations of personality structure depend on the use of a large set of items, and are improved by the inclusion of items which are conceptually complex while still understandable for most of the individuals sampled. The extant IPIP phrased items comprise such a set, though it is not clear whether they are a *representative*, *uniform*, or *cluster-based* sample of the indefinitely large universe of phrased items.

3.1.2 Synthetic Aperture Personality Assessment

The second obstacle to evaluating the structural similarity between the universe of phrased items and the limited set of trait-descriptors stems from the difficulty of administering very large sets of phrased items to large samples. This has obviously been achieved on at least one occasion (the Eugene-Springfield Community Sample [Goldberg, 1999]), though no other samples of similar size *and* breadth are known to the authors. In fact, in the era of "big data" samples, breadth has become the primary issue. Data collection over the internet has made it considerably easier to increase the number of participants in a sample but it has not changed the fact that individual participants are only willing and able to respond to a limited number of items.

Synthetic Aperture Personality Assessment ("SAPA") techniques represent a methodological innovation for dealing with this obstacle (Evans and Revelle, 2008; Revelle et al., 2010b; Wilt et al., 2011; Condon and Revelle, 2014). So named by analogy to the use of similar techniques in radio and optical astronomy, SAPA essentially makes use of modern sampling procedures. Rather than combining signals from the same source using different telescopes as is done in astronomy, the structure of personality can be studied by combining the responses of many people across more items than any one person is willing to answer. Instead of observing celestial objects beyond the visible range, psychologists can observe the relations between personality constructs which would not otherwise be visible given practical assessment constraints. This can be done by sweeping the assessment "telescope" across a wide range of constructs or by focusing for short periods of time on high-priority topics.

This procedure is not without precedent. Lord (1955) has previously described theoretical procedures for the sampling of items (rather than participants) in the context of testing and similar sampling techniques have long been used by the United States census (Navarro and

Griffin, 2002), the Programme for International Student Assessment (Anderson et al., 2007; OECD, 2012), the German Socioeconomic Panel (Haisken-DeNew and Frick, 2005), and the National Assessment of Educational Progress (Mislevy et al., 1992). The current authors have refined these techniques over the last several years for use with public-domain items administered via the internet (sapa-project.org). In addition to the improved geographical diversity which is common to many internet samples, SAPA procedures allow for evaluation of the structural characteristics of personality items across measures (e.g., various IPIP measures administered together) and domains (e.g., vocational, cognitive ability (Condon and Revelle, 2014), and a broad range of criterion variables such as health behaviors and occupations).

3.1.3 Goals of the current studies

The primary goal of the studies described below was to use SAPA sampling procedures and a large set of public-domain phrased items from the International Personality Item Pool in order to evaluate their structure. Given that the items selected for administration were chosen based on their inclusion among several widely-used sets of personality scales, Study 1 evaluated the extent to which these sets were structurally related to one another. This included evaluation of the relationship between scales with theoretically similar structure (e.g., the IPIP100 and the IPIP-NEO) and between scales at different "levels" (i.e., extraction of various number of factors). Several researchers have suggested that these levels can be interpreted hierarchically (DeYoung, 2006, 2010b; Digman, 1997; Saucier, 2009a) and some of the included scales were intentionally designed to produce dendritically nested scales. Empirical evidence regarding such hierarchies was evaluated in Study 1.

Study 2 evaluated the structure of the full item set across all of the scales administered (92 scales in total). This included exploratory factor analyses with three hypotheses in mind. First, given

that the majority of these items represent those used for five and six factor sets of scales, it was expected that five and six factor solutions would be clearly superior to other alternatives (though no explicit predictions were made as to which of these would fit the data better). Second, based on prior work suggesting larger-than-expected correlations between the Big Five factors in a similar data set (Condon et al., 2013), some evidence for higher and lower order factors was expected. The evidence for lower-order factors was expected to loosely reflect the 10 Big Five Aspect Scales (DeYoung et al., 2007), especially given the inclusion of these scales among the items administered. Expectations regarding the higher order factors were more theoretical, based on the work of several authors (De Raad et al., 2010; DeYoung, 2006, 2010b; Digman, 1997; Eysenck, 1981, 1991; Saucier, 2009a). Specifically, it was postulated that correlations between the five and six factor scales would reflect two or three higher order factors.

Third, it was predicted that the structures based on different factor solutions (i.e., solutions with different numbers of factors extracted) would not be clearly related as these nested structures would be neither necessary nor likely *unless they reflected underlying mechanistic distinctions in human behavior across levels of specificity.* This was not expected given that evidence supporting the existence of such mechanisms has not been reported.

It should be noted that inconclusive evidence of hierarchical nesting at different factoring levels would not necessarily mean that the levels are theoretically independent. A plausible (and non-mutually exclusive) alternative is that personality structure constitutes a heterarchy in which the ranking or importance of structural elements are arranged according to context. In other words, the rankings of structural elements need not be fixed. In some cases, the ranking of structural elements might shift because the most relevant constructs for any particular context require more specificity. Inter-personal interactions serve as a ready example: more specific evaluations of conscientiousness are generally made during a job interview than a first date,

while the opposite might be true for agreeableness. In other cases, it may be that the relationships between constructs at different levels will depend on characteristics such as context or group affiliation. In keeping with the prior example, it may be the case that "talkativeness" is more highly correlated with Agreeableness in the context of first dates than with Extraversion even though the opposite is true over a wide range of contexts.

In order to evaluate the hypothesis regarding empirical evidence of hierarchical nesting, the structural analyses of Study 2 were used to identify scales which capture empirically-supported factor structures at multiple levels, including both broad higher-order factors and more narrow traits.

In order to encourage further analysis and re-analysis by the international community of personality researchers, these data have been made available as part of the Supplementary Materials.

3.2 Study 1

Study 1 evaluated the extent to which eight public-domain sets of scales were structurally related to one another. This included evaluation of the relationship between scales with similar levels of structure (e.g., different Big Five scales) and between scales at different levels, including the prospect of hierarchical organizational structure between scales and constructs.

3.2.1 Method

3.2.1.1 Participants

The data were collected from participants who voluntarily visited the SAPA-Project.org website between December 8, 2013 and July 26, 2014 in exchange for customized feedback about their

personalities. This included 23,681 individuals (64% female) from 172 countries (all items on the SAPA-Project website were posted in English though a small subset of participants (of indeterminate size) made use of browser-based translation software). All data were self-reported. The mean self-reported age was 26.6 years (*sd* = 11.1, median = 22) with a range from 14 to 90 years.

Educational attainment levels for the participants are given in Table 3.2. The largest group of participants were current university school students, though a wide range of educational attainment levels were represented. Race/ethnicity is presented for U.S. participants in Table 3.3; participants from outside the United States were not prompted for information regarding race/ethnicity.

Table 3.2: Participants by educational attainment

	SAPA Sample				U.S.
	% of total	Mean age	Median age	% Female	% of total
Less than 12 years	13.2	17.3	17	62.3	14.8
High school graduate	9.3	22.7	18	57.0	28.5
Currently in college/university	41.1	24.3	21	68.5	NA*
Some college/university, but did not graduate	6.1	33.6	30	57.2	21.4*
College/university degree	15.1	33.8	31	62.6	25.1
Currently in graduate or professional school	5.4	29.4	26	65.3	NA
Graduate or professional school degree	9.8	39.5	37	60.1	10.3

U.S. data from the 2009 American Community Sample of the U.S. Census Bureau (Bureau, 2012).
* ACS data do not differentiate between those who are active students and those who are no longer enrolled.

Figure 3.1 shows the geographic distribution of participants from the continental United States who provided optional ZIP code information (97.7% of U.S. participants). The correlation between ZIP code distributions in the sample and the U.S. population based on U.S. Census data (Census, 2011) was 0.7 when using the broader 3 digit ZIP codes (known as the regional prefixes, of which there are approximately 890). Note that these calculations required matching of ZIP codes to the U.S. Census Bureau's Zip Code Tabulation Areas.

Table 3.3: Participants by race/ethnicity

	SAPA Sample		
	Count	%	% of U.S.*
African-American	1,329	9.1	12.2
Asian-American	775	5.3	4.4
Hispanic-American	1,284	8.8	15.7
Native-American	137	0.9	0.9
White/Caucasian	8,291	56.9	64.9
Multi-ethnic	809	5.6	1.7
Other	185	1.3	NA
Not specified	1,763	12.1	NA

* 2009 U.S. data from the U. S. Census Bureau (2012)

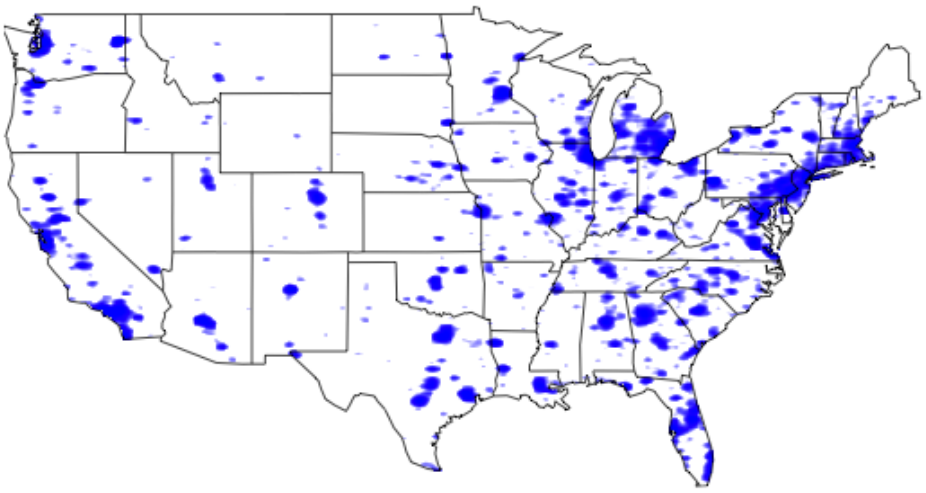


Figure 3.1: Participants by ZIP code for the continental United States

3.2.1.2 Measures

Eight sets of self-report personality scales were administered. Seven of these are based on items from the International Personality Item Pool: the 100 IPIP items corresponding to the Big Five factor markers (Goldberg, 1999), the 100 items of the Big Five Aspect Scales (DeYoung et al., 2007), the 240 items of the IPIP-HEXACO inventory (Ashton et al., 2007), the 48 items of the Questionnaire Big Six scales (Thalmayer et al., 2011), the 300 items of the IPIP-NEO (Goldberg, 1999), the 127 items of the IPIP-Multidimensional Personality Questionnaire ("MPQ" Goldberg, 2014; Tellegen and Waller, 2008), and the 40 items of the Plasticity/Stability scales (DeYoung, 2010b). The eighth set of scales was the 79 items of the Eysenck Personality Questionnaire - Revised (Eysenck et al., 1985). Note that the format of these items was modified to match that of the IPIP items and that the 21 "lie" scale items were intentionally omitted. Administration of these scales also implies the administration of several other measures which are abbreviations of these scales, including the 24 and 36 item Questionnaire Big Six scales (Thalmayer et al., 2011), the 50 item IPIP scales corresponding to the Big Five factor markers (Goldberg, 2014), and the 20 item "mini-IPIP" scales (Donnellan et al., 2006). None of these shorter scales were directly evaluated here.

The 1,034 items from these measures contain 338 duplicates, resulting in a total set of 696 unique items. Of these, 473 items are in only one set of scales, 126 items are included in two sets of scales, 54 items are in three, 22 items in four, 17 items in five, and 4 items are in six of the seven sets of IPIP-based scales ("Have little to say", "Worry about things", "Like order", and "Have a rich vocabulary"). All of the items were administered with the same six response options ("Very Inaccurate", "Moderately Inaccurate", "Slightly Inaccurate", "Slightly Accurate", "Moderately Accurate", "Very Accurate").

The items were administered using the Synthetic Aperture Personality Assessment ("SAPA")

technique (Revelle et al., 2010b), a variant of matrix sampling procedures discussed by Lord (1955). This method produces data which contain "massive missingness" by design (Revelle and Brown, 2013). This missingness qualifies for classification as missing completely at random ("MCAR", Graham, 2009) and it is further described as massively missing because the mean level of missingness by participant was approximately 77%. The items were presented to participants in random order, and participants responded to as many items as they wished. The mean number of items to which participants responded was 86.1 (sd = 58.7; median = 71). The number of items administered to each participant was procedurally independent of participant response characteristics. The number of administrations for each item varied considerably (median = 2554; m = 2931; sd = 781) as did the number of pairwise administrations between any two items in the set (median = 519; m = 528; sd = 117). The minimum number of pairwise administrations among items (281) provided sufficiently high stability in the covariance matrix for the structural analyses described below (Kenny, 2012; Schönbrodt and Perugini, 2013).

3.2.1.3 Analyses

All analyses were conducted using the *psych* package (Revelle, 2014) in R (R Core Team, 2014). The primary method of analysis was simple correlation among scales however the presence of overlapping items among the scales precipitated the need to correct for item overlap (Bashaw and Anderson, 1967) using the score overlap function.

3.2.2 Results

Distributions of the standard errors of the correlations between items and scales are shown in Figures 3.2 and 3.3, respectively. With few exceptions, the standard errors between items were generally below 0.055 and the standard errors of the correlations between scales were less than

0.03. Given the latter result, no explicit discussion is made below of the statistical significance of the correlations between scales (statistical significance values are available in the data set provided as part of the Supplementary Materials).

The correlations among all 8 sets of scales are presented in Figure 3.4, with the scale correlations corrected for item overlap below the diagonal, alpha values presented on the diagonal and scale correlations corrected for item overlap and attenuation above the diagonal. Evidence for construct congruence across the sets of scales can be seen by the diagonal striations of shading in both the lower and upper halves of the correlation matrix (above and below the primary diagonal), though congruence was more evident for some constructs than others. Across the EPQ, Big Five scales (IPIP100 and NEO), Big Six scales (IPIP-HEXACO and QB6), and, to a lesser extent, the BFAS, the most correlated constructs were Neuroticism/Emotional Stability (0.75-0.86) and Extraversion (0.69-0.92). Among the remaining Big Five factors, correlations across sets of scales were also high for Conscientiousness, though its relationship to the Big Three was as part of Psychoticism. Openness and Agreeableness were less highly correlated across the sets of scales, suggesting that these factors are framed less consistently across measures. In the case of Openness, this was expected as it is denoted by different labels, for example, in the IPIP100 (Intellect) and the NEO (Openness to New Experiences). In the case of Agreeableness, the lower correlations were particularly evident between Big Five and Big Six measures, in large part due to overlap with the Honesty/Humility construct.

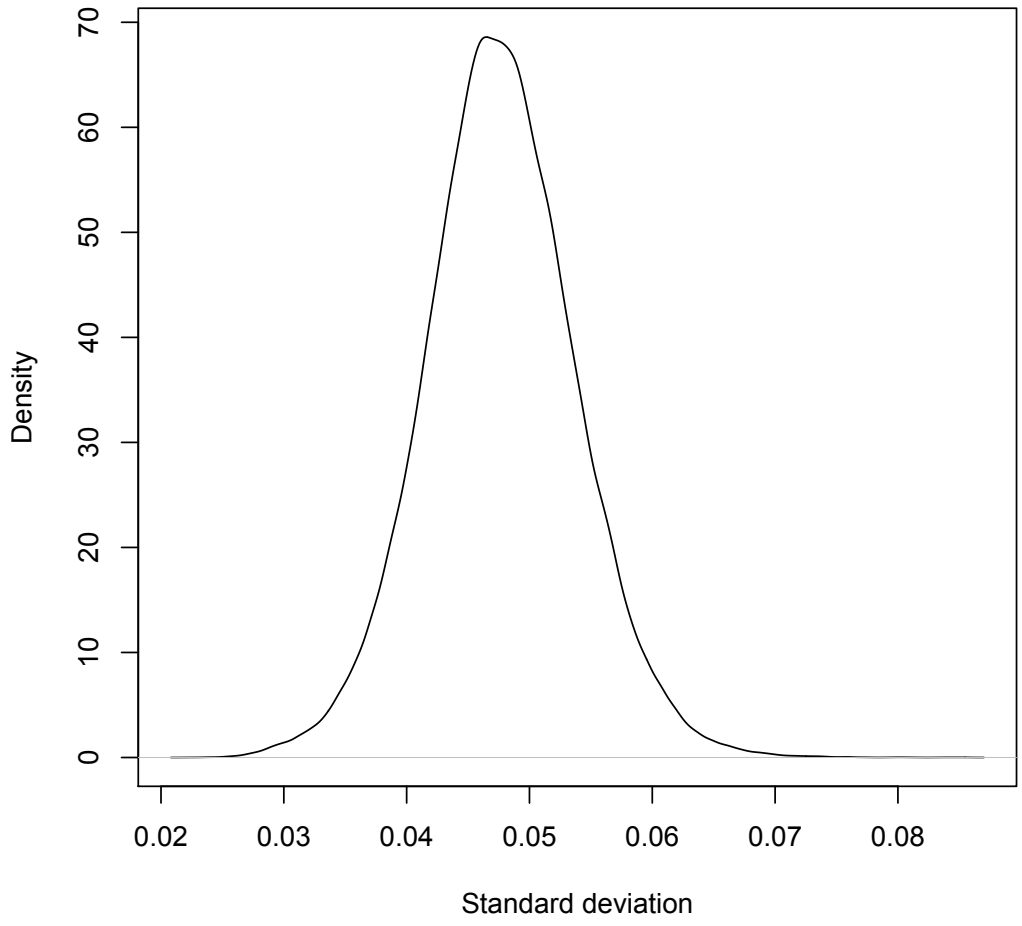


Figure 3.2: Standard errors of the item correlations for 696 items

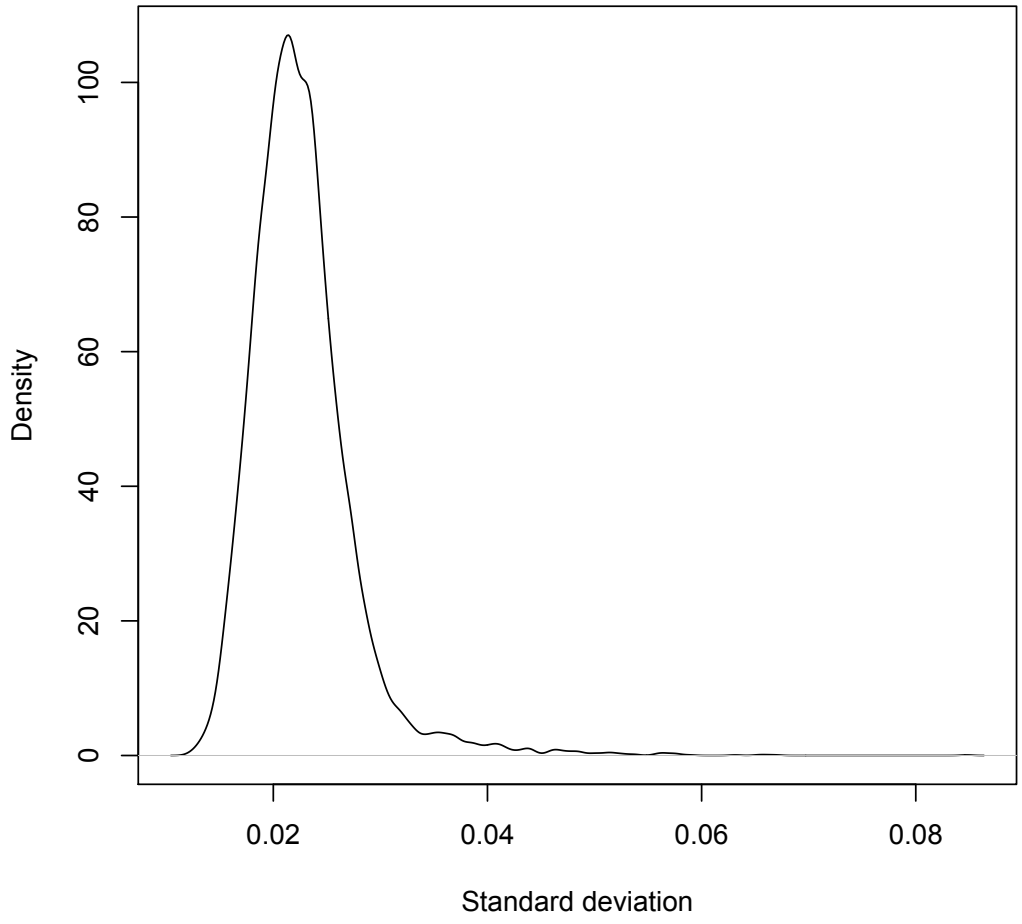


Figure 3.3: Standard errors of the scale correlations

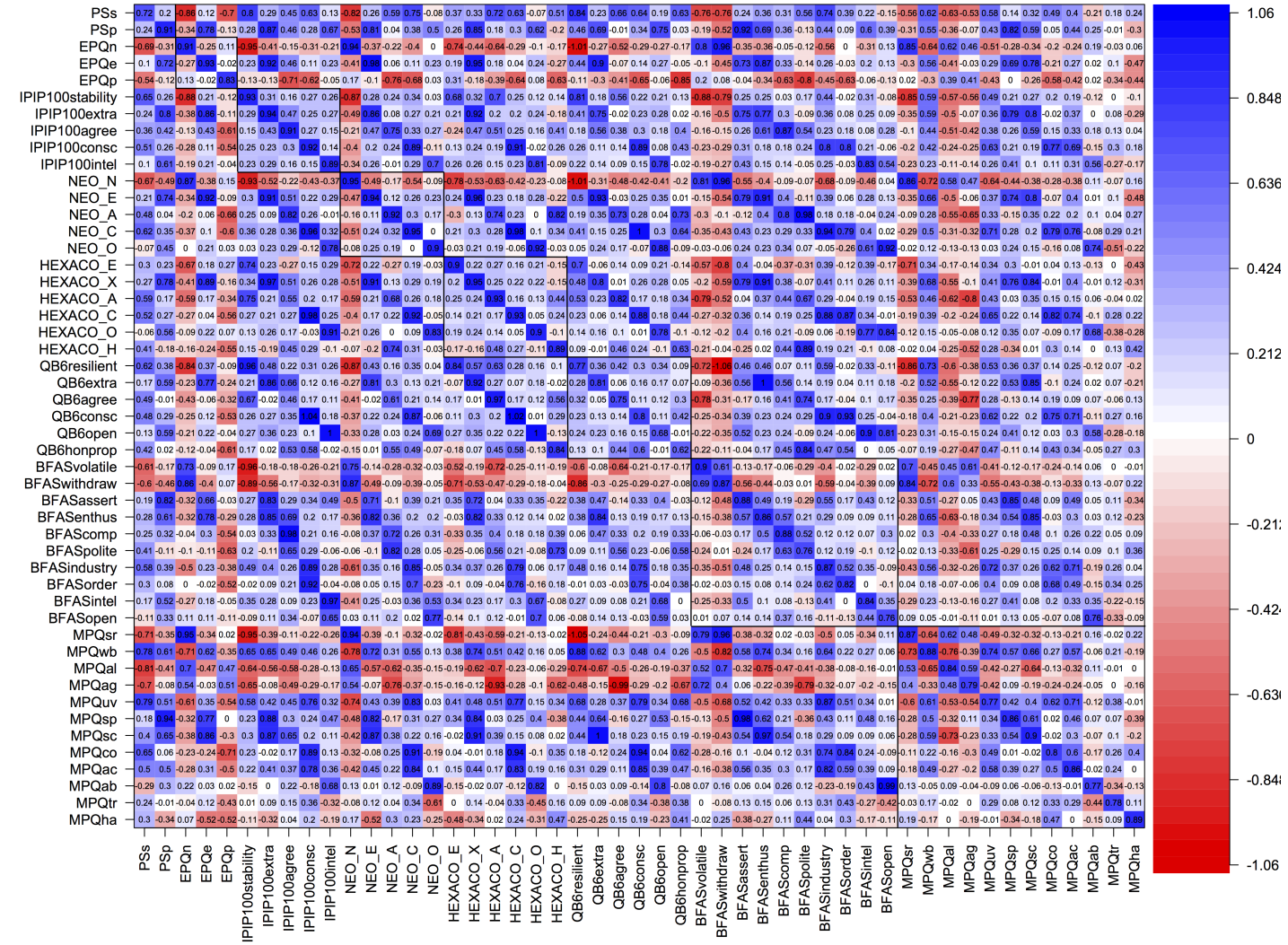


Figure 3.4: Correlations corrected for item overlap; alphas on the diagonal and correlations corrected for attenuation above the diagonal

Evaluation of the correlations above a threshold of 0.3 between specific sets of scales can be seen in Figures 3.5 to 3.11. Figure 3.5 shows the correlations between the 100 items of the five factor IPIP100 and the 300 items of the five factor IPIP-NEO. Across measures, similarly named scales demonstrated correlations ranging from 0.70 for IPIP-NEO Openness and IPIP100 Intellect to 0.89 for Conscientiousness. Several cross-loadings above 0.30 were also evident including those between: IPIP100 Agreeableness and IPIP-NEO Conscientiousness (0.33); IPIP100 Intellect and IPIP-NEO Neuroticism (-0.34); IPIP100 Stability and IPIP-NEO Conscientiousness (0.34); IPIP100 Conscientiousness and IPIP-NEO Neuroticism (-0.4); IPIP100 Agreeableness and IPIP-NEO Extraversion (0.47); and IPIP100 Extraversion and IPIP-NEO Neuroticism (-0.49).

Figure 3.6 shows the correlations between the 240 items of the IPIP-HEXACO and the 48 items of the QB6. Comparable scales demonstrated correlations ranging from 0.63 for IPIP-HEXACO Honesty/Humility and QB6 Honesty/Propriety to 0.88 for Conscientiousness. Several cross-loadings above 0.30 were also evident among these measures including those between: IPIP-HEXACO Agreeableness and QB6 Honesty/Propriety (0.34); IPIP-HEXACO Conscientiousness and QB6 Honesty/Propriety (0.44); IPIP-HEXACO Honesty/Humility and QB6 Agreeableness (0.46); IPIP-HEXACO Extraversion and QB6 Resiliency (0.48); and IPIP-HEXACO Agreeableness and QB6 Resiliency (0.53).

Figure 3.7 shows the correlations between both of the five factor measures with both of the six factor measures. The number of cross-loadings above 0.30 was fewer between the IPIP100 and the IPIP-HEXACO (4) than between the IPIP-NEO and the IPIP-HEXACO (5). Three of these cross-loadings were above 0.50 in magnitude for the IPIP-NEO. Similar circumstances were evident with regards to the QB6 (5 cross-loadings above 0.3 with the IPIP100 and 7 with the IPIP-NEO).

Figures 3.8 and 3.9 demonstrate the extent of cross-loadings among the hierarchically organized

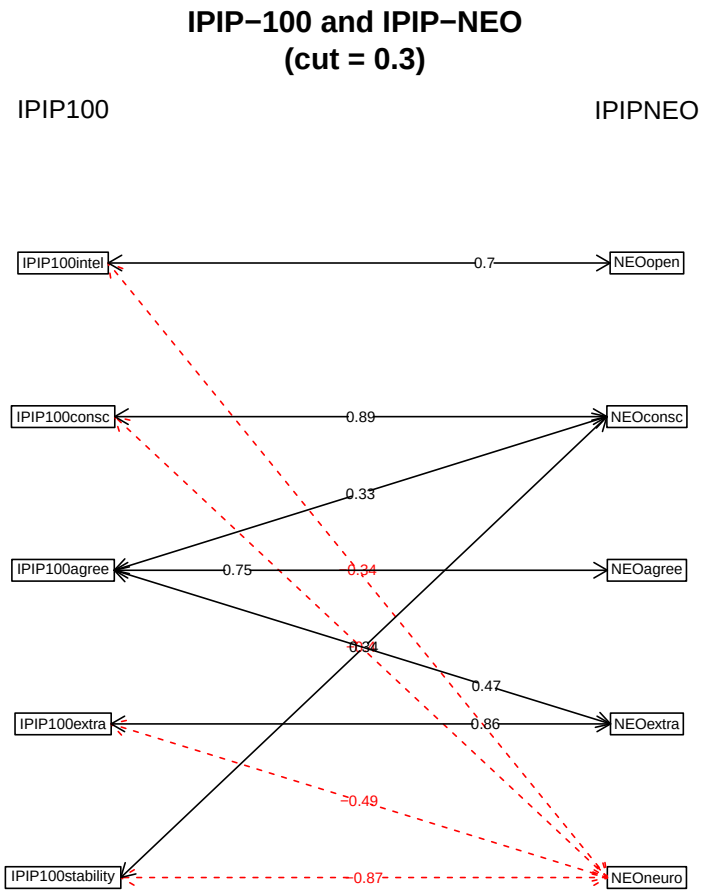


Figure 3.5: Scale correlations corrected for item overlap

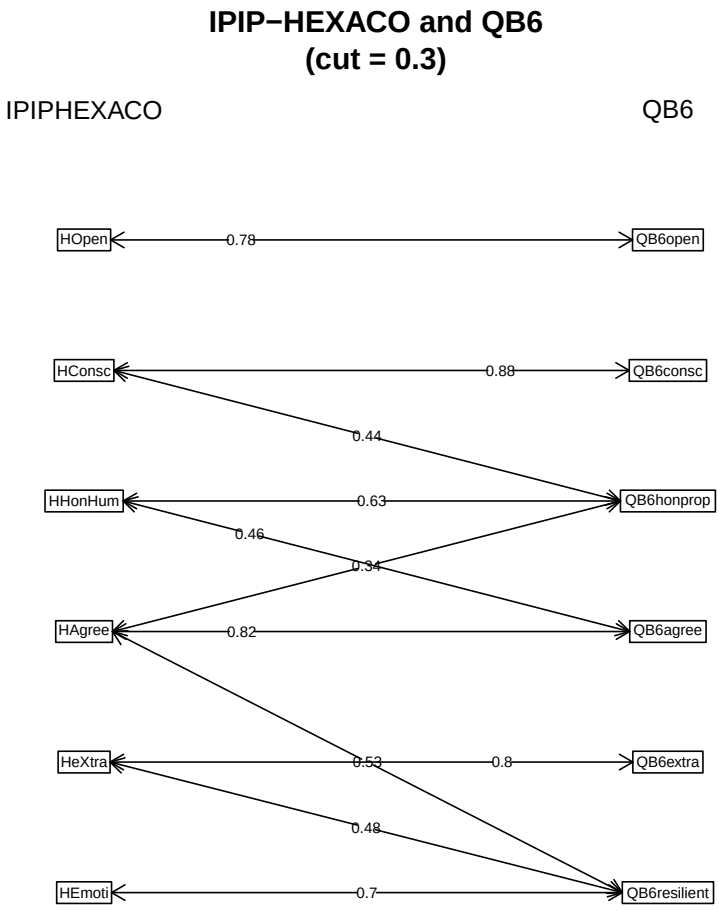


Figure 3.6: Scale correlations corrected for item overlap

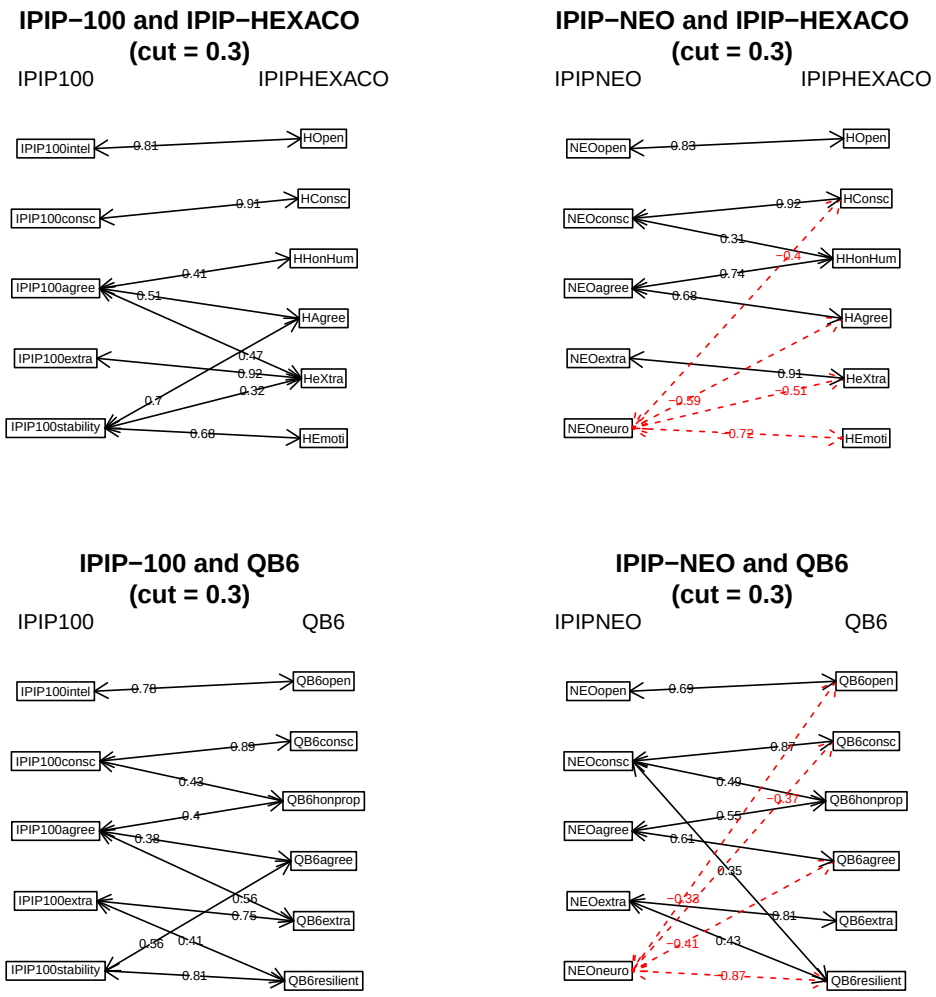


Figure 3.7: Scale correlations corrected for item overlap

scales for the IPIP-NEO and the IPIP-HEXACO. For both measures, the highest correlations are the lower level "facet" scales beneath each of the primary factor level scales, though many significant cross-loadings are also evident, particularly for the IPIP-NEO.

Figures 3.10 and 3.11 demonstrate the hierarchical relationships across multiple levels. In the first case, Figure 3.10, scale correlations are shown using 10, 6, 5, 3, and 2 factor measures. None of the levels clearly demonstrated a nested hierarchy to the level directly above or below, and, in particular, the relationship between 6 and 10 factors has a large number of cross-loadings above 0.3. The evidence for hierarchy among the 10, 5 and 2 factor scales is more clear in Figure 3.11. In this case, each of the five factors is most highly correlated with each of its two hypothesized aspects at the 10 factor level and each of the factors at the two factor level are similarly well-represented by two of the Big Five (Plasticity with Extraversion and Intellect; Stability with Emotional Stability and Conscientiousness), with both factors correlating similarly with Agreeableness.

3.2.3 Discussion

The general finding from Study 1 is that widely-used constructs are not similarly framed across measures. There are many examples of this, even among measures with the same number of scales. Most notable at the five factor level were strong correlations between IPIP100 Extraversion and IPIP-NEO Neuroticism (-0.49) and the IPIP100 Agreeableness and IPIP-NEO Extraversion (0.47). On the six factor level, strong correlations included IPIP-HEXACO Extraversion and QB6 Emotionality (0.48) and IPIP-HEXACO Agreeableness and QB6 Emotionality (0.53). When considering relationships between measures at the five and six factor levels, it appears that the IPIP100 has a cleaner relationship with both of the Big Six measures than the IPIP-NEO.

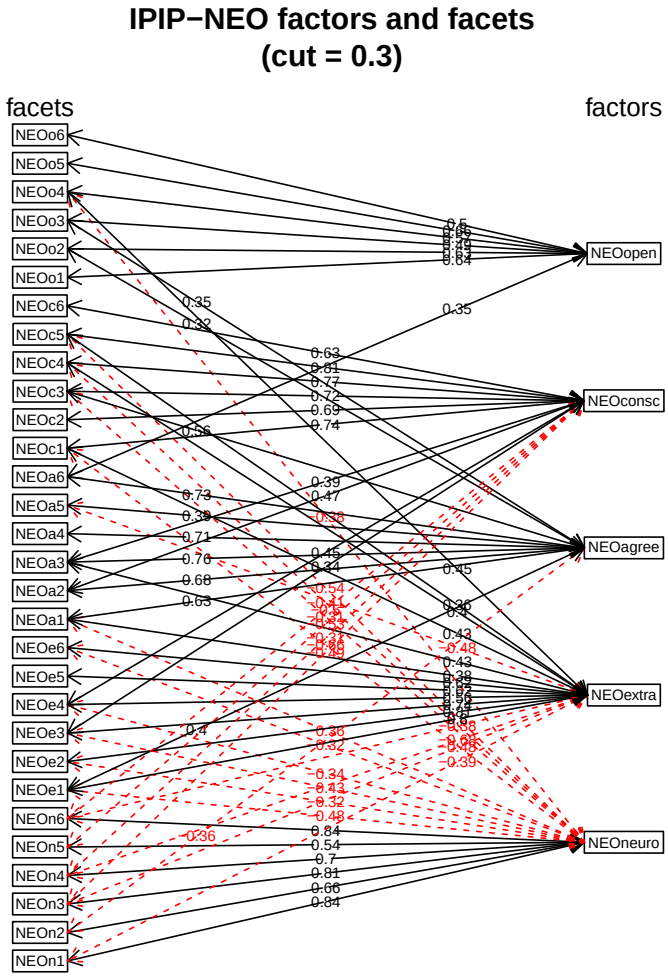


Figure 3.8: Scale correlations corrected for item overlap

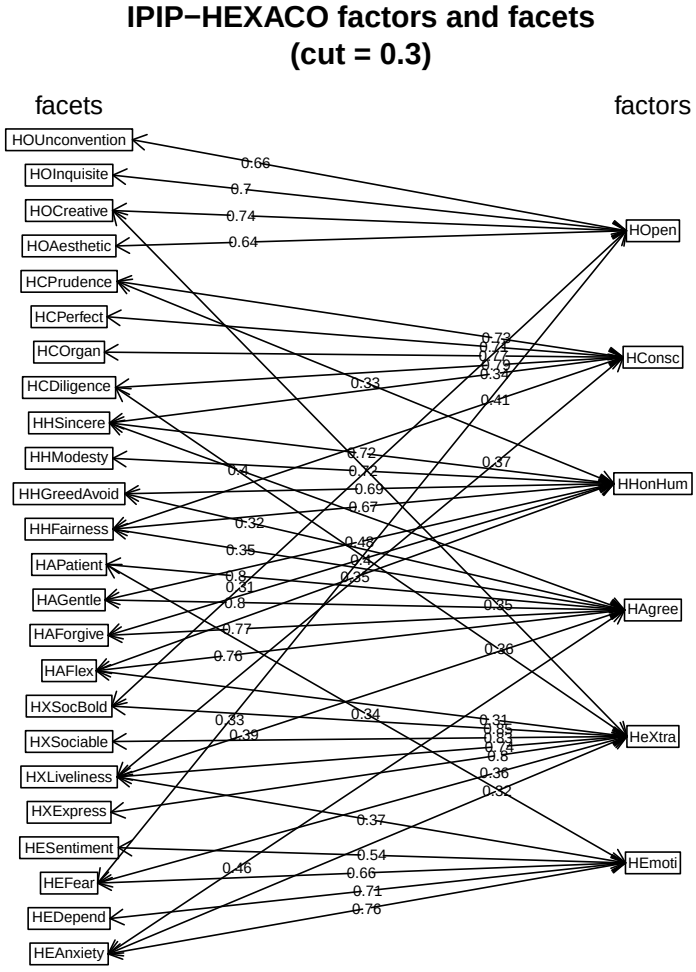


Figure 3.9: Scale correlations corrected for item overlap

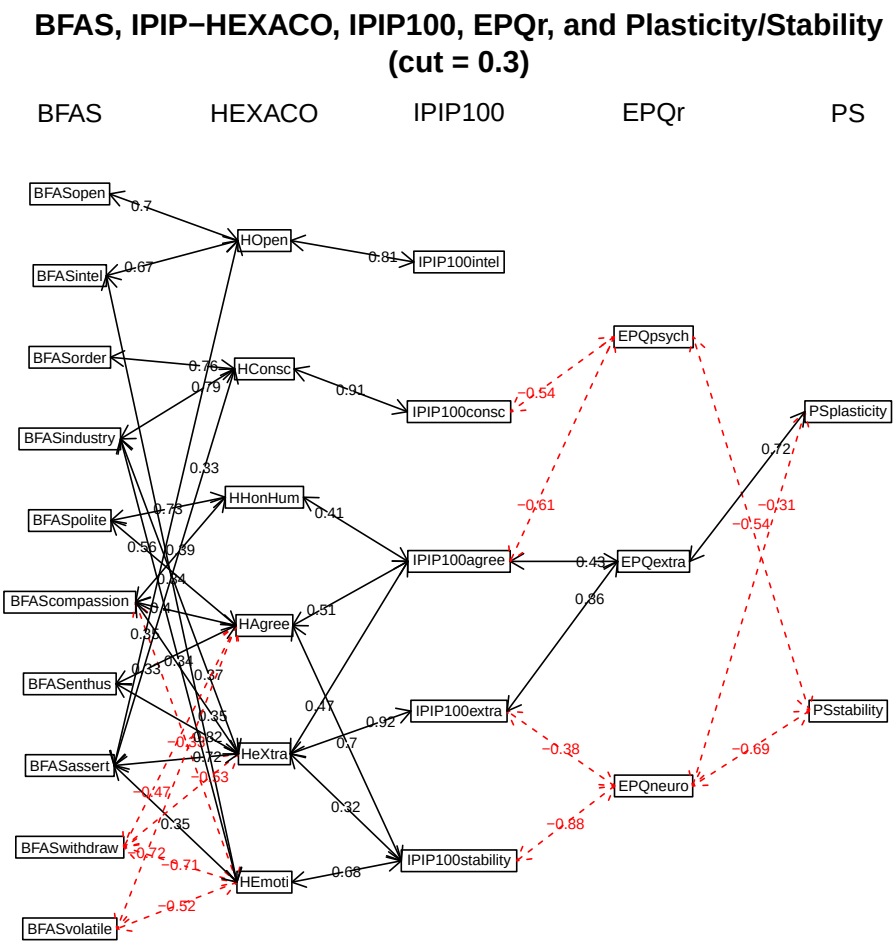


Figure 3.10: Scale correlations between the existing 2, 3, 5, 6, and 10 factor scales corrected for item overlap

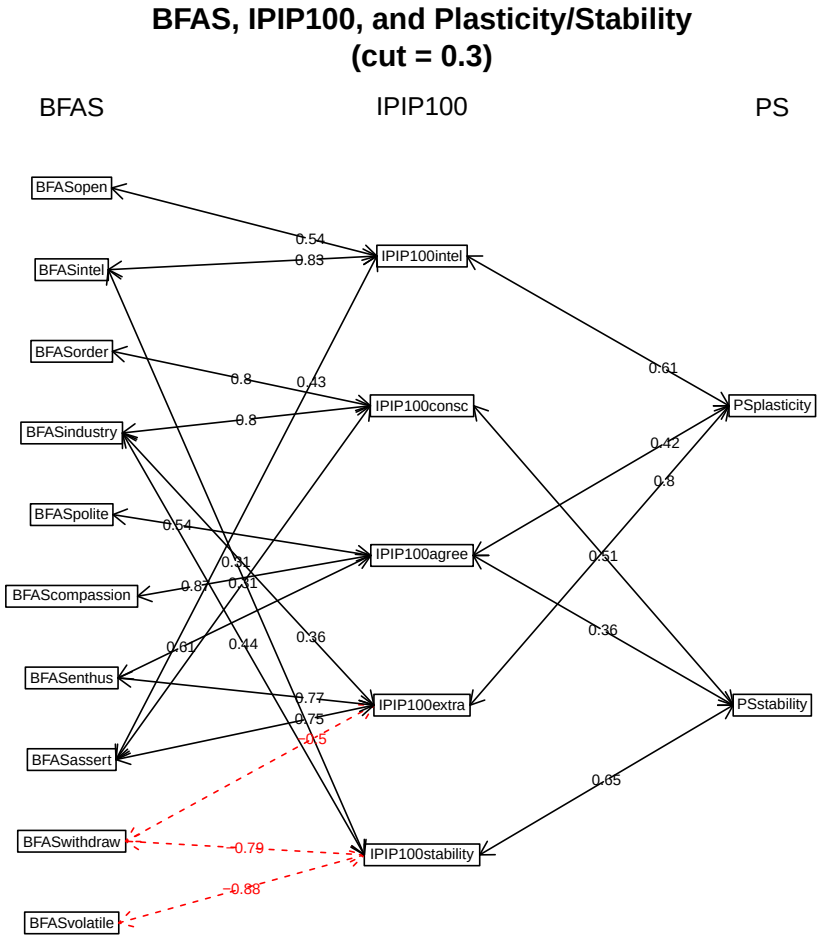


Figure 3.11: Scale correlations corrected for item overlap

With regards to the evidence for hierarchy among different levels, the results were mixed. Both the IPIP-HEXACO and the IPIP-NEO performed as designed in terms of the relationships between the higher level factors and their lower order facets however the large number of cross-loadings among the facets and factors suggests considerable structural complexity. The relationships between the BFAS, IPIP100 and Plasticity/Stability scales demonstrated reasonable nesting across levels, though it is important to acknowledge that the 2 and 10 factor levels were explicitly designed based on the IPIP100. In other words, these results serve essentially as replication of their original design procedures rather than evidence for hierarchy in personality per se.

3.3 Study 2

Study 2 made use of exploratory factor analyses to evaluate the structure of the full set of items administered as part of the eight sets of scales evaluated in Study 1. Three empirical questions were considered. First, it was expected that the five and six factor solutions would provide a superior fit to the data relative to factor extraction at other levels. This hypothesis was mainly driven by the fact that the majority of these items were taken from sets of scales with five or six factors, but also was secondarily motivated by the general consensus that five (and, to a lesser extent, six) factors is optimal. Second, evidence for less well-fitting but clearly identifiable factors at higher and lower levels of extraction was expected. This was, in some sense, inevitable but it was generally expected that these factors would match those which have previously been identified in the literature (e.g., Eysenck's Giant Three and the 10 aspects of the BFAS). Finally, though no *a priori* predictions were made, evaluations were made to consider the extent to which factor structures at different levels were hierarchically nested.

In addition to these empirical questions, attempts were made to identify scales which which

would allow for assessment at the levels of extraction which provide clear and reasonable fits for the full correlational structure. To the extent possible, these scales were meant to draw upon the same items so as to allow for multi-level assessment with as few items as possible.

3.3.1 Method

The participants and measures were the same as those used in Study 1. All analyses were again conducted using the *psych* package (Revelle, 2014) in R (R Core Team, 2014). Latent variable exploratory factor analyses ("EFA") were conducted on all items in order to evaluate the fits for factor solutions based on the extraction of 1 to 30 factors. The EFA results reported below were based on the Pearson correlations between scored responses using Ordinary Least Squares ("OLS") regression models with oblique rotation (Revelle, 2014). Variations on these factor analytic methods are demonstrated in the analytic summary included as part of the Supplementary Materials.

Goodness-of-fit was evaluated using the Minimum Average Partial criterion ("MAP", Velicer, 1976), the Root Mean Squared Error of Approximation ("RMSEA", Hu and Bentler, 1999), the Standardized Root Mean Square of the Residual ("SRMR" Hu and Bentler, 1999), an empirically-derived measure of the Bayesian Information Criterion ("eBIC" Schwarz, 1978; Revelle, 2014), and an index of complexity (Hofmann, 1977, 1978). For all of these fit statistics, lower values indicate a superior fit, though the MAP and BIC will often indicate a localized minimum while the RMSEA and SRMR values will decrease as more factors are extracted. Good fits are typically indicated by RMSEA values of less than 0.05 and SRMR values of less than 0.08 (Kenny, 2012). The complexity reported for a given factor solution reflects the mean of the item-level complexities. It should be noted that the complexity for any single item is not sensitive to the magnitude of factor loadings but rather the degree of similarity in loading

magnitudes. Given that communality is the sum of the squared loadings on all factors, item complexity should be considered in conjunction with its communality.

In some cases, the correlation matrices used for these factor analyses were not invertible. This means that at least one of the eigenvalues for the matrix was negative; matrices of this type are also sometimes described as being "not positive definite" (Rigdon, 1997). While this issue occurs more frequently with polychoric correlation matrices, Pearson correlation matrices can also qualify as "not positive definite" under certain circumstances, despite being based on large samples (Rigdon, 1997; Wothke, 1993). The issue of positive definiteness – and the resulting inability to invert the matrix – is, itself, related to the sign of the determinant of the matrix. In order to be positive definite, a matrix and all of its principal submatrices must have positive determinants. If this is not the case, the matrix will be "not positive definite" and not invertible. If the determinant is zero, the matrix is said to be singular. (It is worth noting that factor analyses in personality research commonly result in the generation of "warning" and "error" messages in many statistical software packages as a result of this phenomenon, as they should. The *psych* package is no exception.) While the cause of this issue can be difficult to identify, linear dependency was suspected in this case as many of the items across measures, and even within measures, contained highly similar content. For example, three of the items used to assess humility are "Consider myself an average person", "See myself as an average person", and "Am just an ordinary person." Evidence for linear dependency as a contributing cause of non-positive definiteness was evaluated, where relevant, by attempting to identify items with highly similar loadings across factors as this suggested that these items were occupying nearly identical locations in the factor space. Note that linear dependency can occur even in the absence of very high ($>.90$) correlations between any two items as it can also result from moderately high correlations between *sets* of related items.

Following EFA for the full item set, scales were created by identifying items with high loadings on each factor at the most viable levels of extraction. Where possible, scales were made using items which were among the highest loading items at more than one level though this was not always possible.

3.3.2 Results

Results from exploratory factor analyses of all items were inconclusive. Table 3.4 shows the fit statistics based on the extraction of 1 to 30 factors. SRMR suggested that anything more than 2 factors provided good fit. Using a threshold of 0.05, RMSEA values suggested a good fit at approximately 27 factors and this corresponded with the best fit solution based on the minimum values for the empirical BIC. The MAP criterion reached a minimum value at 9 factors. Complexity values (see also Figure 3.12) did not indicate a level of substantially improved complexity though the 26, 5, and perhaps 15 factor solutions were slightly superior to other options.

Table 3.5 describes the content of each factor, where identifiable, for the factor solutions with 1 to 15 factors of extraction, with the factors in each case sorted by the magnitude of the eigenvalues. In general, the content of the factors was consistent from one level of extraction to the next though there were several notable exceptions due to occasional "re-orientation" of the factors. For example, sociability was present in nearly every factor solution though there were several levels of extraction (4, 5, and 8) where this factor required a more broad interpretation (referred to as Extraversion) in that it included items related to enthusiasm and high-energy. Impulsivity, which emerged as a clear second factor in the two factor solution, was not clearly identifiable in any of the subsequent solutions until 15 factors were extracted.

Table 3.4: Fit statistics for extraction of 1 to 30 factors from 696 items

Factors	MAP	RMSEA	SRMR	eBIC	complexity
1	0.0169	0.161	0.113	1,589,388	1.0
2	0.0132	0.131	0.093	564,102	1.4
3	0.0104	0.108	0.077	-58,774	1.7
4	0.0085	0.091	0.066	-439,854	1.9
5	0.0074	0.078	0.058	-684,312	2.0
6	0.0069	0.071	0.053	-795,916	2.3
7	0.0067	0.067	0.051	-847,205	2.4
8	0.0066	0.065	0.050	-871,738	2.6
9	**0.0066**	0.064	0.049	-892,027	2.8
10	0.0066	0.062	0.048	-907,511	3.0
11	0.0066	0.061	0.047	-918,480	3.1
12	0.0066	0.060	0.047	-927,418	3.2
13	0.0066	0.059	0.046	-935,854	3.4
14	0.0067	0.058	0.046	-942,399	3.5
15	0.0067	0.057	0.045	-948,740	3.7
16	0.0068	0.056	0.045	-954,363	3.8
17	0.0068	0.055	0.044	-958,845	3.9
18	0.0069	0.055	0.044	-962,201	4.0
19	0.0070	0.054	0.044	-964,981	4.2
20	0.0070	0.053	0.043	-966,675	4.2
21	0.0071	0.053	0.043	-968,057	4.4
22	0.0072	0.052	0.043	-969,093	4.6
23	0.0073	0.052	0.043	-969,806	4.7
24	0.0073	0.052	0.042	-970,433	4.8
25	0.0074	0.051	0.042	-970,864	4.8
26	0.0075	0.051	0.042	-971,276	4.8
27	0.0076	0.050	0.042	**-971,277**	5.1
28	0.0077	0.050	0.041	-971,125	5.2
29	0.0077	0.050	0.041	-970,844	5.3
30	0.0078	0.049	0.041	-970,573	5.4

Fits are based on the “minres” factoring method and “oblimin” rotation procedures using the ‘nfactors’ function in the *psych* package in R. Minimum values are bolded for eBIC and MAP.

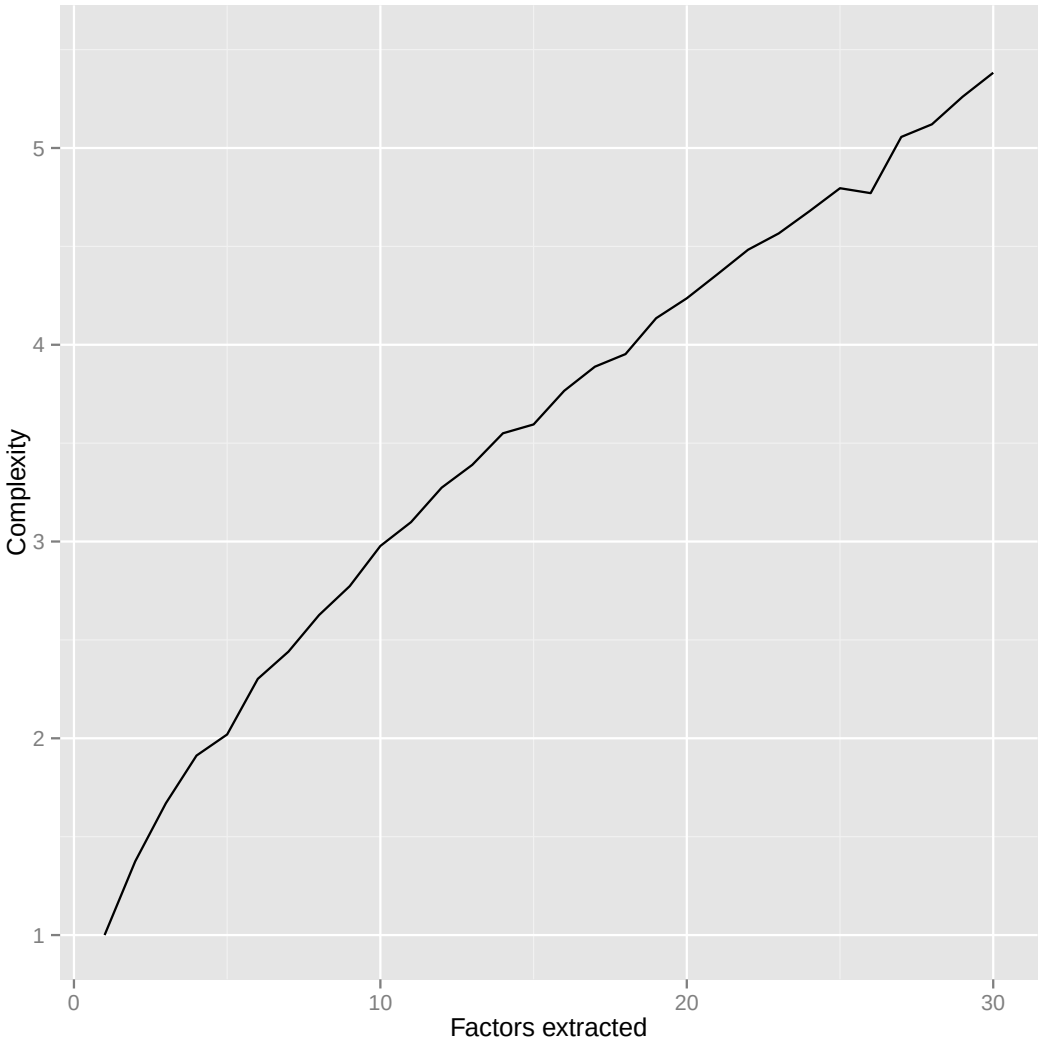


Figure 3.12: Complexity based on extraction of 1 to 30 factors from 696 items

Table 3.5: Item content of extracted factors from 1 through 15

Extracted	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
1	outlook/ sociability														
eigen	88.34														
2	sociability	impulsivity													
eigen	73.61	61.09													
3	emotionality	sociability	social cohesion												
eigen	64.84	60.06	46.91												
4	extraversion	neuroticism	consc	psychopathy											
eigen	60.83	52.41	45.27	41.88											
5	extraversion	neuroticism	consc	compassion /honesty /humility	intellect /openness										
eigen	56.81	51.04	44.53	41.07	30.10										
6	sociability	neuroticism	consc	agreeable	intellect	*unclear*									
eigen	55.14	50.70	43.49	41.15	30.45	18.51									
7	sociability	neuroticism	industrious	humility/ agreeable /honesty	intellect	boldness	trusting/ dependent								
eigen	51.64	46.99	43.77	34.92	27.13	25.99	20.17								
8	extraversion	consc	outlook	humility	neuroticism	intellect	boldness	trusting/ dependent							
eigen	46.06	39.12	36.87	34.26	29.12	27.13	24.50	21.77							
9	sociability	industrious	humility/ compassion	outlook	neuroticism	intellect	boldness	self-esteem	conventional						
eigen	43.98	37.13	34.25	33.94	28.89	26.27	24.79	22.01	15.16						
10	sociability	industrious	outlook	compassion /humility	neuroticism	intellect	boldness	trusting/ dependent	*unclear*	apathy /lie					
eigen	44.20	35.45	35.35	30.80	28.72	25.72	24.63	17.87	15.98	14.65					
11	sociability	compassion	industrious	outlook	neuroticism	boldness	openness	intellect	self-esteem	trusting	*unclear*				
eigen	41.31	34.13	33.57	33.24	27.63	23.34	21.63	21.19	18.12	13.54	11.84				
12	sociability	outlook	industrious	neuroticism	machiavel	boldness	intellect	openness	conventional	compassion	trusting	oddity			
eigen	40.52	34.12	29.95	27.31	25.47	21.42	21.13	20.94	18.34	16.18	15.51	14.40			
13	sociability	outlook	industrious	neuroticism	compassion /empathy	machiavel	openness	boldness	conventional /intellect	confidence	trusting	oddity /lie	easy-going		
eigen	39.81	34.35	29.31	27.17	24.78	21.99	20.13	20.92	16.55	19.14	13.62	12.11	11.03		
14	sociability	outlook	industrious	stability	compassion	machiavel	intellect	openness	boldness	convention	trusting	laugh	easy-going	out-spoken	
eigen	39.26	32.15	29.36	26.00	24.88	21.75	20.05	19.77	19.64	14.75	13.43	12.74	10.80	11.61	
15	sociability	enthusiasm	industrious	compassion	volatility	openness	machiavel	intellect	boldness	convention	fear /worry	impulsive	trusting	easy-going	serious
eigen	38.64	31.85	28.26	25.43	23.63	20.35	20.25	18.72	17.47	15.39	14.33	14.62	11.84	10.32	10.32

Based on visual inspection of the scales and the clarity of the identified factors, scales were developed at 3 levels of extraction: 3, 5 and 15 factors. It should be acknowledged that the fit statistics did not necessarily support these 3 levels to the exclusion of other options. Attempts to develop scales at additional levels of extraction (including levels with 4 factors and 6 through 14 factors) produced scales with low internal consistencies, few items with high loadings, and/or inconsistent item content. Other levels, including those with more than 15 factors, might also provide utility but were deemed less clear and/or less practical.

The scales described here are collectively referred to below as the SAPA Personality Inventory ("SPI") and are individually denoted by the labels SPI-3, SPI-5 and SPI-15. The steps used to develop these scales were as follows: (1) identification of the top 25 items by loading for each of the factors in the 15 factor solution (note that some of the factors did not have 25 items); (2) identification of the top 40 items by loading for each of the factors in the 5 factor solution; (3) identification of the top 100 items by loading for each of the factors in the 3 factor solution; (4) reduction of each of the 15 factor sets down to 8 items each, giving preference to items that have high primary loadings, low secondary loadings, and are also present among the item sets identified at the level of 5 and 3 factors; (5) reduction of the 5 factor sets down to 15 items each, giving preference to items that have high primary loadings, low secondary loadings, and are also present among the item sets identified at the level of 15 and 3 factors; (6) finally, reduction of the 3 factor sets down to 20 items each, giving preference to items that have high primary loadings, low secondary loadings, and are also present among the item sets identified at the level of 15 and 5 factors.

The correlations among these scales are presented in Figure 3.13, with corrections for item overlap below the diagonal, alphas on the diagonal, and corrections for item overlap and attenuation above the diagonal. Internal consistencies for the scales at all three levels were high except for SPI-15 Easy-Goingness ($\alpha = 0.69$, 8 items) and SPI-15 Seriousness ($\alpha = 0.57$, 8 items).

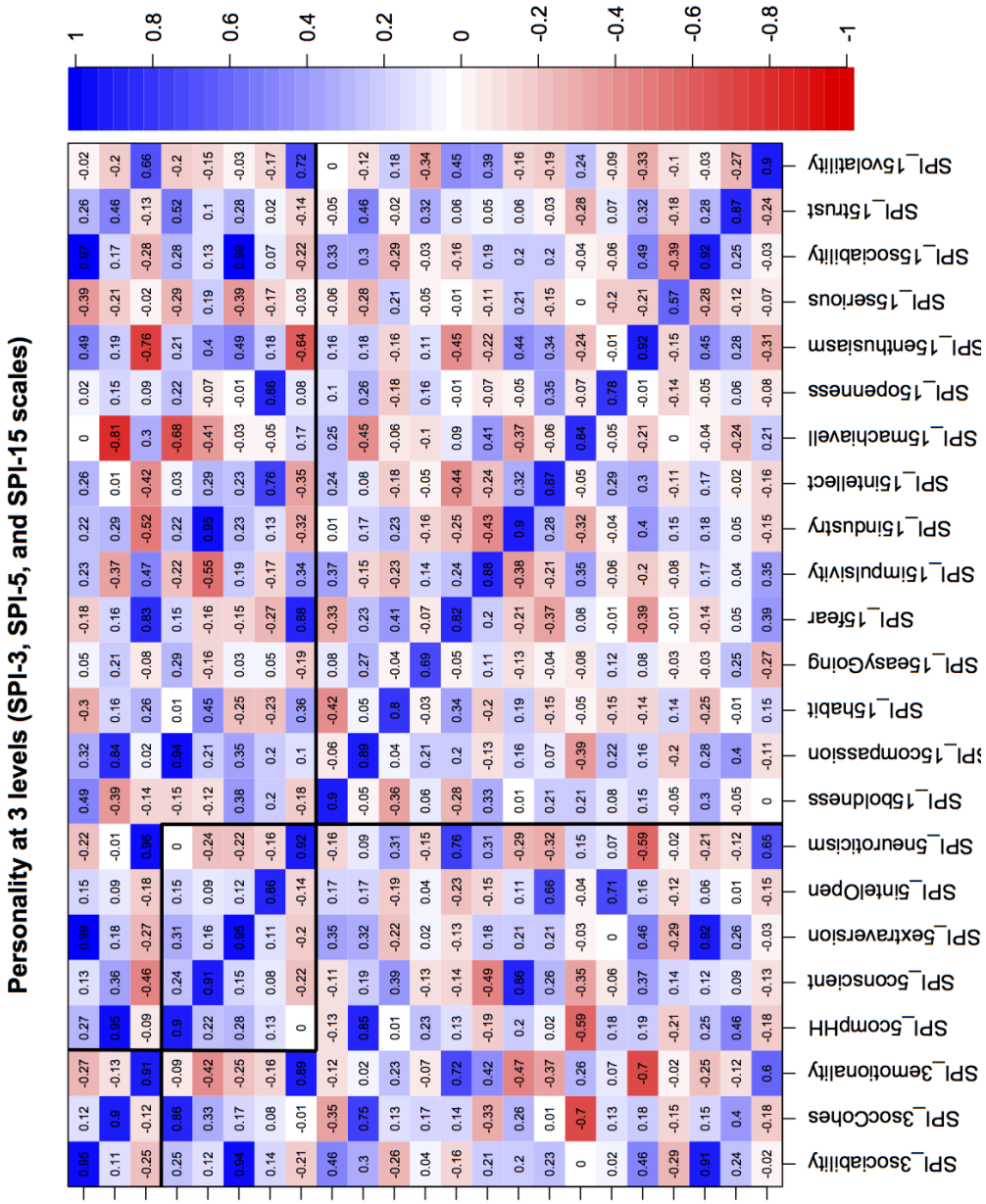


Figure 3.13: Correlations corrected for item overlap; alphas on the diagonal and correlations corrected for attenuation above the diagonal

The correlations between these new scales and the extant scales are presented in Figure 3.14 for the 3 factor solution and Figure 3.15 for the 5 factor solution. In the first case, the correlations between the new SPI-3 scales (Social Cohesion, Sociability, and Emotionality) and the three factor P-E-N model ranged from 0.75 to 0.91. In the second case (Figure 3.15), the SPI-5 scales (Neuroticism, Extraversion, Compassion/Honesty/Humility, Conscientiousness, and Intellect/Openness) also correlated highly with both the IPIP100 and IPIP-NEO scales.

Correlations between the SPI-5 scales and extant scales are presented differently in Figure 3.16 in that only scale correlations above 0.3 are shown. The correlations among construct-related scales are larger between the SPI-5 and both the IPIP100 and the IPIP-NEO than they are between the IPIP100 and the IPIP-NEO. Similarly, while the IPIP-NEO and the IPIP100 contain 6 cross-loading correlations above 0.30, the SPI-5 scales only have 3 cross-loadings above 0.30 with the IPIP-NEO (-0.33 between SPI-5 Neuroticism and IPIP-NEO Conscientiousness; -0.37 between SPI-5 Extraversion and IPIP-NEO Neuroticism; and -0.38 between SPI-5 Conscientiousness and IPIP-NEO Neuroticism) and only one cross-loading with the IPIP100 (0.47 between IPIP Agreeableness and SPI-5 Extraversion).

Correlations above 0.5 between the SPI-15 scales, the BFAS, and the MPQ are shown in Figure 3.17. While there are many correlations above the relatively high value of 0.5, all of the 10 BFAS scales and all but 1 of the MPQ scales correlated above 0.5 with at least 1 of the SPI-15 scales.

The item content of the SPI scales at each level is presented in Tables 3.6 through 3.8. Note that the factor loadings in these tables are based on the independent factoring of the items listed for the scales rather than the first round of factoring including all items. These tables also include complexity and communality for each item. A total of 150 items were used to create the scales across all levels, including 20 item scales for SPI-3, 15 item scales for SPI-5, and 8 item scales for SPI-15; several items were used at more than one level. Appendix A lists all of the 150 items in

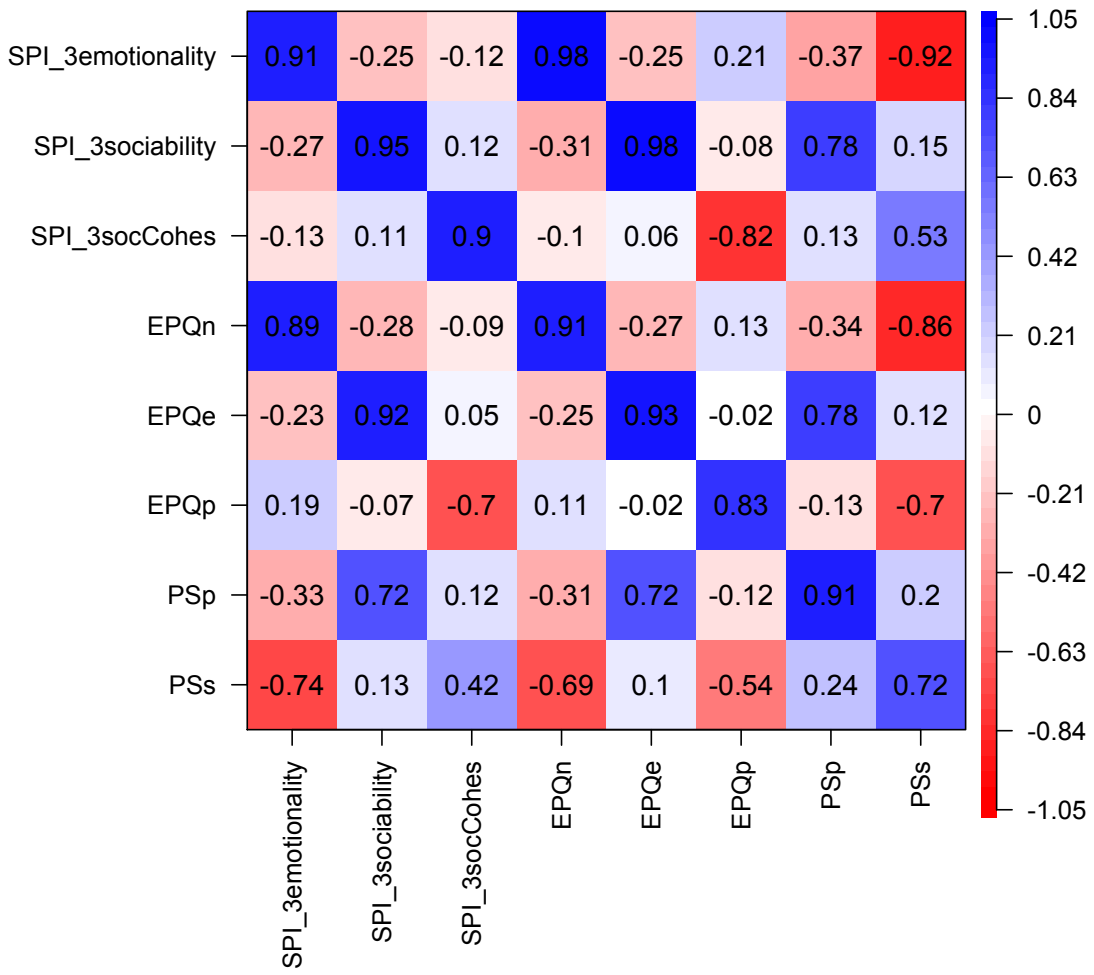


Figure 3.14: Scale correlations corrected for item overlap with alphas shown on the diagonal and correlations corrected for attenuation above the diagonal

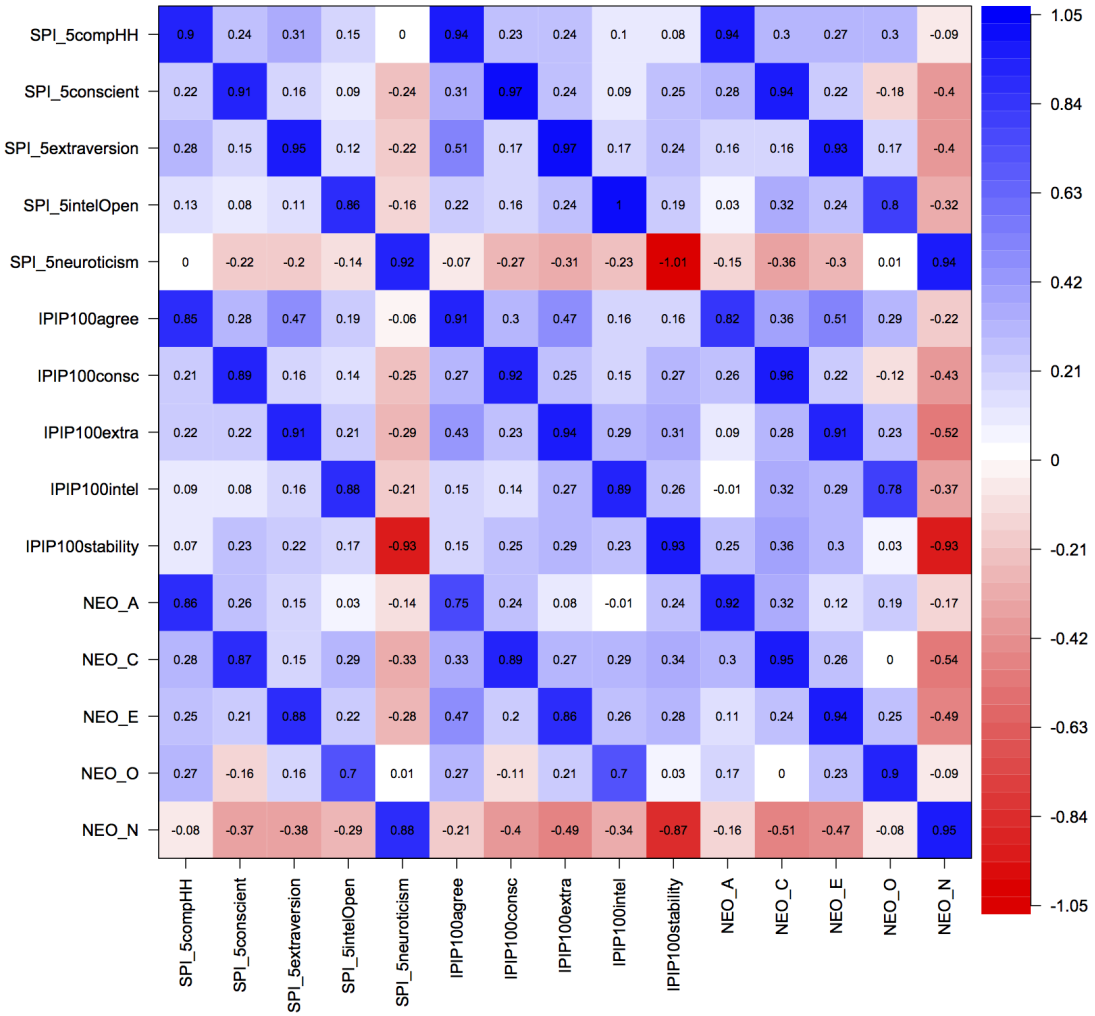


Figure 3.15: Scale correlations corrected for item overlap with alphas shown on the diagonal and correlations corrected for attenuation above the diagonal

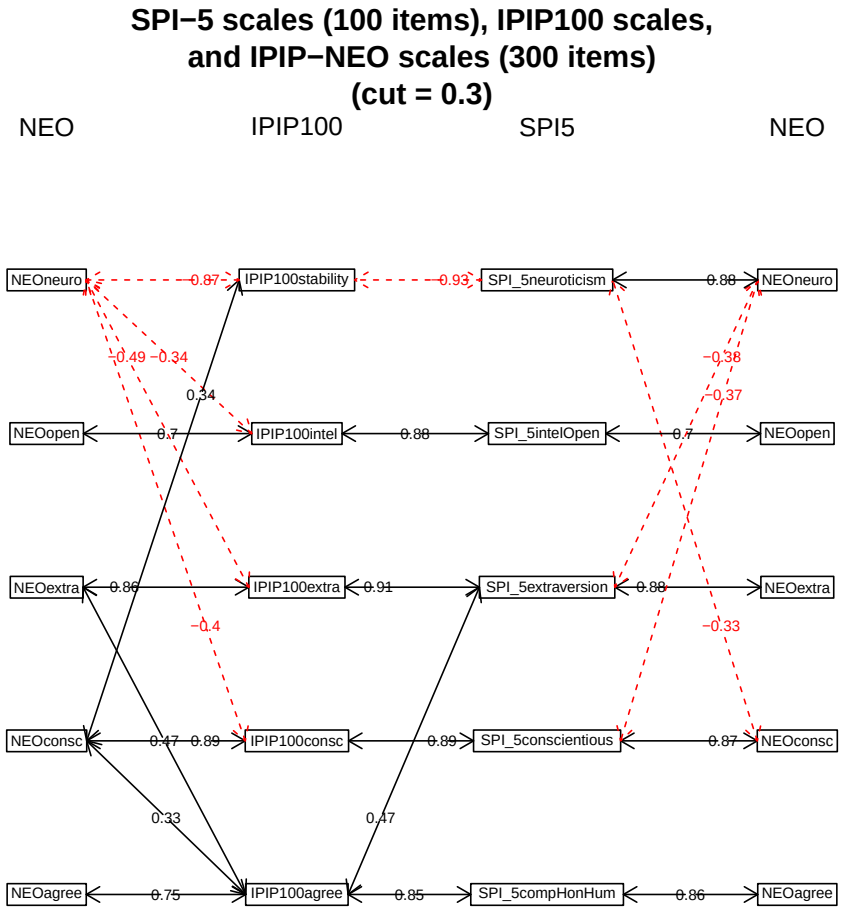


Figure 3.16: Scale correlations corrected for item overlap for the SPI-5, IPIP100, and IPIP-NEO scales

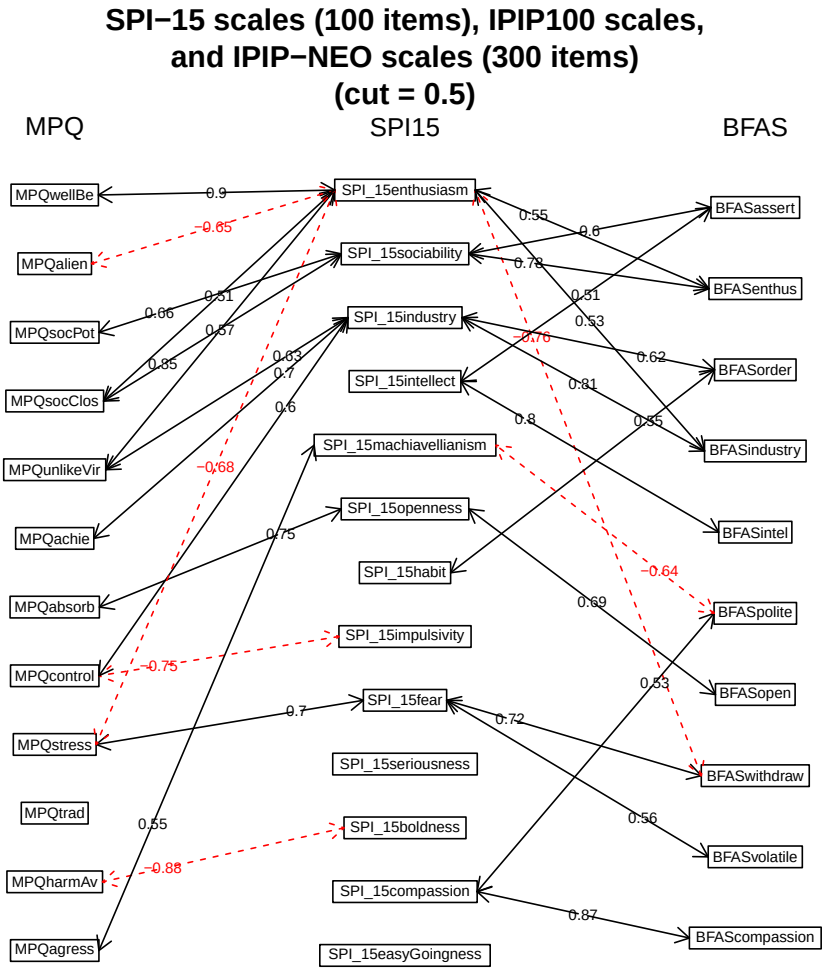


Figure 3.17: Scale correlations corrected for item overlap for the SPI-15, the MPQ, and the BFAS scales

alphabetical order as well as the various levels and scales in which each item is used.

Table 3.6: SPI3 scales and loadings

Item	Emotionality	Sociability	Social Cohesion	c_i	h^2
Emotionality					
Get overwhelmed by emotions.	0.69	0.18	-0.14	1.23	0.45
Get caught up in my problems.	0.66	0.00	-0.01	1.00	0.44
Need reassurance.	0.65	0.05	-0.18	1.16	0.41
Find my feelings are easily hurt.	0.65	0.11	-0.20	1.24	0.42
Have frequent mood swings.	0.64	0.09	0.08	1.07	0.41
Get upset easily.	0.62	0.07	0.02	1.03	0.37
Am a person whose moods go up and down easily.	0.64	0.09	0.10	1.09	0.42
Have a low opinion of myself.	0.62	-0.22	-0.05	1.26	0.48
Am often down in the dumps.	0.61	-0.20	0.09	1.26	0.49
Feel a sense of worthlessness or hopelessness.	0.58	-0.21	0.12	1.35	0.47
Rarely worry.	-0.57	0.03	0.23	1.32	0.35
Am a worrier.	0.56	-0.06	-0.24	1.38	0.36
Get easily agitated.	0.55	0.05	0.11	1.09	0.32
Can be stirred up easily.	0.53	0.19	0.07	1.28	0.28
Get irritated easily.	0.51	0.08	0.23	1.46	0.33
Waste my time.	0.51	-0.09	0.26	1.55	0.39
Find it difficult to get down to work.	0.49	-0.03	0.20	1.34	0.31
Neglect my duties.	0.48	0.02	0.33	1.76	0.38
Make careless mistakes.	0.48	0.11	0.29	1.75	0.34
Let myself be influenced by others.	0.48	0.11	-0.09	1.17	0.21
Sociability					
Like mixing with people.	0.00	0.76	-0.09	1.03	0.59
Enjoy meeting new people.	-0.06	0.73	-0.10	1.05	0.57
Can easily get some life into a dull party.	-0.10	0.72	0.14	1.11	0.56
Am a talkative person.	0.08	0.72	0.00	1.03	0.50
Am rather lively.	-0.09	0.71	0.01	1.03	0.54
Talk to a lot of different people at parties.	-0.05	0.70	0.03	1.01	0.51
Start conversations.	-0.11	0.69	-0.03	1.05	0.52
Love to chat.	0.11	0.68	-0.12	1.12	0.47
Don't talk a lot.	-0.06	-0.68	0.00	1.02	0.45
Make friends easily.	-0.16	0.68	-0.07	1.14	0.54
c_i = item complexity; h^2 = communality	*continued on next page*				

Table 3.6 – SPI-3 scales and loadings (continued from previous page)

Item	Emotionality	Sociability	Social Cohesion	c_i	h^2
Am mostly quiet when with other people.	0.04	-0.66	-0.05	1.02	0.45
Tend to keep in the background on social occasions.	0.12	-0.66	-0.06	1.09	0.48
Seem to derive less enjoyment from interacting with people than others do.	0.02	-0.66	0.21	1.20	0.49
Keep in the background.	0.12	-0.64	-0.08	1.11	0.46
Prefer reading to meeting people.	0.05	-0.57	0.05	1.03	0.34
Seek adventure.	-0.07	0.51	0.27	1.55	0.34
Take risks.	-0.15	0.49	0.37	2.08	0.40
Love action.	-0.15	0.46	0.24	1.72	0.30
Like going out a lot.	0.08	0.64	0.08	1.06	0.39
Say little.	-0.02	-0.61	0.01	1.00	0.37
Social Cohesion					
Take advantage of others.	0.07	-0.03	0.62	1.03	0.40
Am concerned about others.	0.15	0.40	-0.60	1.89	0.52
Use others for my own ends.	0.04	-0.02	0.59	1.01	0.36
Have a soft heart.	0.29	0.32	-0.56	2.16	0.44
Am indifferent to the feelings of others.	-0.12	-0.27	0.55	1.56	0.37
Sympathize with others' feelings.	0.20	0.36	-0.55	2.02	0.44
Am sensitive to the needs of others.	0.15	0.34	-0.55	1.87	0.42
Look down on others.	0.09	-0.09	0.54	1.11	0.33
Seek danger.	-0.04	0.32	0.54	1.64	0.38
Cheat to get ahead.	0.13	0.04	0.52	1.14	0.30
Love dangerous situations.	-0.02	0.29	0.51	1.58	0.33
Can't be bothered with other's needs.	-0.08	-0.23	0.50	1.47	0.31
Am out for my own personal gain.	-0.01	-0.01	0.50	1.00	0.25
Stick to the rules.	-0.08	-0.14	-0.49	1.22	0.26
Tell other people what they want to hear so that they will do what I want them to do.	0.11	0.05	0.49	1.13	0.26
Feel sympathy for those who are worse off than myself.	0.19	0.27	-0.48	1.93	0.31
Believe that I am better than others.	-0.11	-0.03	0.48	1.11	0.23
Tell a lot of lies.	0.25	0.00	0.48	1.51	0.33
Inquire about others' well-being.	0.13	0.44	-0.45	2.17	0.39
Like to do frightening things.	-0.05	0.27	0.44	1.70	0.26

c_i = item complexity; h^2 = communality

Table 3.7: SPI-5 scales and loadings

Item	Extraversion	Neuroticism	Conscientiousness	Comp/Hum/Hon	Intel/Open	c_i	h^2
Extraversion							
Like mixing with people.	0.77	-0.03	-0.04	-0.14	0.00	1.07	0.62
Enjoy meeting new people.	0.73	-0.07	-0.01	-0.15	0.02	1.11	0.60
Can easily get some life into a dull party.	0.72	-0.04	0.03	0.11	0.09	1.09	0.58
Talk to a lot of different people at parties.	0.72	-0.03	0.01	0.00	0.01	1.00	0.53
Tend to keep in the background on social occasions.	-0.72	0.16	0.04	-0.04	0.05	1.13	0.55
Seem to derive less enjoyment from interacting with people than others do.	-0.72	0.10	0.03	0.20	0.17	1.33	0.59
Am rather lively.	0.70	-0.03	0.05	-0.03	0.10	1.06	0.55
Make friends easily.	0.70	-0.12	0.10	-0.06	-0.01	1.12	0.58
Love to chat.	0.69	0.13	0.05	-0.11	-0.04	1.14	0.49
Am a talkative person.	0.69	0.13	0.02	-0.04	0.08	1.11	0.50
Prefer reading to meeting people.	-0.68	0.14	0.03	0.00	0.25	1.36	0.50
Don't talk a lot.	-0.68	-0.09	-0.02	0.01	-0.02	1.04	0.46
Start conversations.	0.67	-0.03	0.10	-0.05	0.09	1.10	0.52
Am mostly quiet when with other people.	-0.67	0.02	0.01	-0.02	-0.05	1.02	0.47
Keep in the background.	-0.66	0.07	-0.06	-0.08	-0.05	1.08	0.48
Neuroticism							
Get upset easily.	0.05	0.71	0.07	0.13	-0.08	1.12	0.52
Get overwhelmed by emotions.	0.09	0.71	-0.08	-0.19	0.07	1.21	0.52
Am a person whose moods go up and down easily.	0.02	0.71	-0.05	0.09	0.08	1.07	0.53
Have frequent mood swings.	0.02	0.69	-0.05	0.07	0.06	1.05	0.50
Am a worrier.	-0.12	0.68	0.17	-0.12	-0.06	1.28	0.50
Get caught up in my problems.	-0.07	0.67	-0.10	-0.04	0.05	1.09	0.51
Find my feelings are easily hurt.	0.05	0.67	-0.02	-0.17	-0.05	1.16	0.47
Get easily agitated.	0.03	0.65	0.06	0.19	-0.03	1.21	0.46
Rarely worry.	0.08	-0.65	-0.10	0.15	0.06	1.21	0.44

c_i = item complexity; h^2 = communality

continued on next page

Table 3.7 – SPI-5 scales and loadings (continued from previous page)

Item	Extraversion	Neuroticism	Conscientiousness	Comp/Hum/Hon	Intel/Open	c_i	h^2
Panic easily.	0.01	0.65	0.00	-0.10	-0.14	1.14	0.46
Need reassurance.	0.02	0.62	-0.07	-0.16	-0.09	1.21	0.44
Get irritated easily.	0.07	0.62	0.06	0.32	-0.01	1.54	0.49
Can be stirred up easily.	0.17	0.61	0.05	0.14	-0.04	1.28	0.39
Am often down in the dumps.	-0.29	0.60	-0.17	0.01	0.11	1.71	0.58
Become overwhelmed by events.	-0.10	0.60	-0.04	-0.17	-0.08	1.27	0.43
Conscientiousness							
Complete my duties as soon as possible.	0.04	-0.02	0.65	0.01	-0.04	1.02	0.43
Do things according to a plan.	-0.05	0.08	0.65	0.01	-0.07	1.07	0.40
Like order.	-0.05	0.17	0.64	-0.04	-0.07	1.19	0.39
Neglect my duties.	0.02	0.19	-0.62	0.08	0.02	1.22	0.49
Do too little work.	-0.06	0.09	-0.61	0.04	-0.04	1.08	0.43
Start tasks right away.	0.08	-0.05	0.61	0.05	0.04	1.07	0.41
Get to work at once.	0.05	-0.05	0.60	0.05	-0.01	1.04	0.37
Keep things tidy.	0.03	0.00	0.59	0.04	-0.10	1.07	0.35
See that rules are observed.	0.03	0.20	0.56	-0.11	-0.19	1.62	0.37
Waste my time.	-0.11	0.25	-0.56	0.03	0.05	1.50	0.48
Make careless mistakes.	0.12	0.21	-0.55	0.08	-0.01	1.45	0.41
Get chores done right away.	0.08	-0.07	0.55	0.01	-0.06	1.10	0.34
Stick to the rules.	-0.10	0.13	0.54	-0.18	-0.27	1.94	0.42
Do things by the book.	-0.15	0.15	0.54	-0.04	-0.28	1.87	0.37
Find it difficult to get down to work.	-0.05	0.25	-0.51	-0.01	0.04	1.48	0.39
Compassion/Honesty/Humility							
Am concerned about others.	0.27	0.22	0.13	-0.65	0.13	1.77	0.58
Am indifferent to the feelings of others.	-0.16	-0.12	-0.04	0.62	-0.08	1.25	0.43
Sympathize with others' feelings.	0.25	0.24	0.09	-0.61	0.08	1.76	0.48
Am sensitive to the needs of others.	0.22	0.21	0.10	-0.60	0.12	1.69	0.47
Take advantage of others.	0.02	0.06	-0.15	0.60	0.05	1.17	0.43

c_i = item complexity; h^2 = communality

continued on next page

Table 3.7 – SPI-5 scales and loadings (continued from previous page)

Item	Extraversion	Neuroticism	Conscientiousness	Comp/Hum/Hon	Intel/Open	c_i	h^2
Use others for my own ends.	0.05	0.03	-0.13	0.59	0.01	1.12	0.40
Have a soft heart.	0.24	0.30	0.06	-0.59	0.01	1.86	0.47
Look down on others.	-0.06	0.16	-0.01	0.58	0.06	1.21	0.40
Feel sympathy for those who are worse off than myself.	0.16	0.21	0.03	-0.56	0.13	1.57	0.37
Can't be bothered with other's needs.	-0.15	-0.07	-0.04	0.56	-0.05	1.21	0.35
Am out for my own personal gain.	0.07	0.01	-0.02	0.55	-0.04	1.05	0.31
Believe that I am better than others.	-0.03	0.03	0.08	0.51	0.16	1.27	0.29
Feel others' emotions.	0.25	0.26	0.09	-0.49	0.17	2.48	0.39
Feel that most people can't be trusted.	-0.15	0.23	0.15	0.49	0.02	1.83	0.34
Inquire about others' well-being.	0.32	0.23	0.15	-0.48	0.14	2.70	0.43
Intellect/Openness							
Am able to come up with new and different ideas.	0.07	-0.04	0.10	-0.04	0.64	1.09	0.46
Am full of ideas.	0.05	0.02	0.09	-0.04	0.64	1.07	0.44
Have a rich vocabulary.	-0.13	-0.04	0.04	-0.06	0.59	1.14	0.34
Am not interested in abstract ideas.	0.18	0.04	0.13	0.13	-0.55	1.47	0.32
Have a vivid imagination.	0.04	0.15	-0.05	-0.13	0.54	1.29	0.31
Enjoy thinking about things.	-0.15	0.06	0.03	-0.20	0.53	1.49	0.29
Can handle complex problems.	-0.09	-0.14	0.30	0.03	0.51	1.87	0.39
Know the answers to many questions.	-0.05	-0.12	0.12	0.11	0.49	1.39	0.30
Believe in the importance of art.	-0.01	0.12	-0.10	-0.28	0.49	1.84	0.29
Spend time reflecting on things.	-0.21	0.19	0.05	-0.23	0.49	2.22	0.29
Need a creative outlet.	0.02	0.17	-0.12	-0.17	0.48	1.70	0.27
Am quick to understand things.	-0.04	-0.15	0.23	0.00	0.47	1.71	0.34
Can handle a lot of information.	-0.03	-0.11	0.27	0.07	0.44	1.87	0.32
Love to reflect on things.	-0.08	0.20	0.06	-0.27	0.45	2.21	0.26
Come up with good solutions.	0.03	-0.16	0.33	-0.03	0.51	1.95	0.47

c_i = item complexity; h^2 = communality

Table 3.8: SPI-15 scales and loadings

Item	Soc	Out	Ind	Comp	Vol	Ope	Mac	Int	Bold	Con	Fear	Imp	Trus	Easy	Seri	c_i	h^2
Sociability																	
Am mostly quiet when with other people.	-0.76	0.02	0.00	0.00	-0.06	-0.03	0.04	-0.06	0.10	0.00	0.10	-0.11	0.10	0.12	0.08	1.27	0.67
Tend to keep in the background on social occasions.	-0.75	0.01	-0.05	0.02	0.04	0.10	0.04	-0.03	-0.03	0.14	0.03	-0.08	0.07	0.09	0.01	1.21	0.67
Don't talk a lot.	-0.73	0.02	-0.05	-0.03	-0.10	-0.01	0.06	0.00	0.11	-0.08	0.07	-0.21	0.13	0.16	0.13	1.59	0.70
Like mixing with people.	0.73	0.00	0.02	0.07	-0.07	0.05	-0.01	0.05	0.14	-0.01	0.08	-0.02	0.09	0.17	0.00	1.31	0.70
Talk to a lot of different people at parties.	0.73	-0.08	0.13	0.04	0.03	0.02	0.05	0.04	0.04	-0.05	-0.08	0.06	0.01	0.14	-0.04	1.26	0.63
Keep in the background.	-0.70	-0.06	-0.01	0.07	-0.03	0.02	0.05	-0.14	0.05	0.05	-0.05	-0.05	0.10	0.18	0.01	1.37	0.63
Prefer reading to meeting people.	-0.68	-0.01	0.02	-0.01	0.07	0.23	0.04	0.06	-0.19	0.00	-0.03	0.09	0.01	0.00	0.03	1.5	0.59
Seem to derive less enjoyment from interacting with people than others do.	-0.68	-0.09	0.00	-0.10	0.02	0.09	0.10	0.04	-0.10	0.06	-0.06	0.06	-0.11	-0.03	0.07	1.35	0.64
Enthusiasm																	
Love life.	0.05	0.73	0.07	0.12	0.04	0.07	0.01	-0.01	0.06	0.03	0.09	0.06	0.06	0.03	0.01	1.2	0.62
Feel a sense of worthlessness or hopelessness.	-0.06	-0.72	-0.08	0.03	0.01	0.02	0.06	0.03	0.01	0.00	0.08	0.08	0.06	0.05	0.09	1.16	0.66
Am often down in the dumps.	-0.03	-0.71	-0.14	0.05	0.11	0.08	-0.05	0.06	-0.06	0.02	0.11	0.03	-0.02	0.00	0.13	1.34	0.71
Dislike myself.	-0.07	-0.70	0.00	0.04	0.01	-0.01	0.13	-0.04	0.03	0.08	0.00	0.12	0.06	0.04	-0.02	1.22	0.64
Am very pleased with myself.	0.07	0.69	0.15	-0.03	0.02	0.06	0.06	0.07	-0.05	-0.01	0.05	0.02	-0.07	0.08	0.17	1.39	0.66
Am happy with my life.	0.04	0.69	0.08	0.11	-0.04	0.01	-0.02	-0.05	0.01	0.15	0.01	0.08	0.05	0.08	-0.04	1.3	0.59
Have a low opinion of myself.	-0.05	-0.67	-0.01	0.05	-0.01	-0.02	0.02	-0.14	0.05	0.11	0.02	0.10	0.14	0.05	-0.04	1.36	0.64
Have a dark outlook on the future.	-0.07	-0.65	-0.05	-0.15	0.06	0.13	0.06	-0.07	-0.06	0.08	-0.15	0.07	-0.04	0.05	0.08	1.55	0.65
Industry																	

c_i = item complexity; h^2 = communality

continued on next page

Table 3.8 – SPI-15 scales and loadings (continued from previous page)

Item	Soc	Out	Ind	Comp	Vol	Ope	Mac	Int	Bold	Con	Fear	Imp	Trus	Easy	Seri	c_i	h^2
Complete my duties as soon as possible.	0.06	-0.07	0.76	-0.06	-0.04	0.00	-0.04	0.03	-0.07	0.05	0.00	0.04	0.03	0.11	0.03	1.14	0.60
Get chores done right away.	0.10	-0.03	0.75	-0.04	-0.04	0.04	-0.02	-0.11	-0.09	-0.07	-0.03	0.05	0.03	0.10	0.12	1.26	0.55
Start tasks right away.	-0.02	0.03	0.74	-0.02	0.03	0.01	-0.04	0.09	-0.05	-0.06	-0.01	0.12	0.04	0.05	0.09	1.18	0.58
Get to work at once.	-0.03	-0.02	0.68	-0.02	0.04	-0.03	-0.07	0.07	0.04	0.02	-0.07	0.06	0.08	0.08	0.07	1.19	0.52
Keep things tidy.	0.05	-0.02	0.66	0.01	-0.01	-0.02	0.03	-0.10	-0.08	0.00	0.06	-0.09	-0.07	0.11	0.11	1.29	0.48
Find it difficult to get down to work.	0.01	-0.03	-0.62	0.03	0.06	0.14	0.11	-0.08	-0.04	0.11	0.04	-0.01	0.04	0.03	0.06	1.37	0.52
Waste my time.	-0.08	-0.16	-0.56	0.04	0.09	0.08	0.16	-0.01	-0.03	0.07	-0.09	0.09	0.08	0.14	-0.03	1.92	0.58
Neglect my duties.	0.05	-0.10	-0.56	-0.05	0.02	0.10	0.18	-0.04	-0.02	-0.05	0.05	0.09	0.11	0.09	0.10	1.71	0.57
Compassion																	
Am concerned about others.	0.03	-0.05	0.06	0.74	-0.03	0.03	-0.14	0.02	0.04	0.04	-0.04	0.03	0.09	0.00	-0.01	1.17	0.68
Feel others' emotions.	0.00	0.00	0.02	0.70	0.01	0.05	0.03	0.00	-0.04	-0.04	0.10	-0.01	-0.12	0.08	0.03	1.16	0.53
Am sensitive to the needs of others.	0.06	-0.07	-0.02	0.69	-0.06	0.04	-0.02	0.07	-0.03	-0.03	0.06	-0.12	0.10	0.05	0.03	1.23	0.58
Sympathize with others' feelings.	0.07	-0.04	-0.01	0.68	-0.05	0.04	-0.04	-0.02	-0.05	0.03	0.05	-0.05	0.02	0.10	0.01	1.15	0.57
Have a soft heart.	-0.01	0.01	-0.06	0.62	-0.03	-0.03	-0.12	0.01	-0.07	0.04	0.17	0.06	0.04	0.10	0.07	1.42	0.55
Feel sympathy for those who are worse off than myself.	-0.02	-0.02	-0.08	0.62	0.03	0.05	-0.13	0.01	0.01	0.05	-0.03	-0.02	0.04	0.01	-0.04	1.19	0.45
Can't be bothered with other's needs.	0.02	0.01	0.02	-0.60	0.07	0.08	0.10	-0.01	-0.08	-0.01	0.04	0.04	-0.06	0.17	0.20	1.64	0.49
Inquire about others' well-being.	0.11	-0.01	0.06	0.58	0.06	0.06	-0.07	0.06	0.04	0.09	-0.03	0.01	0.05	0.09	-0.06	1.35	0.49
Volatility																	
Rarely show my anger.	-0.11	-0.08	0.01	-0.04	-0.73	0.05	0.11	0.04	0.04	0.10	0.09	-0.05	0.05	0.01	0.07	1.26	0.52
Get angry easily.	-0.06	-0.05	-0.04	0.00	0.72	-0.09	0.03	0.01	0.07	-0.01	0.03	0.08	0.00	0.02	0.11	1.16	0.64
Lose my temper.	-0.01	0.04	-0.02	0.04	0.71	-0.09	-0.01	0.00	0.02	-0.04	-0.04	0.16	0.01	-0.04	0.11	1.22	0.57
Get irritated easily.	-0.04	-0.08	-0.06	-0.07	0.68	-0.02	-0.02	0.08	0.00	0.05	0.15	0.08	-0.04	0.00	-0.02	1.25	0.64
Can be stirred up easily.	-0.05	0.02	-0.02	0.11	0.59	0.01	0.01	0.00	-0.01	0.09	0.08	0.17	0.15	-0.05	0.06	1.54	0.51
Find that it takes a lot to make me feel angry at someone.	-0.01	0.02	-0.04	0.06	-0.68	0.04	-0.06	0.03	0.02	0.11	-0.01	0.05	0.06	0.09	0.09	1.2	0.54
Seldom get mad.	0.00	0.05	-0.05	-0.05	-0.67	0.12	0.08	-0.01	-0.06	0.04	-0.02	-0.04	0.07	0.06	0.10	1.25	0.51

c_i = item complexity; h^2 = communality

continued on next page

Table 3.8 – SPI-15 scales and loadings (continued from previous page)

Item	Soc	Out	Ind	Comp	Vol	Ope	Mac	Int	Bold	Con	Fear	Imp	Trus	Easy	Seri	c_i	h^2
Get easily agitated.	-0.03	-0.17	0.02	-0.01	0.55	0.01	-0.04	-0.01	-0.06	0.10	0.14	0.15	-0.02	0.05	0.02	1.64	0.54
Openness																	
Enjoy thinking about things.	-0.08	0.07	0.01	0.03	-0.03	0.60	0.01	0.08	-0.08	0.06	-0.02	-0.03	0.09	0.04	0.05	1.24	0.40
Believe in the importance of art.	0.02	-0.03	0.02	0.10	0.07	0.57	-0.05	0.03	0.03	-0.16	0.06	-0.06	0.16	0.08	0.02	1.56	0.40
Am not interested in abstract ideas.	0.04	0.03	0.07	0.01	0.05	-0.55	0.03	-0.02	-0.07	0.02	0.08	0.08	0.02	0.09	0.07	1.3	0.37
Spend time reflecting on things.	-0.10	-0.07	0.01	0.19	0.01	0.54	0.04	-0.04	-0.05	0.10	-0.04	-0.13	-0.05	0.04	0.03	1.67	0.38
Need a creative outlet.	-0.04	0.05	-0.03	0.08	0.07	0.53	-0.05	-0.02	0.09	-0.06	0.11	0.05	0.03	0.02	0.06	1.39	0.34
Have a vivid imagination.	0.04	-0.02	0.01	-0.01	0.03	0.52	-0.07	0.12	0.05	0.02	0.14	0.06	-0.16	0.14	-0.08	1.84	0.39
Love to reflect on things.	-0.08	-0.01	0.02	0.19	-0.04	0.49	0.06	0.06	-0.04	0.10	0.06	-0.05	0.06	0.08	0.04	1.76	0.34
Try to understand myself.	-0.04	0.03	0.02	0.06	0.01	0.46	0.00	0.07	-0.06	-0.02	0.07	-0.09	0.04	0.08	0.00	1.34	0.25
Machiavellianism																	
Tell other people what they want to hear so that they will do what I want them to do.	-0.08	0.03	-0.04	-0.04	-0.05	-0.08	0.68	0.07	0.03	0.03	0.06	0.05	-0.05	0.05	0.03	1.17	0.52
Use others for my own ends.	-0.01	0.03	-0.08	-0.14	0.08	-0.03	0.60	-0.02	0.05	0.01	-0.11	-0.05	-0.03	0.02	0.12	1.42	0.54
Take advantage of others.	-0.03	-0.03	-0.02	-0.16	0.07	-0.01	0.58	-0.03	0.04	-0.06	-0.08	0.02	-0.04	-0.06	0.15	1.46	0.54
Tell a lot of lies.	-0.05	-0.14	-0.10	-0.04	0.00	-0.03	0.57	-0.03	0.06	-0.03	0.00	0.05	-0.05	0.11	-0.02	1.37	0.47
Cheat to get ahead.	-0.02	0.03	-0.10	-0.08	0.02	-0.08	0.57	-0.04	0.07	-0.06	0.02	0.01	-0.07	0.04	0.05	1.29	0.44
Play a role in order to impress people.	0.10	-0.12	-0.04	-0.02	-0.09	0.00	0.54	0.04	0.03	0.13	0.08	0.10	0.20	-0.11	0.04	1.98	0.45
Use flattery to get ahead.	0.16	0.02	-0.01	0.02	0.00	-0.10	0.52	0.12	0.03	0.02	0.06	0.06	-0.02	0.05	-0.02	1.48	0.36
Cannot imagine lying or cheating.	0.06	0.01	0.09	0.04	-0.01	0.00	-0.52	0.03	-0.01	0.12	0.04	0.00	0.10	-0.05	0.23	1.76	0.38
Intellect																	
Am quick to understand things.	-0.04	0.02	0.01	-0.02	-0.05	0.19	0.00	0.61	-0.05	-0.02	-0.03	-0.06	0.00	0.03	-0.06	1.28	0.47
Catch on to things quickly.	-0.04	0.04	0.04	0.01	0.00	0.11	-0.02	0.59	-0.01	-0.01	-0.07	0.03	0.01	0.06	-0.08	1.21	0.42
Think quickly.	0.10	-0.02	0.04	0.04	0.02	0.08	0.01	0.57	0.04	0.03	-0.12	0.05	-0.11	0.06	-0.06	1.4	0.49
Can handle complex problems.	-0.13	-0.03	0.04	0.08	-0.03	0.15	-0.04	0.57	0.08	0.03	-0.14	-0.07	-0.02	-0.10	0.02	1.63	0.49
Can handle a lot of information.	0.00	-0.09	0.15	-0.03	-0.07	0.18	0.09	0.52	-0.03	0.01	-0.10	-0.02	0.02	-0.02	-0.07	1.77	0.42

c_i = item complexity; h^2 = communality

continued on next page

Table 3.8 – SPI-15 scales and loadings (continued from previous page)

Item	Soc	Out	Ind	Comp	Vol	Ope	Mac	Int	Bold	Con	Fear	Imp	Trus	Easy	Seri	c_i	h^2
Learn things slowly.	0.01	0.02	0.02	0.04	0.06	-0.02	0.02	-0.50	-0.01	0.06	0.03	0.02	0.08	0.20	0.23	1.93	0.38
Come up with good solutions.	0.00	0.06	0.08	0.08	-0.06	0.23	-0.05	0.50	0.01	0.01	-0.04	-0.11	-0.13	0.01	0.04	1.97	0.52
Am more capable than most others.	-0.07	0.14	0.12	-0.10	-0.03	0.11	0.18	0.44	0.02	0.04	0.06	0.01	-0.06	-0.08	0.07	2.35	0.39
Boldness																	
Love dangerous situations.	-0.02	-0.06	0.02	0.03	0.03	0.01	0.13	-0.06	0.69	-0.01	-0.13	0.12	-0.10	-0.04	0.10	1.34	0.62
Seek danger.	0.01	-0.03	0.03	-0.03	-0.01	0.03	0.17	-0.04	0.67	-0.09	-0.09	0.16	-0.05	-0.05	0.04	1.4	0.63
Seek adventure.	0.19	0.05	0.02	0.03	0.01	0.15	-0.05	0.02	0.62	-0.03	0.04	0.04	-0.03	0.08	0.05	1.44	0.58
Would never go hang gliding or bungee jumping.	-0.02	0.06	-0.02	0.02	0.10	-0.05	-0.07	0.05	-0.60	-0.03	0.03	-0.01	0.00	0.00	0.10	1.23	0.38
Like to do frightening things.	0.03	-0.05	0.09	-0.01	0.00	0.07	0.17	-0.08	0.58	0.01	-0.11	0.09	-0.09	0.03	0.07	1.56	0.47
Take risks.	0.10	0.11	0.04	0.07	0.01	0.05	0.05	0.07	0.57	-0.15	0.01	0.11	-0.10	0.01	0.08	1.61	0.58
Love action.	0.26	0.03	0.02	-0.02	0.06	-0.02	-0.05	0.14	0.52	0.00	-0.01	-0.04	0.01	0.06	0.10	1.83	0.47
Am willing to take risks.	0.03	0.09	0.03	0.13	0.03	0.08	-0.01	0.16	0.51	-0.08	-0.09	0.09	-0.06	0.01	0.10	1.84	0.48
Habit																	
Dislike changes.	-0.15	-0.04	0.01	0.04	0.04	-0.15	0.09	-0.14	-0.27	0.45	0.01	0.14	-0.13	-0.03	-0.01	3.22	0.45
Prefer to stick with things that I know.	-0.10	0.01	0.03	0.00	0.00	-0.21	0.07	-0.18	-0.24	0.43	0.08	0.07	-0.09	0.16	0.02	3.43	0.45
Don't like the idea of change.	-0.08	-0.04	0.02	0.02	-0.01	-0.19	0.05	-0.18	-0.25	0.42	0.03	0.14	-0.15	0.01	0.04	3.53	0.41
Do things by the book.	-0.05	-0.05	0.25	0.01	0.01	-0.15	-0.01	0.04	-0.17	0.40	-0.10	-0.14	0.21	0.07	0.11	4.17	0.49
Want everything to add up perfectly.	-0.03	-0.03	0.19	-0.13	0.06	0.05	0.02	0.14	-0.02	0.40	0.21	-0.05	0.03	0.03	0.09	3.04	0.35
Stick to the rules.	0.03	-0.01	0.20	-0.02	-0.04	-0.13	-0.19	0.08	-0.18	0.39	0.03	-0.15	0.20	0.05	0.09	4.42	0.50
Am a creature of habit.	-0.17	0.04	0.09	0.02	0.07	-0.10	0.11	-0.03	-0.21	0.38	0.01	0.08	0.04	0.10	-0.08	3.16	0.31
Want everything to be "just right."	-0.01	-0.06	0.22	-0.05	0.08	-0.01	-0.02	0.14	-0.06	0.36	0.21	-0.03	0.06	0.02	0.10	3.41	0.35
Fear																	
Begin to panic when there is danger.	0.06	0.00	0.02	0.02	0.13	-0.02	0.01	-0.16	-0.30	0.15	0.47	0.12	0.05	0.04	-0.01	2.73	0.52
Worry a lot about my looks.	0.16	-0.11	-0.01	-0.01	0.01	-0.06	0.30	0.08	0.12	0.15	0.44	-0.16	0.07	0.07	-0.09	3.54	0.43
Panic easily.	0.04	-0.28	0.11	0.03	0.08	0.01	0.03	-0.15	-0.16	0.10	0.44	0.18	0.07	-0.11	0.03	3.48	0.55
Become overwhelmed by events.	-0.11	-0.21	-0.01	0.09	0.12	0.04	-0.01	-0.08	-0.12	0.07	0.40	0.09	0.12	-0.03	0.03	2.9	0.47

c_i = item complexity; h^2 = communality

continued on next page

Table 3.8 – SPI-15 scales and loadings (continued from previous page)

Item	Soc	Out	Ind	Comp	Vol	Ope	Mac	Int	Bold	Con	Fear	Imp	Trus	Easy	Seri	c_i	h^2
Tremble in dangerous situations.	-0.12	0.01	0.07	0.02	0.02	-0.01	0.10	-0.12	-0.25	0.03	0.40	0.11	0.15	0.05	0.01	3.07	0.38
Find my feelings are easily hurt.	-0.05	-0.15	-0.03	0.22	0.26	0.03	0.05	-0.02	-0.16	-0.01	0.40	0.08	0.12	-0.02	0.06	3.79	0.53
Need protection.	-0.08	-0.06	0.01	0.13	0.03	-0.01	0.11	-0.18	-0.13	0.05	0.36	0.15	0.04	0.01	0.18	3.82	0.38
Am not easily disturbed by events.	0.02	0.02	-0.02	-0.17	-0.21	-0.01	0.08	0.13	0.15	0.04	-0.33	-0.03	-0.06	0.16	0.06	4.34	0.36
Impulsivity																	
Do things without thinking of the consequences.	0.05	0.00	-0.08	-0.12	0.17	-0.02	0.04	-0.01	0.21	-0.14	0.04	0.55	0.09	0.12	-0.04	2.1	0.58
Act without thinking.	0.06	0.07	-0.13	-0.10	0.19	-0.08	0.03	-0.04	0.14	-0.09	0.12	0.55	0.12	0.05	-0.02	2.09	0.56
Jump into things without thinking.	0.10	-0.01	-0.21	-0.01	0.19	-0.09	0.00	0.08	0.22	-0.08	0.03	0.46	0.12	0.05	0.02	3.01	0.51
Don't see the consequences of things.	0.06	-0.07	-0.14	-0.18	0.04	-0.07	0.10	-0.06	0.04	-0.11	0.04	0.45	0.23	0.04	0.12	2.99	0.48
Rush into things.	0.08	0.06	-0.13	-0.03	0.17	-0.06	0.08	0.07	0.20	-0.16	0.19	0.44	0.12	-0.04	0.04	3.48	0.52
Make rash decisions.	0.05	-0.07	-0.13	-0.05	0.20	-0.08	0.04	0.05	0.22	-0.09	0.03	0.42	0.10	0.13	-0.02	3.12	0.46
People have said that I sometimes act rashly.	0.08	-0.01	-0.11	-0.06	0.28	0.04	0.01	0.00	0.20	-0.04	-0.03	0.41	0.02	0.00	0.08	2.87	0.44
Pay too little attention to details.	0.02	0.03	-0.22	-0.08	-0.02	-0.07	0.07	-0.06	-0.06	-0.16	0.06	0.36	0.15	0.05	0.07	3.38	0.33
Trust																	
Trust what people say.	0.23	0.15	0.03	0.13	-0.16	0.05	-0.03	-0.03	-0.15	-0.01	-0.04	0.15	0.54	0.04	0.02	2.38	0.54
Trust people to mainly tell the truth.	0.07	0.18	0.00	0.14	-0.10	0.05	-0.08	0.03	-0.05	0.04	-0.06	0.19	0.49	0.05	0.12	2.4	0.42
Trust others.	0.12	0.31	-0.02	0.26	-0.06	0.07	-0.03	-0.08	-0.09	-0.03	-0.07	0.09	0.45	0.05	0.12	3.42	0.51
Believe that others have good intentions.	0.10	0.22	0.01	0.25	-0.15	0.05	-0.11	0.01	-0.04	-0.01	-0.01	0.13	0.43	-0.02	0.08	3.22	0.47
Believe that people are basically moral.	0.02	0.24	0.02	0.22	-0.01	0.08	0.02	0.04	-0.09	-0.06	-0.06	0.13	0.42	0.10	0.00	2.99	0.37
Let myself be influenced by others.	0.04	-0.07	-0.03	0.08	-0.01	-0.02	0.29	-0.11	-0.03	0.13	0.09	0.17	0.41	-0.02	-0.09	3.16	0.42
Distrust people.	-0.16	-0.29	0.15	-0.14	0.15	-0.02	0.08	0.02	0.04	0.14	0.06	0.06	-0.39	0.09	0.07	4.24	0.48
Feel that most people can't be trusted.	-0.09	-0.21	0.08	-0.20	0.18	-0.07	0.07	-0.04	0.15	0.16	0.04	-0.04	-0.38	0.08	0.11	4.63	0.46
Easy-Goingness																	
Take things as they come.	0.00	0.14	-0.06	0.04	-0.05	0.05	-0.08	0.14	0.13	-0.02	-0.17	0.05	0.02	0.47	-0.05	2.09	0.38

c_i = item complexity; h^2 = communality

continued on next page

Table 3.8 – SPI-15 scales and loadings (continued from previous page)

Item	Soc	Out	Ind	Comp	Vol	Ope	Mac	Int	Bold	Con	Fear	Imp	Trus	Easy	Seri	c_i	h^2
Like to take it easy.	-0.06	0.17	-0.37	0.07	-0.01	0.04	0.03	0.00	-0.15	0.12	0.06	0.03	-0.17	0.46	-0.06	3.18	0.39
Prefer to just let things happen.	-0.15	0.12	-0.32	0.02	-0.07	-0.01	-0.01	-0.01	0.09	-0.02	0.03	0.19	-0.01	0.44	0.07	3	0.42
Let things proceed at their own pace.	-0.07	0.03	-0.20	0.07	-0.12	0.10	-0.08	-0.05	0.01	0.03	-0.08	-0.09	0.05	0.44	0.06	2.25	0.31
Am easy to satisfy.	0.03	0.18	-0.03	0.18	-0.19	-0.04	-0.10	-0.10	-0.01	0.10	-0.08	0.10	0.12	0.36	0.01	4.3	0.39
Value cooperation over competition.	0.00	-0.07	0.01	0.23	-0.06	0.21	-0.12	-0.02	-0.19	-0.06	0.00	0.01	0.14	0.30	0.00	4.87	0.31
Rarely put people under pressure.	-0.03	-0.10	-0.09	0.03	-0.24	0.06	-0.19	0.03	-0.07	-0.02	0.16	-0.02	0.03	0.30	-0.02	4.32	0.26
Am more easy-going about right and wrong than most people.	0.06	-0.02	0.06	-0.10	-0.19	0.04	0.21	-0.01	0.14	-0.23	0.11	0.23	-0.05	0.29	0.04	6.31	0.36
Seriousness																	
Seldom joke around.	-0.15	-0.11	0.14	-0.02	-0.01	-0.10	-0.07	-0.08	0.02	-0.20	0.06	-0.16	0.13	-0.04	0.47	2.85	0.36
Have never engaged in gossip.	-0.16	-0.08	0.03	-0.17	-0.23	0.19	-0.26	-0.07	-0.02	-0.09	-0.11	0.12	0.02	-0.06	0.43	4.45	0.40
Believe in one true religion.	0.07	0.15	-0.12	0.23	-0.01	-0.28	-0.09	-0.08	0.05	0.23	0.05	0.00	-0.14	-0.10	0.40	4.99	0.38
Am not easily amused.	-0.14	-0.26	0.14	-0.15	0.00	-0.06	0.07	-0.01	0.05	-0.15	-0.04	-0.17	0.03	-0.03	0.35	4.51	0.34
Most things taste the same to me.	-0.17	-0.11	0.04	-0.05	-0.07	-0.11	0.15	-0.11	-0.01	-0.04	-0.06	0.13	0.01	-0.02	0.34	3.49	0.26
Rarely overindulge.	0.01	-0.01	0.24	-0.02	-0.12	0.07	-0.20	-0.11	-0.04	-0.07	-0.17	-0.13	0.04	0.05	0.32	5	0.32
Never splurge.	-0.05	-0.14	0.18	-0.10	-0.13	0.17	-0.10	-0.21	-0.11	0.08	-0.17	-0.08	0.03	0.06	0.29	7.14	0.28
There are several people who keep trying to avoid me.	-0.07	-0.16	0.05	-0.01	0.15	0.05	0.12	-0.20	0.11	0.03	-0.13	0.14	0.07	-0.04	0.29	5.87	0.29

c_i = item complexity; h^2 = communality

Figures 3.18 shows the hierarchical organization of the SPI scales. Two of the SPI-15 scales do not correlate above 0.3 with any of the SPI-5 scales (Easy-Goingness and Seriousness). Six additional factors at this level correlate above 0.3 with more than one of the SPI-5 scales: SPI-15 Intellect with SPI-5 Intellect/Openness (0.66) and SPI-5 Neuroticism (-0.32); SPI-15 Habit with SPI-5 Conscientiousness (0.39) and SPI-5 Neuroticism (0.31); SPI-15 Impulsive with SPI-5 Conscientiousness (-0.49) and SPI-5 Neuroticism (0.39); and SPI-15 Machiavellianism with SPI-5 Conscientiousness (-0.35) and SPI-5 Compassion/Honesty/Humility (-0.59); SPI-15 Compassion with SPI-5 Compassion/Honesty/Humility (0.85) and SPI-5 Extraversion (0.32); and SPI-15 Enthusiasm with SPI-5 Neuroticism (-0.59), SPI-5 Conscientiousness (0.37) and SPI-5 Extraversion (0.46). The remainder of the SPI-15 scales (7) correlate above 0.3 with only one of the SPI-5 scales. The SPI-5 scales correlate with the SPI-3 scales as expected though SPI-5 Conscientiousness does have a moderate cross-loading with SPI-3 Emotionality (-0.42). Correlations above 0.3 between the SPI-15 and SPI-3 scales are generally limited such that each of the SPI-15 scales relate to only one of the SPI-3 scales, with three exceptions: SPI-15 Boldness correlates with both SPI-3 Social Cohesion (-0.35) and SPI-3 Sociability (0.46); SPI-15 Enthusiasm correlates with SPI-3 Sociability (0.46) and SPI-3 Emotionality (-0.70); and SPI-15 Impulsivity correlates with SPI-3 Social Cohesion (-0.47) and SPI-3 Emotionality (0.42).

3.3.3 Discussion

Attempts to identify the underlying structure of all 696 phrased items based on empirical fit statistics following exploratory factor analyses were largely inconclusive as these indicated drastically different values, including 3 (SRMR), 5 (complexity), 9 (MAP), 15 (complexity), and 27 factors (complexity and eBIC). Visual inspection of the factors at each level supported the development of scales at 3, 5, and 15 factors, but not 9 as the content of the factors at this level were not as clear or coherent. The prospect of creating scales using even more factors (for

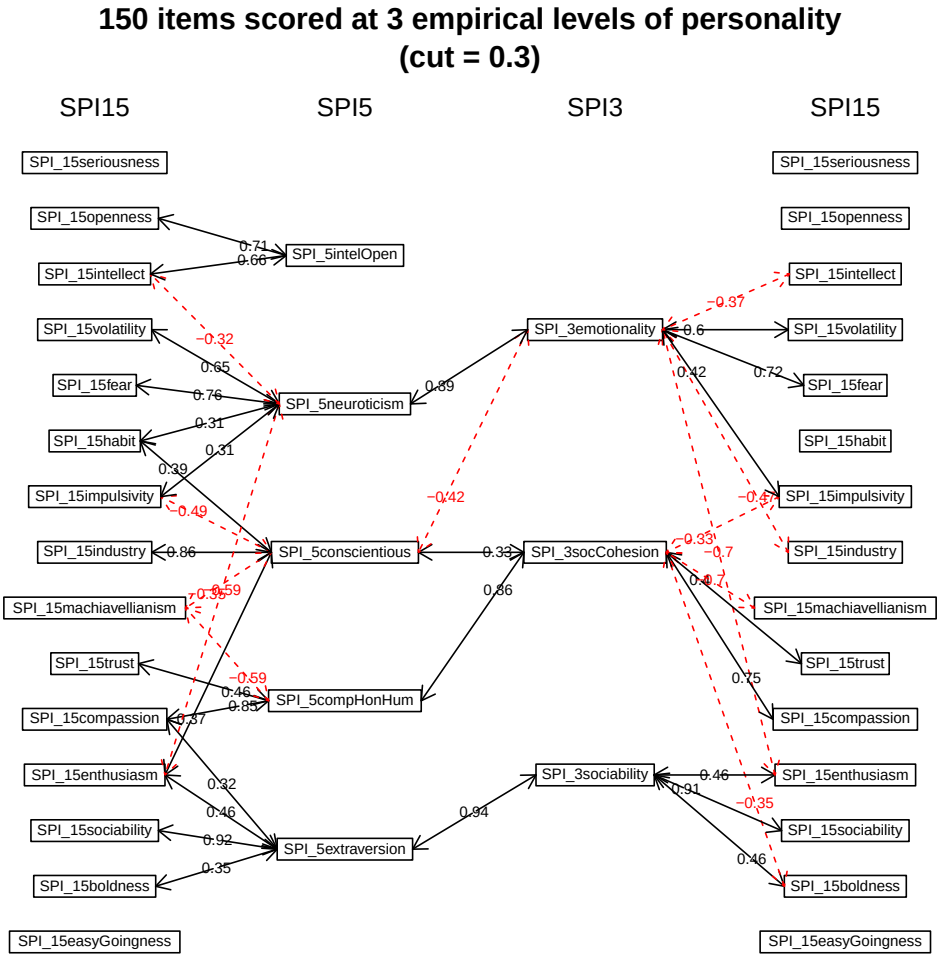


Figure 3.18: Scale correlations above 0.3 between the SPI-3, SPI-5, and SPI-15 scales after correcting for item overlap. The structure does not reflect strict hierarchical nesting though there are some scales which reflect very high correlations across levels such as Sociability/Extraversion and Compassion/Social Cohesion. Other scales (Seriousness and Easy-Goingness in the SPI-15 and Intellect/Openness in the SPI-5) provide incremental validity in that they are not highly correlated with higher level scales.

example, 27) was not supported as these led to one or more small factors, several of which did not have any items with primary factor loadings.

The primary limitation to these results is that a larger percentage of the public-domain phrased item universe was not administered. While the use of 696 items and 8 major scales is considerably more than are typically used in studies of this type, this represents fewer than one-sixth of the public-domain phrased items which have been identified to date and many more are likely available. It remains a priority to replicate the structure of the SPI-3/5/15 scales with larger item sets and additional samples.

The primary benefit of the SPI scales proposed here relative to existing scales are three-fold. First, they allow for assessment at multiple levels of specificity in a single measure using a relatively small number of total items (150). Second, they offer a noticeably cleaner relationship to both one another and existing scales at each level (e.g., relative to extant 3 and 5 factor scales). Finally, the SPI were directly and empirically derived on the basis of the structure of the *phrased item* universe rather than indirectly derived on the basis of the structure of trait-descriptors (i.e., the IPIP100) or the iterative evolution of theory (i.e., the NEO).

3.4 Conclusion

Given the opening remarks regarding the wide array of measures already available for personality assessment, the standards for introducing a novel set of scales should be quite high. This means that there should be at least some empirical evidence to suggest that extant measures do not describe the structure of personality well enough to preclude improvement and that there is a strong theoretical basis for deviating from the procedures which have been previously used for scale development.

To that end, the theoretical basis for the research described here stemmed from the possibility that the structure of the phrased item universe is different from the structure of the trait descriptor universe. While the indefinite size of the phrased item universe precludes the possibility of making definitive claims about its structure, the structure of the full set of items administered to this large international sample did not clearly support the superiority of a five (or six) factor structure as it apparently has in the trait descriptor universe. Regarding the potential for improvement of existing scales, the evidence from Study 1 suggested that the existing scales do not converge in their description of the primary dimensions despite the general presumption of consensus.

The SPI scales described here reflect both of these points and others. They are directly and empirically derived on the basis of administration of a large set of phrased items to an international sample. The fact that they demonstrate a relatively clean relationship with existing scales suggests that they provide a definitional "blend" for the constructs assessed by these existing measures (at least at the 3 and 5 factor levels). Perhaps most importantly, the SPI scales offer incremental utility over some measures in it that they allow for simultaneous assessment at 3 different levels. These levels are not explicitly designed to reflect strict nesting of constructs among levels. This is due to both the empirical methodology used to derive the SPI scales and the lack of evidence for strict nesting among the various factor analytic solutions.

As an alternative to the frequently-proposed hierarchical structure, the primary constructs of personality may be more usefully considered to demonstrate heterarchical structure. Unlike a strict hierarchy, a heterarchy allows for the possibility that the predominant organizational elements can shift depending on context. During a professional interview for a narrowly-defined office job, for example, the interviewer may be primarily concerned with the evaluation of a candidate's industriousness and impulsivity (on the SPI-15 level), slightly less interested in his

or her intellect/openness (on the SPI-5 level), and only superficially concerned with the candidate's sociability (on the SPI-3 level). In other contexts, such as when looking for a strong leader or a dinner date, the arrangement of salient constructs is likely to be entirely different. The qualities of industriousness and impulsivity can still be measured but the level of specificity to which they *should* be measured may depend on the extent to which they are relevant.

The SPI-3/5/15 allows for the potential to develop context-specific assessment models which simultaneously use different levels of specificity, and it does so with a relatively manageable number of items (150) for such an exhaustive assessment. In light of these characteristics and circumstances, it seems warranted to advocate for the use and further development of the SPI scales in hopes that they will offer a flexible and incrementally useful measure of the phrased-item universe in personality.

3.5 Appendix A – *The SAPA Personality Inventory (SPI3/5/15)*

Table 3.9: Appendix A - 150 items of the SPI scales by level

Item	SPI-15	SPI-5	SPI-3
Act without thinking.	impulsivity		
Am a creature of habit.	conventionality		
Am a person whose moods go up and down easily.		neuroticism	emotionality
Am a talkative person.		extraversion	sociability
Am a worrier.		neuroticism	emotionality
Am able to come up with new and different ideas.		intellect/openness	
Am concerned about others.	compassion	compassion/honesty/humility	social cohesion
Am easy to satisfy.	easy-going		
Am full of ideas.		intellect/openness	
Am happy with my life.	outlook		
Am indifferent to the feelings of others.		compassion/honesty/humility	social cohesion
Am more capable than most others.	intellect		
Am more easy-going about right and wrong than most people.	easy-going		
Am mostly quiet when with other people.	sociability	extraversion	sociability
Am not easily amused.	seriousness		
Am not easily disturbed by events.	fear/worry		
Am not interested in abstract ideas.	openness	intellect/openness	
Am often down in the dumps.	outlook	neuroticism	emotionality
Am out for my own personal gain.		compassion/honesty/humility	social cohesion
Am quick to understand things.	intellect	intellect/openness	
Am rather lively.		extraversion	sociability
Am sensitive to the needs of others.	compassion	compassion/honesty/humility	social cohesion
Am very pleased with myself.	outlook		
Am willing to take risks.	boldness		
Become overwhelmed by events.	fear/worry	neuroticism	
Begin to panic when there is danger.	fear/worry		
Believe in one true religion.	seriousness		
Believe in the importance of art.	openness	intellect/openness	
Believe that I am better than others.		compassion/honesty/humility	social cohesion
Believe that others have good intentions.	trusting		
Believe that people are basically moral.	trusting		
Can be stirred up easily.	volatility	neuroticism	emotionality
Can easily get some life into a dull party.		extraversion	sociability
Can handle a lot of information.	intellect	intellect/openness	

continued on next page

Table 3.9 – 150 items of the SPI scales by level (continued from previous page)

Item	SPI-15	SPI-5	SPI-3
Can handle complex problems.	intellect	intellect/openness	
Can't be bothered with other's needs.	compassion	compassion/honesty/humility	social cohesion
Cannot imagine lying or cheating.	machiavellianism		
Catch on to things quickly.	intellect		
Cheat to get ahead.	machiavellianism		social cohesion
Come up with good solutions.	intellect	intellect/openness	
Complete my duties as soon as possible.	industriousness	conscientiousness	
Dislike changes.	conventionality		
Dislike myself.	outlook		
Distrust people.	trusting		
Do things according to a plan.		conscientiousness	
Do things by the book.	conventionality	conscientiousness	
Do things without thinking of the consequences.	impulsivity		
Do too little work.		conscientiousness	
Don't like the idea of change.	conventionality		
Don't see the consequences of things.	impulsivity		
Don't talk a lot.	sociability	extraversion	sociability
Enjoy meeting new people.		extraversion	sociability
Enjoy thinking about things.	openness	intellect/openness	
Feel a sense of worthlessness or hopelessness.	outlook		emotionality
Feel others' emotions.	compassion	compassion/honesty/humility	
Feel sympathy for those who are worse off than myself.	compassion	compassion/honesty/humility	social cohesion
Feel that most people can't be trusted.	trusting	compassion/honesty/humility	
Find it difficult to get down to work.	industriousness	conscientiousness	emotionality
Find that it takes a lot to make me feel angry at someone.	volatility		
Get angry easily.	volatility		
Get caught up in my problems.		neuroticism	emotionality
Get chores done right away.	industriousness	conscientiousness	
Get easily agitated.	volatility	neuroticism	emotionality
Get irritated easily.	volatility	neuroticism	emotionality
Get overwhelmed by emotions.		neuroticism	emotionality
Get to work at once.	industriousness	conscientiousness	
Get upset easily.		neuroticism	emotionality
Have a dark outlook on the future.	outlook		
Have a low opinion of myself.	outlook		emotionality
Have a rich vocabulary.		intellect/openness	

continued on next page

Table 3.9 – 150 items of the SPI scales by level (continued from previous page)

Item	SPI-15	SPI-5	SPI-3
Have a soft heart.	compassion	compassion/honesty/humility	social cohesion
Have a vivid imagination.	openness	intellect/openness	
Have frequent mood swings.		neuroticism	emotionality
Have never engaged in gossip.	seriousness		
Inquire about others' well-being.	compassion	compassion/honesty/humility	social cohesion
Jump into things without thinking.	impulsivity		
Keep in the background.	sociability	extraversion	sociability
Keep things tidy.	industriousness	conscientiousness	
Know the answers to many questions.		intellect/openness	
Learn things slowly.	intellect		
Let myself be influenced by others.	trusting		emotionality
Let things proceed at their own pace.	easy-going		
Like going out a lot.			sociability
Like mixing with people.	sociability	extraversion	sociability
Like order.		conscientiousness	
Like to do frightening things.	boldness		social cohesion
Like to take it easy.	easy-going		
Look down on others.		compassion/honesty/humility	social cohesion
Lose my temper.	volatility		
Love action.	boldness		sociability
Love dangerous situations.	boldness		social cohesion
Love life.	outlook		
Love to chat.		extraversion	sociability
Love to reflect on things.	openness	intellect/openness	
Make careless mistakes.		conscientiousness	emotionality
Make friends easily.		extraversion	sociability
Make rash decisions.	impulsivity		
Most things taste the same to me.	seriousness		
My feelings are easily hurt.	fear/worry	neuroticism	emotionality
Need a creative outlet.	openness	intellect/openness	
Need protection.	fear/worry		
Need reassurance.		neuroticism	emotionality
Neglect my duties.	industriousness	conscientiousness	emotionality
Never splurge.	seriousness		
Panic easily.	fear/worry	neuroticism	
Pay too little attention to details.	impulsivity		

continued on next page

Table 3.9 – 150 items of the SPI scales by level (continued from previous page)

Item	SPI-15	SPI-5	SPI-3
People have said that I sometimes act rashly.	impulsivity		
Play a role in order to impress people.	machiavellianism		
Prefer reading to meeting people.	sociability	extraversion	sociability
Prefer to just let things happen.	easy-going		
Prefer to stick with things that I know.	conventionality		
Rarely overindulge.	seriousness		
Rarely put people under pressure.	easy-going		
Rarely show my anger.	volatility		
Rarely worry.		neuroticism	emotionality
Rush into things.	impulsivity		
Say little.			sociability
See that rules are observed.		conscientiousness	
Seek adventure.	boldness		sociability
Seek danger.	boldness		social cohesion
Seem to derive less enjoyment from interacting with people than others do.	sociability	extraversion	sociability
Seldom get mad.	volatility		
Seldom joke around.	seriousness		
Spend time reflecting on things.	openness	intellect/openness	
Start conversations.		extraversion	sociability
Start tasks right away.	industriousness	conscientiousness	
Stick to the rules.	conventionality	conscientiousness	social cohesion
Sympathize with others' feelings.	compassion	compassion/honesty/humility	social cohesion
Take advantage of others.	machiavellianism	compassion/honesty/humility	social cohesion
Take risks.	boldness		sociability
Take things as they come.	easy-going		
Talk to a lot of different people at parties.	sociability	extraversion	sociability
Tell a lot of lies.	machiavellianism		social cohesion
Tell other people what they want to hear so that they will do what I want them to do.	machiavellianism		social cohesion
Tend to keep in the background on social occasions.	sociability	extraversion	sociability
There are several people who keep trying to avoid me.	seriousness		
Think quickly.	intellect		
Tremble in dangerous situations.	fear/worry		
Trust others.	trusting		
Trust people to mainly tell the truth.	trusting		

continued on next page

Table 3.9 – 150 items of the SPI scales by level (continued from previous page)

Item	SPI-15	SPI-5	SPI-3
Trust what people say.	trusting		
Try to understand myself.	openness		
Use flattery to get ahead.	machiavellianism		
Use others for my own ends.	machiavellianism	compassion/honesty/humility	social cohesion
Value cooperation over competition.	easy-going		
Want everything to add up perfectly.	conventionality		
Want everything to be “just right.”	conventionality		
Waste my time.	industriousness	conscientiousness	emotionality
Worry a lot about my looks.	fear/worry		social cohesion
Would never go hang gliding or bungee jumping.	boldness		

Chapter 4

The cognitive domain: Developing a public-domain cognitive ability measure

4.1 Introduction

The domain of cognitive ability assessment is now populated with dozens, possibly hundreds, of proprietary measures (Camara et al., 2000; Carroll, 1993; Cattell, 1943a; Eliot and Smith, 1983; Goldstein and Beers, 2004; Murphy et al., 2011). While many of these are no longer maintained or administered, the variety of tests in active use remains quite broad, providing those who want to assess cognitive abilities with a large menu of options. In spite of this diversity, however, assessment challenges persist for researchers attempting to evaluate the structure and correlates of cognitive ability. We argue that it is possible to address these challenges through the use of well-established test development techniques and report on the development and validation of an item pool which demonstrates the utility of a public-domain measure of cognitive ability for basic intelligence research. We conclude by imploring other researchers to contribute to the on-going development, aggregation and maintenance of many more item types

as part of a broader, public-domain tool – the International Cognitive Ability Resource ("ICAR").

4.2 The Case For A Public Domain Measure

To be clear, the science of intelligence has historically been well-served by commercial measures. Royalty income streams (or their prospect) have encouraged the development of testing "products" and have funded their ongoing production, distribution and maintenance for decades. These assessments are broadly marketed for use in educational, counseling and industrial contexts and their administration and interpretation is a core service for many applied psychologists. Their proprietary nature is fundamental to the perpetuation of these royalty streams and to the privileged status of trained psychologists. For industrial and clinical settings, copyright-protected commercial measures offer clear benefits.

However, the needs of primary researchers often differ from those of commercial test users. These differences relate to issues of score interpretation, test content and administrative flexibility. In the case of score interpretation, researchers are considerably less concerned about the nature and quality of interpretative feedback. Unlike test-takers in selection and clinical settings, research participants are typically motivated by monetary rewards, course credit or, perhaps, a casual desire for informal feedback about their performance. This does not imply that researchers are less interested in quality norming data – it is often critical for evaluating the degree to which a sample is representative of a broader population. It simply means that, while many commercial testing companies have attempted to differentiate their products by providing materials for individual score interpretation, these materials have relatively little value for administration in research contexts.

The motivation among commercial testing companies to provide useful interpretative feedback

is directly related to test content however, and the nature of test content is of critical importance for intelligence researchers. The typical rationale for cognitive ability assessment in research settings is to evaluate the relationship between constructs and a broad range of other attributes. As such, the variety and depth of a test's content are very meaningful criteria for intelligence researchers – ones which are somewhat incompatible with the provision of meaningful interpretative feedback for each type of content. In other words, the ideal circumstance for many researchers would include the ability to choose from a variety of broadly-assessed cognitive ability constructs (or perhaps to choose a single measure which includes the assessment of a broad variety of constructs). While this ideal can sometimes be achieved through the administration of multiple commercial measures, this is rarely practical due to issues of cost and/or a lack of administrative flexibility.

The cost of administering commercial tests in research settings varies considerably across measures. While published rates are typically high, many companies allow for the qualified use of their copyright-protected materials at reduced rates or free-of-charge in research settings (e.g., the ETS Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976)). Variability in administration and scoring procedures is similarly high across measures. A small number of extant tests allow for brief, electronic assessment with automated scoring conducted within the framework of proprietary software, though none of these measures allow for customization of test content. The most commonly-used batteries are more arduous to administer, requiring one-to-one administration for over an hour followed by an additional 10 to 20 minutes for scoring (Camara et al., 2000). All too often, the result of the combination of challenges posed by these constraints is the omission of cognitive ability assessment in psychological research.

Several authors have suggested that the pace of scientific progress is diminished by reliance on proprietary measures (Gambardella and Hall, 2006; Goldberg, 1999; Liao et al., 2008). While it is

difficult to evaluate this claim empirically in the context of intelligence research, the circumstances surrounding development of the International Personality Item Pool ("IPIP") (Goldberg, 1999; Goldberg et al., 2006) provide a useful analogy. Prior to the development of the IPIP, personality researchers were forced to choose between validated but restrictive proprietary measures and a disorganized collection of narrow-bandwidth public-domain scales (these having been developed by researchers who were either unwilling to deal with copyright issues or whose needs were not met by the content of proprietary options). In the decade ending in 2012, at least 500 journal articles and book chapters using IPIP measures were published (Goldberg, 2014).

In fact, most of the arguments set forth in Goldberg's (1999) proposal for public-domain measures are directly applicable here. His primary point was that unrestricted use of public-domain instruments would make it less costly and difficult for researchers to administer scales which are flexible and widely-used. Secondary benefits would include a collaborative medium through which researchers could contribute to test development, refinement, and validation. The research community as a whole would benefit from an improved means of empirically comparing hypotheses across many diverse criteria.

Critics of the IPIP proposal expressed concern that a lack of copyright protection would impair the validity of personality measures (Goldberg et al., 2006). This argument would seem even more germane for tests of cognitive ability given the "maximal performance/typical behavior" distinction between intelligence and personality measures. The widely-shared presumption is that copyright restrictions on proprietary tests maintain validity by enhancing test security. Testing materials are, in theory, only disseminated to authorized users who have purchased licensed access and further dissemination is discouraged by the enforcement of intellectual property laws. Unfortunately, it is difficult to ascertain the extent to which test validity would be compromised in the general population without these safeguards. Concerns about disclosure

have been called into question with several prominent standardized tests (Field, 2012). There is also debate about the efficacy of intellectual property laws for protection against the unauthorized distribution of testing materials via the internet (Field, 2012; Kaufmann, 2009; McCaffrey and Lynch, 2009). Further evaluation of the relationship between copyright-protection and test validity seems warranted by these concerns, particularly for research applications where individual outcomes are less consequential.

Fortunately, copyright protection is not a prerequisite for test validity. Modern item-generation techniques (Arendasy et al., 2006; Dennis et al., 2002) present an alternate strategy that is less dependent on test security. Automatic item-generation makes use of algorithms which dictate the parameters of new items with predictable difficulty and in many alternate forms. These techniques allow for the creation of item types where the universe of *possible* items is very large. This, in turn, reduces the threat to validity that results from item disclosure. It can even be used to enhance test validity under administration paradigms that expose participants to sample items prior to testing and use alternate forms during assessment as this methodology reduces the effects of differential test familiarity across participants.

While automatic item-generation techniques represent the optimal method for developing public-domain cognitive ability items, this approach is often considerably more complicated than traditional development methods and it may be some time before a sizable number of automatically-generated item types is available for use in the public domain. For item types developed by traditional means, the maintenance of test validity depends on implementation of the more practical protocols used by commercial measures (i.e., those which do not invoke the credible threat of legal action). A public domain resource should set forth clear expectations for researchers regarding appropriate and ethical usage and make use of "warnings for nonprofessionals" (Goldberg et al., 2006). Sample test items should be made easily available to

the general public to further discourage wholesale distribution of testing materials. Given the current barriers to enforcement for intellectual property holders, these steps are arguably commensurate with protocols in place for copyright-protected commercial measures.

To the extent that traditional and automatic item-generation methods maintain adequate validity, there are many applications in which a non-proprietary measure would be useful. The most demanding of these applications would involve distributed, un-proctored assessments *in situ*, presumably conducted via online administration. Validity concerns would be most acute in these situations as there would be no safeguards against the use of external resources, including those available on the internet.

The remainder of this paper is dedicated to the evaluation of a public-domain measure developed for use under precisely these circumstances. This measure, the International Cognitive Ability Resource ("ICAR"), has been developed in stages over several years and further development is on-going. The first four item types (described below) were initially designed to provide an estimation of general cognitive ability for participants completing personality surveys at SAPA-Project.org, previously test.personality-project.org.

The primary goals when developing these initial item types were to: (1) briefly assess a small number of cognitive ability domains which were relatively distinct from one another (though considerable overlap between scores on the various types was anticipated); (2) avoid the use of "timed" items in light of potential technical issues resulting from telemetric assessment (Wilt et al., 2011); and (3) avoid item content that could be readily referenced elsewhere given the intended use of un-proctored online administrations. The studies described below were conducted to evaluate the degree to which these goals of item development were achieved.

The first study evaluated the item characteristics, reliability and structural properties of a

60-item ICAR measure. The second study evaluated the validity of the ICAR items when administered online in the context of self-reported achievement test scores and university majors. The third study evaluated the construct validity of the ICAR items when administered offline, using a brief commercial measure of cognitive ability.

4.3 Study 1

We investigated the structural properties of the initial version of the International Cognitive Ability Resource based on internet administration to a large international sample. This investigation was based on 60 items representing four item types developed in various stages since 2006 (and does not include deprecated items or item types currently under development). We hypothesized that the factor structure would demonstrate four distinct but highly correlated factors, with each type of item represented by a separate factor. This implied that, while individual items might demonstrate moderate or strong cross-loadings, the primary loadings would be consistent among items of each type.

4.3.1 Method

4.3.1.1. Participants

Participants were 96,958 individuals (66% female) from 199 countries who completed an online survey at SAPA-project.org (previously test.personality-project.org) between August 18, 2010 and May 20, 2013 in exchange for customized feedback about their personalities. All data were self-reported. The mean age was 26 years (*sd* = 10.6, median = 22) with a range from 14 to 90 years. Educational attainment levels for the participants are given in Table 4.1. Most participants were current university or secondary school students, although a wide range of

educational attainment levels were represented. Among the 75,740 participants from the United States (78.1%), 67.5% identified themselves as White/Caucasian, 10.3% as African-American, 8.5% as Hispanic-American, 4.8% as Asian-American, 1.1% as Native-American, and 6.3% as multi-ethnic (the remaining 1.5% did not specify). Participants from outside the United States were not prompted for information regarding race/ethnicity.

Table 4.1: Study 1 participants by educational attainment

Educational attainment	% of total	Mean age	Median age
Less than 12 years	14.5%	17.3	17
High school graduate	6.2%	23.7	18
Currently in college/university	51.4%	24.2	21
Some college/university, but did not graduate	5.0%	33.2	30
College/university degree	11.7%	33.2	30
Currently in graduate or professional school	4.4%	30.0	27
Graduate or professional school degree	6.9%	38.6	36

4.3.1.2. Measures

Four item types from the International Cognitive Ability Resource were administered, including: 9 Letter and Number Series items, 11 Matrix Reasoning items, 16 Verbal Reasoning items and 24 Three-Dimensional Rotation items. A 16 item subset of the measure, hereafter referred to as the *ICAR Sample Test*, is included as Appendix B. [1] Letter and Number Series items prompt participants with short digit or letter sequences and ask them to identify the next position in the sequence from among six choices. Matrix Reasoning items contain stimuli that are similar to those used in Raven's Progressive Matrices. The stimuli are 3x3 arrays of geometric shapes with one of the nine shapes missing. Participants are instructed to identify which of six geometric shapes presented as response choices will best complete the stimuli. The Verbal Reasoning items

[1] In addition to the sample items available in Appendix B, the remaining ICAR items can be accessed through ICAR-Project.org. A sample data set based on the items listed in Appendix B is also available ('iqitems') through the *psych* package (Revelle, 2014) in the R computing environment (R Core Team, 2014).

include a variety of logic, vocabulary and general knowledge questions. The Three-Dimensional Rotation items present participants with cube renderings and ask participants to identify which of the response choices is a possible rotation of the target stimuli. None of the items were timed in these administrations as untimed administration was expected to provide more stringent and conservative evaluation of the items' utility when given online (there are no specific reasons precluding timed administrations of the ICAR items, whether online or offline).

Participants were administered 12 to 16 item subsets of the 60 ICAR items using the Synthetic Aperture Personality Assessment ("SAPA") technique (Revelle et al., 2010b), a variant of matrix sampling procedures discussed by Lord (1955). The number of items administered to each participant varied over the course of the sampling period and was independent of participant characteristics. The number of administrations for each item varied considerably (median = 21,764) as did the number of pairwise administrations between any two items in the set (median = 2,610). This variability reflected the introduction of newly developed items over time and the fact that item sets include unequal numbers of items. The minimum number of pairwise administrations among items (422) provided sufficiently high stability in the covariance matrix for the structural analyses described below (Kenny, 2012).

4.3.1.3. Analyses

Internal consistency measures were assessed by using the Pearson correlations between ICAR items to calculate α, ω_h, and ω_{total} reliability coefficients (Revelle, 2014; Revelle and Zinbarg, 2009; Zinbarg et al., 2005). The use of tetrachoric correlations for reliability analyses is discouraged on the grounds that it typically over-estimates both alpha and omega (Revelle and Condon, 2012).

Two latent variable exploratory factor analyses ("EFA") were conducted to evaluate the structure

of the ICAR items. The first of these included all 60 items (9 Letter and Number Series items, 11 Matrix Reasoning items, 16 Verbal Reasoning items and 24 Three-Dimensional Rotation items). A second EFA was required to address questions regarding the structural impact of including disproportionate numbers of items by type. This was done by using only the subset of participants (n = 4,574) who were administered the 16 item *ICAR Sample Test*. This subset included four items each from the four ICAR item types. These items were selected as a representative set on the basis of their difficulty relative to the full set of 60 items and their factor loadings relative to other items of the same type. Note that the factor analysis of this 16 item subset was not independent from that conducted on the full 60 item set. EFA results were then used to evaluate the omega hierarchical general factor saturation (Revelle and Zinbarg, 2009; Zinbarg et al., 2006) of the 16 item *ICAR Sample Test*.

Both of these exploratory factor analyses were based on the Pearson correlations between scored responses using Ordinary Least Squares ("OLS") regression models with oblique rotation (Revelle, 2014). The factoring method used here minimizes the χ^2 value rather than minimizing the sum of the squared residual values (as is done by default with most statistical software). Note that in cases where the number of administrations is consistent across items, as with the 16 item *ICAR Sample Test*, these methods are identical. The methods differ in cases where the number of pairwise administrations between items varies because the squared residuals are weighted by sample size rather than assumed to be equivalent across variables. Goodness-of-fit was evaluated using the Root Mean Square of the Residual, the Root Mean Squared Error of Approximation (Hu and Bentler, 1999), and the Tucker Lewis Index of factoring reliability (Kenny, 2012; Tucker and Lewis, 1973).

Analyses based on two-parameter Item Response Theory (Baker, 1985; Embretson, 1996; Revelle, 2014) were used to evaluate the unidimensional relationships between items on several levels,

including (1) all 60 items, (2) each of the four item types independently, and (3) for the 16 item *ICAR Sample Test*. In these cases, the tetrachoric correlations between items were used. These procedures allow for estimation of the correlations between items as if they had been measured continuously (Uebersax, 2000).

4.3.2 Results

Descriptive statistics for all 60 ICAR items are given in Table 4.2. Mean values indicate the proportion of participants who provided the correct response for an item relative to the total number of participants who were administered that item. The Three-Dimensional Rotation items had the lowest proportion of correct responses (m = 0.19, sd = 0.08), followed by Matrix Reasoning (m = 0.52, sd = 0.15), then Letter and Number Series (m = 0.59, sd = 0.13), and Verbal Reasoning (m = 0.64, sd = 0.22). Internal consistencies for the ICAR item types are given in Table 4.3. These values are based on the composite correlations between items as individual participants completed only a subset of the items (as is typical when using SAPA sampling procedures).

Results from the first exploratory factor analysis using all 60 items suggested factor solutions of three to five factors based on inspection of the scree plots in Figure 4.1. The fit statistics were similar for each of these solutions. The four factor model was slightly superior in fit (RMSEA = 0.058, RMSR = 0.05) and reliability (TLI = 0.71) to the three factor model (RMSEA = 0.059, RMSR = 0.05, TLI = 0.7) and was slightly inferior to the five factor model (RMSEA = 0.055, RMSR = 0.05, TLI = 0.73). Factor loadings and the correlations between factors for each of these solutions are included in Tables 4.4 to 4.9.

Table 4.2: Descriptive statistics for the ICAR items administered in Study 1

Item	*n*	*mean*	*sd*	Item	*n*	*mean*	*sd*
LN.01	31,239	0.79	0.41	R3D.11	7,165	0.09	0.29
LN.03	31,173	0.59	0.49	R3D.12	7,168	0.13	0.34
LN.05	31,486	0.75	0.43	R3D.13	7,291	0.10	0.30
LN.06	34,097	0.46	0.50	R3D.14	7,185	0.14	0.35
LN.07	*36,346*	*0.62*	*0.49*	R3D.15	7,115	0.22	0.42
LN.33	*39,384*	*0.59*	*0.49*	R3D.16	7,241	0.30	0.46
LN.34	*36,655*	*0.62*	*0.48*	R3D.17	7,085	0.15	0.36
LN.35	34,372	0.47	0.50	R3D.18	6,988	0.13	0.34
LN.58	*39,047*	*0.42*	*0.49*	R3D.19	7,103	0.16	0.37
MR.43	29,812	0.77	0.42	R3D.20	7,203	0.39	0.49
MR.44	17,389	0.66	0.47	R3D.21	7,133	0.08	0.28
MR.45	*24,689*	*0.52*	*0.50*	R3D.22	7,369	0.30	0.46
MR.46	*34,952*	*0.60*	*0.49*	R3D.23	7,210	0.19	0.39
MR.47	*34,467*	*0.62*	*0.48*	R3D.24	7,000	0.19	0.39
MR.48	17,450	0.53	0.50	*VR.04*	*29,975*	*0.67*	*0.47*
MR.50	19,155	0.28	0.45	VR.09	25,402	0.70	0.46
MR.53	29,548	0.61	0.49	VR.11	26,644	0.86	0.35
MR.54	19,246	0.39	0.49	VR.13	24,147	0.24	0.43
MR.55	*24,430*	*0.36*	*0.48*	VR.14	26,100	0.74	0.44
MR.56	19,380	0.40	0.49	*VR.16*	*31,727*	*0.69*	*0.46*
R3D.01	7,537	0.08	0.28	*VR.17*	*31,552*	*0.73*	*0.44*
R3D.02	7,473	0.16	0.37	VR.18	26,474	0.96	0.20
R3D.03	*12,701*	*0.17*	*0.37*	*VR.19*	*30,556*	*0.61*	*0.49*
R3D.04	*12,959*	*0.21*	*0.41*	VR.23	24,928	0.27	0.44
R3D.05	7,526	0.24	0.43	VR.26	13,108	0.38	0.49
R3D.06	*12,894*	*0.29*	*0.46*	VR.31	26,272	0.90	0.30
R3D.07	7,745	0.12	0.33	VR.32	25,419	0.55	0.50
R3D.08	*12,973*	*0.17*	*0.37*	VR.36	25,076	0.40	0.49
R3D.09	7,244	0.28	0.45	VR.39	26,433	0.91	0.28
R3D.10	7,350	0.14	0.35	VR.42	25,108	0.66	0.47

Note: “LN" denotes Letter and Number Series, “MR" is Matrix Reasoning, “R3D” is Three-Dimensional Rotation, and “VR" is Verbal Reasoning. Italicized items denote those included in the 16-Item *ICAR Sample Test.*

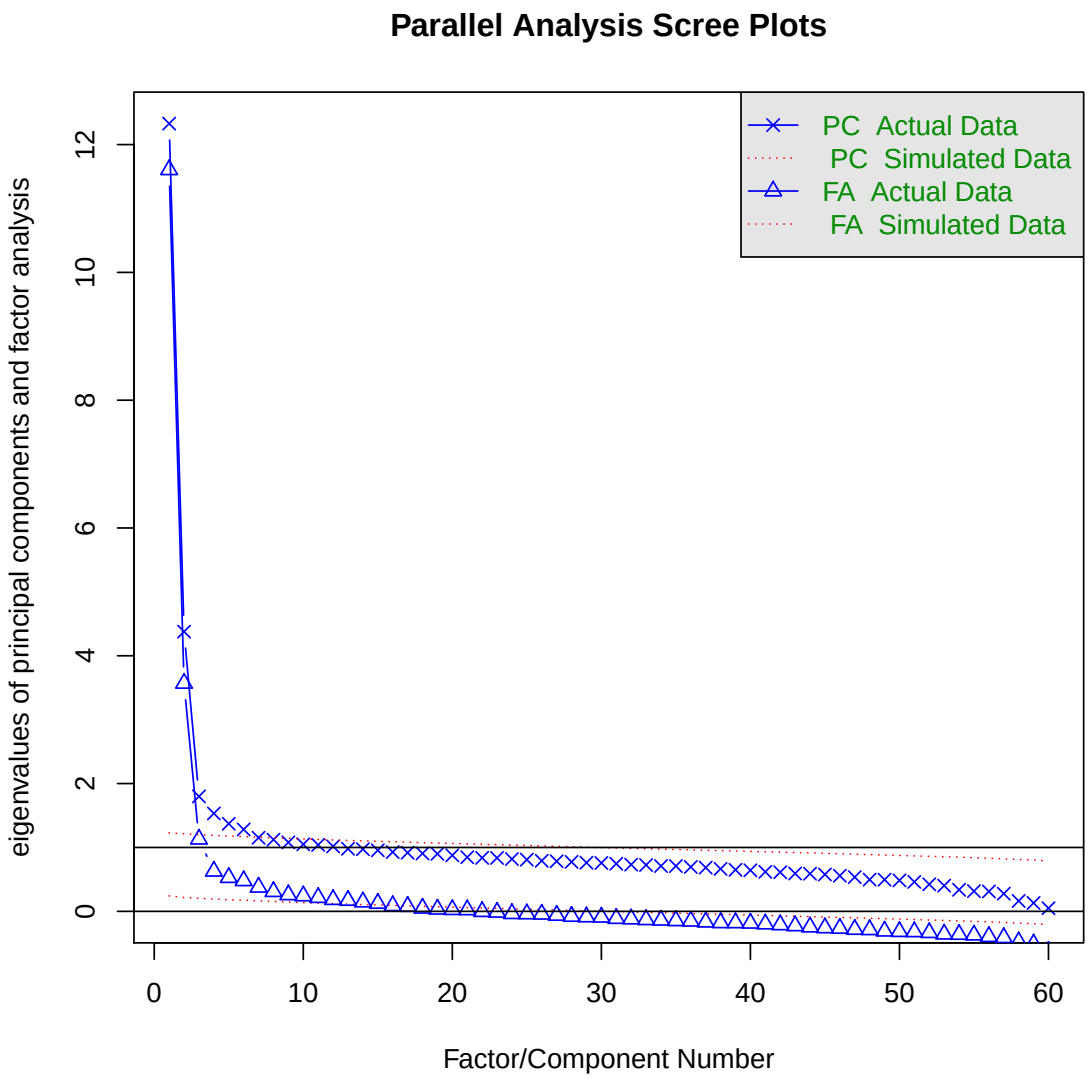


Figure 4.1: Scree plots based on factoring of all 60 ICAR items

Table 4.3: Alpha and omega for the ICAR item types

	α	ω_h	ω_t	items
ICAR60	0.93	0.61	0.94	60
LN items	0.77	0.66	0.80	9
MR items	0.68	0.58	0.71	11
R3D items	0.93	0.78	0.94	24
VR items	0.76	0.64	0.77	16
ICAR16	0.81	0.66	0.83	16

Note: ω_h = omega hierarchical, ω_t = omega total. Values are based on composites of Pearson correlations between items.

Table 4.4: Three factor solution based on all 60 ICAR items

ICAR item	Factor 1	Factor 2	Factor 3
LN.01	**0.54**	-0.01	-0.03
VR.17	**0.54**	0.00	0.00
VR.04	**0.53**	-0.01	0.05
LN.07	**0.52**	-0.01	0.00
LN.34	**0.52**	0.01	0.01
VR.14	**0.52**	0.00	0.03
LN.03	**0.51**	0.04	0.03
LN.58	**0.47**	0.05	0.06
VR.19	**0.46**	0.01	0.02
VR.16	**0.46**	0.00	0.02
LN.33	**0.45**	0.02	0.03
LN.05	**0.45**	0.01	-0.03
VR.31	**0.43**	-0.03	-0.07
VR.32	**0.42**	0.04	0.00
MR.47	**0.41**	0.04	0.06
LN.06	**0.39**	0.06	0.07
MR.43	**0.38**	0.00	0.04
MR.46	**0.37**	0.01	0.06
VR.11	**0.35**	-0.01	-0.01
LN.35	**0.35**	0.06	0.05
MR.45	**0.35**	0.03	0.07
VR.09	**0.32**	-0.03	0.01
VR.36	**0.31**	0.06	0.08
VR.39	**0.30**	0.02	-0.06
MR.53	**0.30**	0.03	0.05
VR.42	**0.29**	-0.01	0.04
VR.23	**0.28**	0.06	0.10
MR.54	**0.28**	-0.02	0.05
MR.44	**0.26**	0.00	0.12
MR.56	**0.24**	0.04	0.08
VR.13	**0.23**	0.01	0.10
VR.26	**0.23**	0.01	0.03
MR.48	**0.22**	0.04	0.09
MR.55	**0.21**	0.09	0.11
VR.18	**0.21**	-0.01	-0.04
MR.50	**0.19**	0.10	0.10
R3D.14	0.13	**0.85**	-0.16

continued on next page

Table 4.4 – continued from previous page

ICAR item	Factor 1	Factor 2	Factor 3
R3D.07	0.03	**0.83**	-0.07
R3D.19	0.00	**0.70**	0.02
R3D.21	0.05	**0.62**	-0.02
R3D.11	-0.02	**0.55**	0.14
R3D.02	-0.07	**0.51**	0.31
R3D.18	-0.03	**0.50**	0.30
R3D.17	0.01	**0.46**	0.32
R3D.12	-0.02	**0.46**	0.21
R3D.13	-0.02	**0.41**	0.26
R3D.01	-0.13	**0.41**	0.28
R3D.15	0.02	**0.33**	0.32
R3D.10	-0.03	**0.31**	0.28
R3D.06	0.07	-0.06	**0.74**
R3D.22	0.12	-0.08	**0.69**
R3D.05	0.06	-0.01	**0.66**
R3D.09	0.07	0.00	**0.61**
R3D.16	0.07	0.00	**0.61**
R3D.20	0.19	-0.04	**0.53**
R3D.24	-0.01	0.32	**0.49**
R3D.04	0.07	0.22	**0.49**
R3D.08	-0.01	0.29	**0.45**
R3D.03	-0.01	0.38	**0.40**
R3D.23	0.07	0.27	**0.36**
SS Loadings	5.88	5.38	5.37
% of Variance	0.35	0.32	0.32
Score Correlation	0.94	0.95	0.94

Fit statistics: RMSR = 0.05; RMSEA = 0.059;
TLI = 0.70

Table 4.5: Correlations between factors for the three factor solution

	Factor 1	Factor 2	Factor 3
Factor 1	1.00		
Factor 2	0.27	1.00	
Factor 3	0.43	0.54	1.00

Table 4.6: Four factor solution based on all 60 ICAR items

ICAR item	Factor 1	Factor 2	Factor 3	Factor 4
LN.07	**0.56**	0.02	-0.06	-0.03
LN.34	**0.55**	0.02	-0.03	-0.01
LN.01	**0.54**	-0.06	-0.01	0.01
VR.04	**0.53**	-0.04	0.06	0.01
VR.17	**0.53**	-0.09	0.03	0.06
LN.03	**0.52**	0.04	0.01	-0.02
VR.14	**0.52**	-0.04	0.05	0.01
LN.58	**0.49**	0.04	0.02	0.04
LN.05	**0.48**	0.02	-0.08	-0.01
LN.33	**0.48**	0.05	-0.01	-0.04
VR.19	**0.46**	-0.07	0.06	0.06
VR.16	**0.45**	-0.02	0.04	-0.02
VR.31	**0.43**	-0.10	-0.02	0.01
VR.32	**0.43**	0.01	0.00	0.02
MR.47	**0.42**	0.06	0.03	-0.02
LN.06	**0.41**	0.11	0.02	-0.04
MR.46	**0.38**	0.06	0.03	-0.06
MR.43	**0.38**	0.00	0.04	-0.02
LN.35	**0.36**	0.04	0.02	0.04
MR.45	**0.35**	0.02	0.06	0.00
VR.11	**0.35**	-0.04	0.02	-0.01
VR.09	**0.32**	-0.04	0.02	-0.01
MR.53	**0.31**	0.06	0.04	-0.04
VR.39	**0.30**	-0.02	-0.03	0.00
VR.36	**0.30**	-0.01	0.09	0.08
VR.42	**0.30**	0.03	0.01	-0.05
VR.23	**0.29**	0.09	0.06	-0.01
MR.54	**0.28**	0.00	0.05	-0.03
MR.44	**0.27**	0.02	0.10	0.00
VR.13	**0.25**	0.11	0.05	-0.11
MR.56	**0.24**	0.08	0.05	-0.02
VR.26	**0.24**	0.02	0.02	-0.02
MR.48	**0.23**	0.08	0.06	-0.03
MR.55	**0.22**	0.09	0.08	0.02
MR.50	**0.21**	0.19	0.04	-0.09
VR.18	**0.21**	-0.05	-0.02	0.03
R3D.19	0.05	**0.77**	-0.09	-0.06

continued on next page

Table 4.6 – continued from previous page

ICAR item	Factor 1	Factor 2	Factor 3	Factor 4
R3D.07	0.06	**0.73**	-0.10	0.07
R3D.02	-0.03	**0.63**	0.15	0.00
R3D.18	0.01	**0.56**	0.14	0.08
R3D.01	-0.08	**0.50**	0.10	0.10
R3D.14	0.11	**0.49**	-0.11	0.37
R3D.10	0.01	**0.46**	0.15	-0.07
R3D.24	0.03	**0.45**	0.32	0.02
R3D.03	0.01	**0.44**	0.30	0.00
R3D.12	0.00	**0.44**	0.11	0.15
R3D.08	0.02	**0.42**	0.29	0.02
R3D.11	-0.02	**0.39**	0.08	0.28
R3D.17	0.02	**0.36**	0.23	0.25
R3D.23	0.10	**0.33**	0.22	0.09
R3D.13	0.00	**0.33**	0.16	0.24
R3D.06	0.03	0.02	**0.71**	0.00
R3D.22	0.07	-0.06	**0.69**	0.05
R3D.05	0.03	0.10	**0.64**	-0.06
R3D.09	0.02	0.03	**0.63**	0.01
R3D.16	0.03	0.01	**0.60**	0.05
R3D.20	0.14	-0.05	**0.57**	0.04
R3D.04	0.06	0.29	**0.42**	-0.01
R3D.15	0.01	0.28	**0.29**	0.11
R3D.21	0.01	0.00	0.01	**0.91**
SS Loadings	6.01	5.17	4.24	1.71
% of Variance	0.35	0.30	0.25	0.10
Score Correlation	0.94	0.94	0.92	0.93

Fit statistics: RMSR = 0.05; RMSEA = 0.058; TLI = 0.71

Table 4.7: Correlations between factors for the four factor solution

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1.00			
Factor 2	0.30	1.00		
Factor 3	0.49	0.57	1.00	
Factor 4	0.24	0.54	0.31	1.00

Table 4.8: Five factor solution based on all 60 ICAR items

ICAR item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
LN.07	**0.56**	-0.05	0.03	-0.04	-0.01
VR.04	**0.55**	0.02	-0.01	-0.02	0.01
LN.34	**0.55**	-0.03	0.01	-0.01	0.00
LN.01	**0.54**	-0.02	-0.06	0.00	0.00
VR.17	**0.54**	0.00	-0.04	-0.02	0.04
LN.03	**0.53**	-0.01	0.03	0.02	-0.02
VR.14	**0.52**	0.02	-0.03	-0.01	0.00
LN.58	**0.49**	0.02	0.03	-0.01	0.05
LN.05	**0.48**	-0.07	0.02	-0.02	0.01
LN.33	**0.48**	-0.01	0.01	0.02	-0.03
VR.19	**0.46**	0.04	-0.05	-0.02	0.04
VR.16	**0.46**	0.02	-0.04	0.04	-0.04
MR.47	**0.44**	0.01	0.05	0.02	-0.02
VR.32	**0.43**	-0.01	-0.01	0.03	0.00
VR.31	**0.42**	-0.01	-0.13	0.02	-0.01
LN.06	**0.41**	0.02	0.07	0.03	-0.03
MR.43	**0.39**	0.01	0.01	0.00	-0.02
MR.46	**0.39**	0.03	0.00	0.06	-0.06
MR.45	**0.37**	0.03	0.03	0.01	0.00
LN.35	**0.37**	0.02	0.02	0.01	0.05
VR.11	**0.35**	0.01	-0.05	0.02	-0.03
VR.09	**0.32**	0.02	-0.05	0.00	-0.02
VR.42	**0.32**	-0.02	0.05	-0.02	-0.04
VR.23	**0.31**	0.02	0.12	-0.02	-0.01
MR.53	**0.31**	0.05	-0.01	0.06	-0.04
VR.36	**0.30**	0.09	-0.04	0.04	0.07
VR.39	**0.29**	-0.02	-0.08	0.06	-0.02
MR.54	**0.29**	0.03	0.00	0.00	-0.03
MR.44	**0.28**	0.07	0.06	-0.03	0.01
VR.13	**0.26**	0.03	0.09	0.02	-0.10
MR.56	**0.26**	0.03	0.07	0.01	-0.01
VR.26	**0.25**	-0.02	0.05	-0.02	-0.01
MR.48	**0.25**	0.04	0.07	0.02	-0.03
MR.55	**0.23**	0.07	0.07	0.03	0.03
MR.50	**0.22**	0.04	0.11	0.09	-0.07
VR.18	**0.22**	-0.04	-0.03	-0.02	0.02
R3D.06	-0.01	**0.78**	0.00	-0.03	-0.01

continued on next page

Table 4.8 – continued from previous page

ICAR item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
R3D.22	0.06	**0.72**	0.00	-0.10	0.04
R3D.09	-0.01	**0.69**	-0.04	0.05	-0.02
R3D.05	0.00	**0.69**	0.06	0.02	-0.09
R3D.16	-0.01	**0.67**	-0.04	0.02	0.05
R3D.20	0.15	**0.55**	-0.01	-0.03	0.03
R3D.04	0.06	**0.45**	0.17	0.16	-0.02
R3D.15	-0.01	**0.35**	0.11	0.21	0.07
R3D.03	0.01	**0.30**	0.28	0.27	-0.02
R3D.01	0.00	-0.06	**0.68**	-0.02	0.14
R3D.24	0.08	0.24	**0.54**	-0.01	0.06
R3D.18	0.06	0.07	**0.52**	0.17	0.10
R3D.02	-0.01	0.15	**0.47**	0.24	0.01
R3D.23	0.15	0.14	**0.43**	-0.05	0.14
R3D.08	0.06	0.24	**0.42**	0.05	0.06
R3D.13	0.04	0.08	**0.42**	0.04	0.25
R3D.10	0.05	0.08	**0.41**	0.14	-0.04
R3D.12	0.01	0.13	**0.31**	0.19	0.14
R3D.11	-0.01	0.09	**0.28**	0.22	0.25
R3D.17	0.02	0.26	**0.26**	0.17	0.24
R3D.14	0.04	0.05	-0.17	**0.85**	0.23
R3D.07	0.03	-0.02	0.21	**0.72**	-0.05
R3D.19	0.02	0.01	0.25	**0.66**	-0.14
R3D.21	0.01	-0.01	0.08	0.05	**0.88**
SS Loadings	6.11	4.34	3.54	2.76	1.47
% of Variance	0.34	0.24	0.19	0.15	0.08
Score Correlation	0.94	0.94	0.92	0.95	0.93

Fit statistics: RMSR = 0.05; RMSEA = 0.058; TLI = 0.71

Table 4.9: Correlations between factors for the five factor solution

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	1.00				
Factor 2	0.53	1.00			
Factor 3	0.24	0.53	1.00		
Factor 4	0.26	0.41	0.48	1.00	
Factor 5	0.22	0.29	0.30	0.47	1.00

The second EFA, based on a balanced number of items by type, demonstrated very good fit for the four-factor solution (RMSEA = 0.014, RMSR = 0.01, TLI = 0.99). Factor loadings by item for the four-factor solution are shown in Table 4.10. Each of the item types was represented by a different factor and the cross-loadings were small. Correlations between factors (Table 4.11) ranged from 0.41 to 0.70.

Table 4.10: Four-factor item loadings for the *ICAR Sample Test*

Item	Factor 1	Factor 2	Factor 3	Factor 4
R3D.03	**0.69**	-0.02	-0.04	0.01
R3D.08	**0.67**	-0.04	-0.01	0.02
R3D.04	**0.66**	0.03	0.01	0.00
R3D.06	**0.59**	0.06	0.07	-0.02
LN.34	-0.01	**0.68**	-0.01	-0.02
LN.07	-0.03	**0.60**	-0.01	0.05
LN.33	0.04	**0.52**	0.01	0.00
LN.58	0.08	**0.43**	0.07	0.01
VR.17	-0.04	0.00	**0.65**	-0.02
VR.04	0.06	-0.01	**0.51**	0.05
VR.16	0.02	0.05	**0.41**	0.00
VR.19	0.03	0.02	**0.38**	0.06
MR.45	-0.02	-0.01	0.01	**0.56**
MR.46	0.02	0.02	0.01	**0.50**
MR.47	0.05	0.18	0.10	**0.24**
MR.55	0.14	0.09	-0.04	**0.21**

Table 4.11: Correlations between factors for the *ICAR Sample Test*

	R3D Factor	LN Factor	VR Factor	MR Factor
R3D Factor	1.00			
LN Factor	0.44	1.00		
VR Factor	0.70	0.45	1.00	
MR Factor	0.63	0.41	0.59	1.00

Note: R3D = Three-Dimensional Rotation, LN = Letter and Number Series, VR = Verbal Reasoning, MR = Matrix Reasoning

General factor saturation for the 16 item *ICAR Sample Test* is depicted in Figures 4.2 and 4.3. Figure 4.2 shows the primary factor loadings for each item consistent with the values presented in Table 4.10 and also shows the general factor loading for each of the second-order factors. Figure 4.3 shows the general factor loading for each item and the residual loading of each item to its primary second-order factor after removing the general factor.

The results of IRT analyses for the 16 item *ICAR Sample Test* are presented in Table 4.12 as well as Figures 4.4 and 4.5. Table 4.12 provides item information across levels of the latent trait and summary information for the test as a whole. The item information functions are depicted graphically in Figure 4.4. Figure 4.5 depicts the test information function for the *ICAR Sample Test* as well as reliability in the vertical axis on the right (reliability in this context is calculated as one minus the reciprocal of the test information). The results of IRT analyses for the full 60 item set and for each of the item types independently are given in Tables 4.13 to 4.17. The pattern of results was similar to those for the *ICAR Sample Test* in terms of the relationships between item types and the spread of item difficulties across levels of the latent trait, though the reliability was higher for the full 60 item set across the range of difficulties (Figure 4.6).

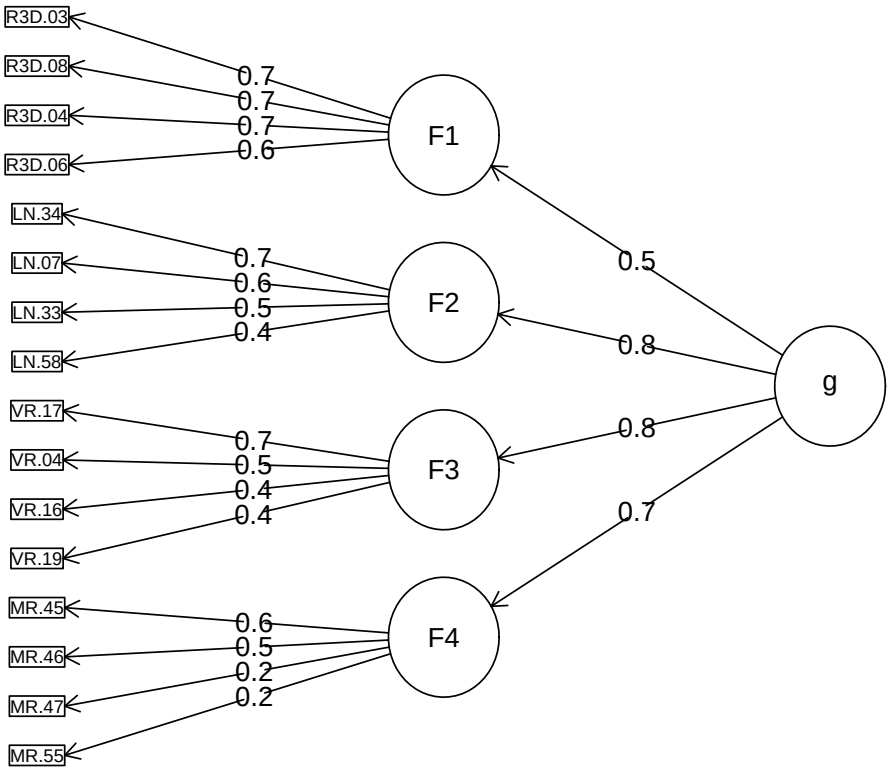


Figure 4.2: Omega hierarchical for the *ICAR Sample Test*

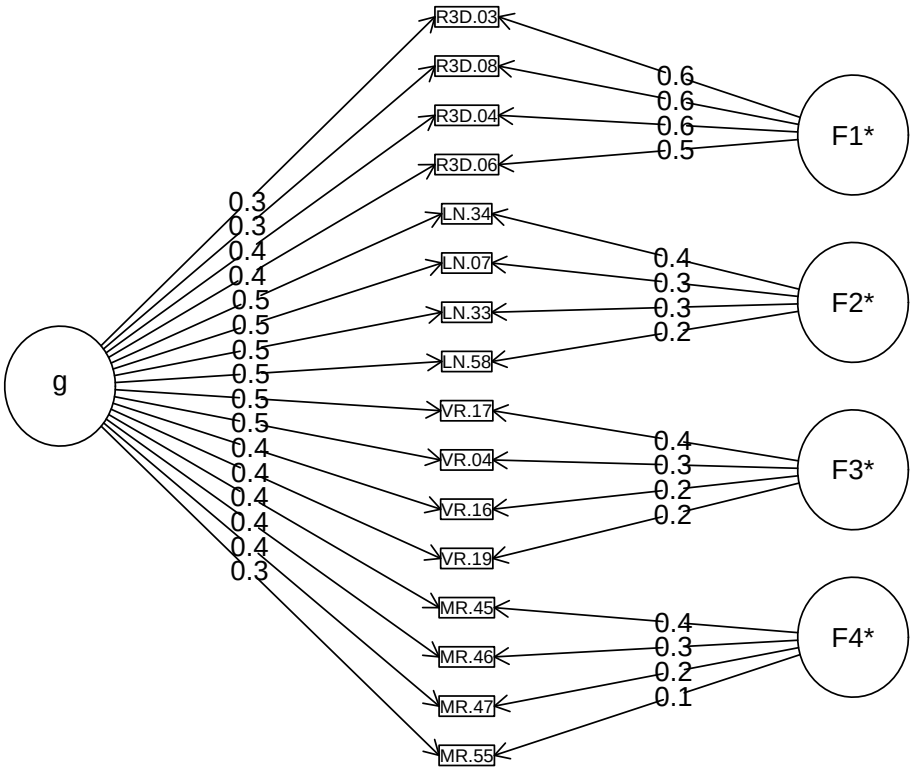


Figure 4.3: Omega with Schmid-Leiman transformation for the *ICAR Sample Test*

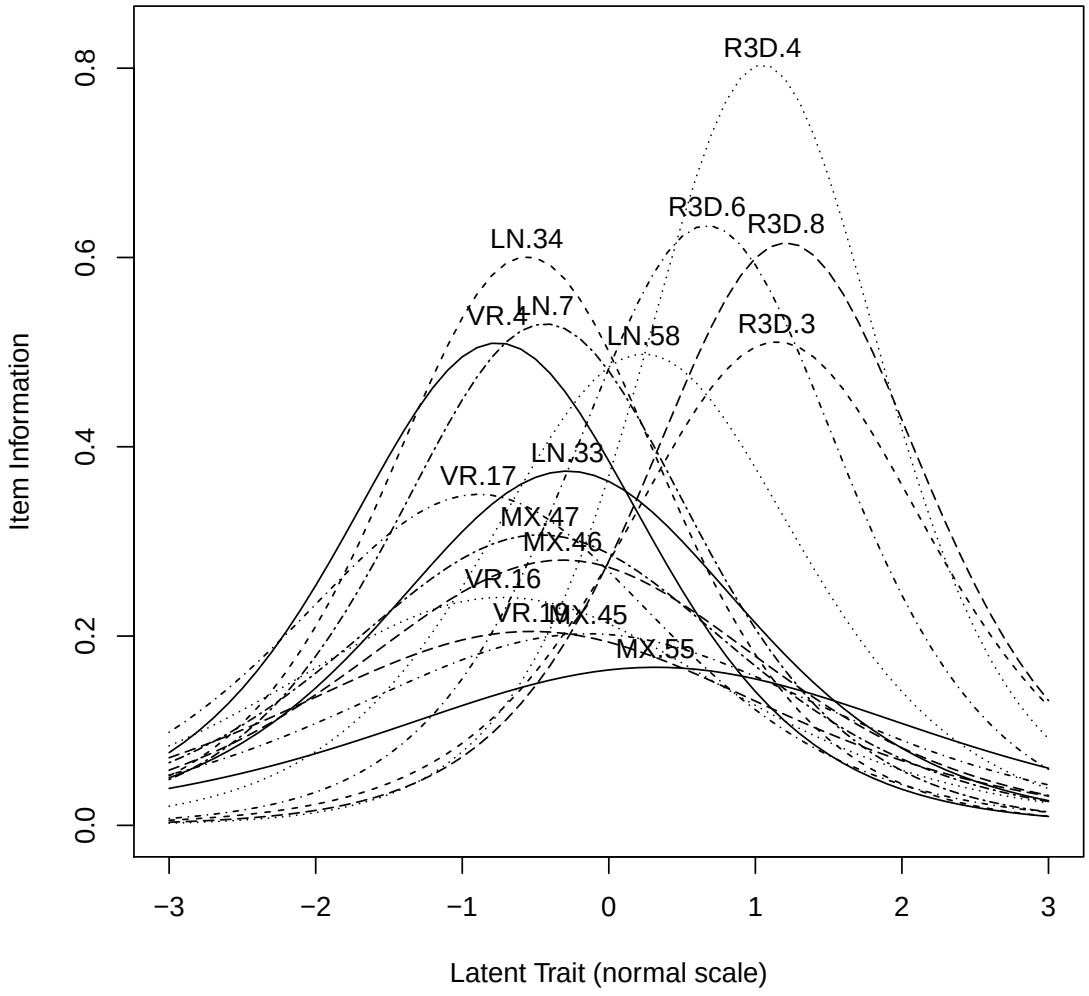


Figure 4.4: Item Information Functions for the 16 item *ICAR Sample Test*

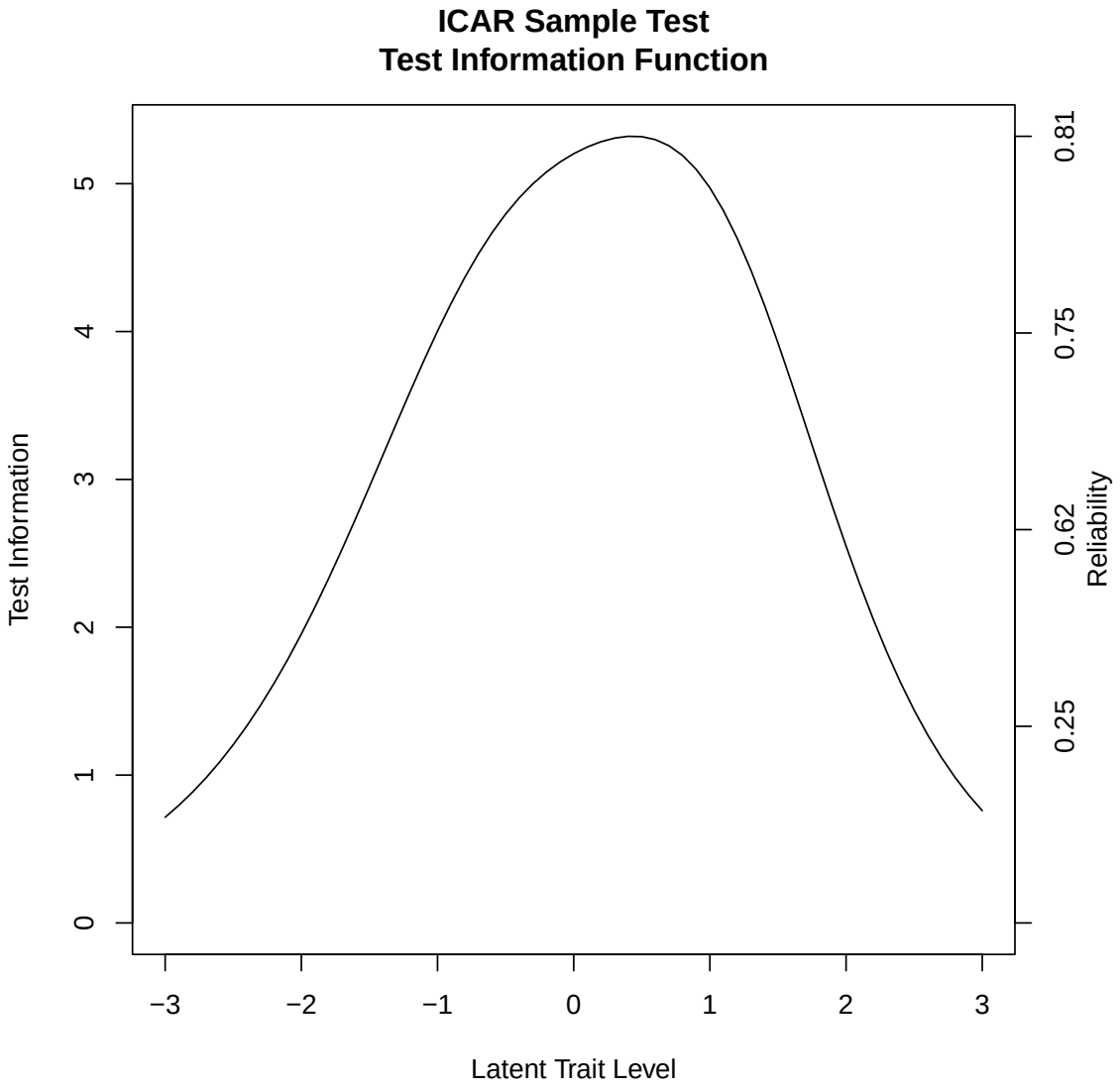


Figure 4.5: Test Information Function for the 16 item *ICAR Sample Test*

Table 4.12: Item and test information for the 16 item *ICAR Sample Test*

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
VR.04	0.07	0.23	0.49	0.42	0.16	0.04	0.01
VR.16	0.08	0.17	0.25	0.23	0.13	0.06	0.02
VR.17	0.09	0.27	0.46	0.34	0.13	0.04	0.01
VR.19	0.07	0.14	0.24	0.25	0.16	0.07	0.03
LN.07	0.06	0.18	0.38	0.39	0.19	0.06	0.02
LN.33	0.05	0.15	0.32	0.37	0.21	0.08	0.02
LN.34	0.05	0.20	0.46	0.45	0.19	0.05	0.01
LN.58	0.03	0.09	0.26	0.43	0.32	0.13	0.04
MR.45	0.05	0.11	0.17	0.20	0.16	0.09	0.04
MR.46	0.06	0.13	0.22	0.24	0.17	0.08	0.04
MR.47	0.06	0.16	0.31	0.32	0.18	0.07	0.02
MR.55	0.04	0.07	0.11	0.14	0.13	0.10	0.06
R3D.03	0.00	0.01	0.06	0.27	0.64	0.47	0.14
R3D.04	0.00	0.01	0.07	0.35	0.83	0.45	0.10
R3D.06	0.00	0.03	0.14	0.53	0.73	0.26	0.05
R3D.08	0.00	0.01	0.06	0.26	0.64	0.48	0.14
TIF	0.72	1.95	4.00	5.20	4.97	2.55	0.76
SEM	1.18	0.72	0.50	0.44	0.45	0.63	1.15
Reliability	NA	0.49	0.75	0.81	0.80	0.61	NA

Table 4.13: Item and test information for the 60 ICAR items

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
LN.01	0.11	0.23	0.31	0.24	0.11	0.04	0.01
LN.03	0.05	0.14	0.31	0.37	0.22	0.08	0.03
LN.05	0.08	0.14	0.18	0.16	0.11	0.06	0.03
LN.06	0.04	0.1	0.19	0.26	0.21	0.12	0.05
LN.07	0.06	0.14	0.23	0.25	0.16	0.08	0.03
LN.33	0.06	0.12	0.21	0.24	0.17	0.09	0.04
LN.34	0.06	0.15	0.27	0.28	0.18	0.08	0.03
LN.35	0.04	0.09	0.16	0.2	0.17	0.10	0.05
LN.58	0.03	0.09	0.21	0.35	0.29	0.14	0.05

continued on next page

Table 4.13 – continued from previous page

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
MR.43	0.09	0.16	0.20	0.18	0.11	0.05	0.02
MR.44	0.07	0.11	0.15	0.15	0.12	0.07	0.04
MR.45	0.05	0.10	0.16	0.19	0.16	0.09	0.05
MR.46	0.06	0.11	0.17	0.18	0.14	0.08	0.04
MR.47	0.06	0.15	0.26	0.28	0.17	0.08	0.03
MR.48	0.05	0.08	0.10	0.12	0.10	0.07	0.05
MR.50	0.03	0.06	0.10	0.14	0.15	0.12	0.07
MR.53	0.06	0.1	0.14	0.14	0.11	0.07	0.04
MR.54	0.04	0.05	0.07	0.08	0.08	0.06	0.05
MR.55	0.04	0.07	0.12	0.16	0.15	0.11	0.06
MR.56	0.04	0.07	0.10	0.12	0.12	0.09	0.06
R3D.01	0.00	0.01	0.03	0.13	0.38	0.53	0.27
R3D.02	0.00	0.00	0.03	0.18	0.75	0.76	0.19
R3D.03	0.00	0.00	0.02	0.16	0.83	0.91	0.19
R3D.04	0.00	0.00	0.03	0.23	1.07	0.73	0.12
R3D.05	0.00	0.01	0.06	0.37	0.95	0.47	0.09
R3D.06	0.00	0.01	0.09	0.53	1.05	0.33	0.05
R3D.07	0.00	0.00	0.02	0.12	0.57	0.88	0.28
R3D.08	0.00	0.00	0.02	0.17	0.8	0.85	0.19
R3D.09	0.00	0.02	0.11	0.46	0.75	0.32	0.07
R3D.10	0.00	0.02	0.07	0.22	0.45	0.39	0.16
R3D.11	0.00	0.00	0.02	0.10	0.44	0.75	0.33
R3D.12	0.00	0.01	0.04	0.18	0.54	0.58	0.21
R3D.13	0.00	0.00	0.02	0.10	0.46	0.82	0.34
R3D.14	0.00	0.00	0.03	0.17	0.65	0.74	0.22
R3D.15	0.00	0.02	0.09	0.34	0.61	0.36	0.10
R3D.16	0.01	0.03	0.14	0.49	0.66	0.27	0.06
R3D.17	0.00	0.00	0.02	0.14	0.74	0.95	0.22
R3D.18	0.00	0.00	0.01	0.09	0.61	1.12	0.30
R3D.19	0.00	0.01	0.06	0.24	0.58	0.48	0.16
R3D.20	0.01	0.04	0.22	0.70	0.61	0.17	0.03
R3D.21	0.00	0.00	0.02	0.11	0.40	0.68	0.33
R3D.22	0.00	0.02	0.11	0.56	0.94	0.30	0.05
R3D.23	0.00	0.01	0.04	0.25	0.83	0.61	0.14
R3D.24	0.00	0.00	0.01	0.15	0.98	1.02	0.16
VR.04	0.07	0.19	0.38	0.36	0.17	0.06	0.02

continued on next page

Table 4.13 – continued from previous page

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
VR.09	0.05	0.07	0.08	0.08	0.07	0.05	0.03
VR.11	0.1	0.14	0.16	0.13	0.08	0.04	0.02
VR.13	0.03	0.05	0.08	0.11	0.12	0.10	0.07
VR.14	0.09	0.23	0.38	0.31	0.14	0.04	0.01
VR.16	0.08	0.15	0.23	0.22	0.14	0.07	0.03
VR.17	0.08	0.22	0.36	0.30	0.14	0.05	0.02
VR.18	0.06	0.06	0.06	0.05	0.04	0.03	0.02
VR.19	0.06	0.13	0.22	0.24	0.16	0.08	0.03
VR.23	0.03	0.06	0.12	0.19	0.21	0.15	0.08
VR.26	0.03	0.05	0.06	0.06	0.06	0.05	0.04
VR.31	0.11	0.15	0.16	0.12	0.07	0.04	0.02
VR.32	0.05	0.11	0.17	0.20	0.16	0.09	0.04
VR.36	0.04	0.08	0.15	0.20	0.18	0.12	0.06
VR.39	0.09	0.11	0.11	0.09	0.06	0.04	0.02
VR.42	0.05	0.08	0.09	0.09	0.08	0.06	0.04
TIF	2.16	4.38	7.78	13.01	21.55	17.76	5.60
SEM	0.68	0.48	0.36	0.28	0.22	0.24	0.42
Reliability	0.54	0.77	0.87	0.92	0.95	0.94	0.82

Table 4.14: Item and test information for the 9 Letter and Number Series items

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
LN.01	0.11	0.30	0.45	0.29	0.10	0.03	0.01
LN.03	0.04	0.15	0.36	0.43	0.23	0.08	0.02
LN.05	0.09	0.31	0.58	0.38	0.11	0.03	0.01
LN.06	0.03	0.09	0.27	0.45	0.33	0.13	0.04
LN.07	0.03	0.17	0.69	0.74	0.20	0.03	0.01
LN.33	0.04	0.15	0.40	0.49	0.24	0.07	0.02
LN.34	0.03	0.18	0.67	0.71	0.20	0.04	0.01
LN.35	0.04	0.10	0.21	0.28	0.23	0.12	0.05
LN.58	0.01	0.05	0.26	0.76	0.55	0.14	0.02
TIF	0.42	1.51	3.89	4.54	2.21	0.66	0.17
SEM	1.55	0.82	0.51	0.47	0.67	1.24	2.41
Reliability	NA	0.34	0.74	0.78	0.55	NA	NA

Table 4.15: Item and test information for the 11 Matrix Reasoning items

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
MR.43	0.10	0.22	0.32	0.25	0.12	0.05	0.02
MR.44	0.07	0.15	0.25	0.25	0.16	0.07	0.03
MR.45	0.03	0.12	0.33	0.49	0.29	0.10	0.03
MR.46	0.05	0.15	0.33	0.38	0.21	0.08	0.02
MR.47	0.06	0.16	0.34	0.37	0.20	0.07	0.02
MR.48	0.05	0.10	0.17	0.20	0.16	0.09	0.04
MR.50	0.03	0.06	0.10	0.14	0.14	0.11	0.07
MR.53	0.06	0.13	0.20	0.22	0.15	0.08	0.04
MR.54	0.04	0.08	0.14	0.19	0.18	0.12	0.06
MR.55	0.03	0.08	0.15	0.22	0.21	0.13	0.07
MR.56	0.04	0.08	0.15	0.21	0.19	0.12	0.06
TIF	0.55	1.33	2.49	2.92	2.02	1.02	0.45
SEM	1.34	0.87	0.63	0.58	0.70	0.99	1.49
Reliability	NA	0.25	0.6	0.66	0.51	0.02	NA

Table 4.16: Item and test information for the 24 Three Dimensional Rotation items

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
R3D.01	0.00	0.00	0.00	0.04	0.30	1.08	0.60
R3D.02	0.00	0.00	0.00	0.03	0.51	1.98	0.34
R3D.03	0.00	0.00	0.00	0.05	0.65	1.77	0.28
R3D.04	0.00	0.00	0.02	0.19	1.15	0.88	0.12
R3D.05	0.00	0.01	0.06	0.36	1.01	0.49	0.08
R3D.06	0.00	0.01	0.08	0.53	1.15	0.35	0.05
R3D.07	0.00	0.00	0.00	0.04	0.38	1.55	0.50
R3D.08	0.00	0.00	0.01	0.08	0.76	1.37	0.24
R3D.09	0.00	0.02	0.10	0.47	0.86	0.34	0.06
R3D.10	0.00	0.01	0.04	0.20	0.61	0.60	0.19
R3D.11	0.00	0.00	0.00	0.03	0.29	1.23	0.63
R3D.12	0.00	0.00	0.01	0.10	0.60	1.01	0.29
R3D.13	0.00	0.00	0.00	0.04	0.33	1.27	0.57
R3D.14	0.00	0.00	0.01	0.11	0.66	1.08	0.26
R3D.15	0.00	0.01	0.06	0.33	0.88	0.50	0.10
R3D.16	0.00	0.02	0.13	0.53	0.77	0.28	0.06
R3D.17	0.00	0.00	0.00	0.05	0.56	1.72	0.33
R3D.18	0.00	0.00	0.00	0.02	0.26	1.95	0.65
R3D.19	0.00	0.00	0.02	0.15	0.79	0.94	0.20
R3D.20	0.01	0.05	0.23	0.63	0.55	0.17	0.03
R3D.21	0.00	0.00	0.01	0.05	0.33	1.03	0.54
R3D.22	0.00	0.02	0.11	0.55	0.87	0.29	0.05
R3D.23	0.00	0.00	0.03	0.23	0.90	0.69	0.14
R3D.24	0.00	0.00	0.00	0.04	0.70	2.03	0.23
TIF	0.02	0.15	0.94	4.84	15.87	24.6	6.54
SEM	6.34	2.55	1.03	0.45	0.25	0.20	0.39
Reliability	NA	NA	NA	0.79	0.94	0.96	0.85

Table 4.17: Item and test information for the 21 Verbal Reasoning items

	Latent Trait Level (normal scale)						
Item	-3	-2	-1	0	1	2	3
VR.04	0.05	0.24	0.69	0.56	0.15	0.03	0.01
VR.09	0.07	0.13	0.18	0.17	0.12	0.06	0.03
VR.11	0.15	0.31	0.36	0.21	0.08	0.03	0.01
VR.13	0.03	0.05	0.09	0.12	0.13	0.10	0.07
VR.14	0.08	0.34	0.71	0.42	0.10	0.02	0.00
VR.16	0.07	0.22	0.43	0.38	0.16	0.05	0.01
VR.17	0.07	0.33	0.77	0.45	0.10	0.02	0.00
VR.18	0.14	0.16	0.14	0.09	0.05	0.02	0.01
VR.19	0.04	0.17	0.45	0.50	0.22	0.06	0.01
VR.23	0.02	0.05	0.14	0.28	0.32	0.20	0.09
VR.26	0.04	0.07	0.10	0.13	0.12	0.09	0.06
VR.31	0.20	0.39	0.36	0.17	0.06	0.02	0.00
VR.32	0.04	0.13	0.30	0.39	0.24	0.09	0.03
VR.36	0.03	0.08	0.17	0.24	0.22	0.13	0.06
VR.39	0.15	0.22	0.21	0.14	0.07	0.03	0.01
VR.42	0.06	0.10	0.14	0.14	0.11	0.07	0.04
TIF	1.25	3.00	5.25	4.39	2.26	1.03	0.44
SEM	0.90	0.58	0.44	0.48	0.67	0.99	1.50
Reliability	0.20	0.67	0.81	0.77	0.56	0.03	NA

4.3.3 Discussion

A key finding from Study 1 relates to the broad range of means and standard deviations for the ICAR items as these values demonstrated that the un-proctored and untimed administration of cognitive ability items online does not lead to uniformly high scores with insufficient variance. To the contrary, all of the Three-Dimensional Rotation items and more than half of all 60 items were answered incorrectly more often than correctly and the weighted mean for all items was only 0.53. This point was further supported by the IRT analyses in that the item information functions demonstrate a relatively wide range of item difficulties.

Internal consistency was good for the Three-Dimensional Rotation item type, adequate for the Letter and Number Series and the Verbal Reasoning item types, and marginally adequate for the Matrix Reasoning item type. This suggests that the 11 Matrix Reasoning items were not uniformly measuring a singular latent construct whereas performance on the Three-Dimensional Rotation items was highly consistent. For the composites based on both 16 and 60 items however, internal consistencies were adequate (α=0.81; ω_{total}=0.83) and good (α=0.93; ω_{total}=0.94), respectively. While higher reliabilities reflect the greater number of items in the ICAR60, it should be noted that the general factor saturation was slightly higher for the shorter 16-item measure (ICAR16 ω_h=0.66; ICAR60 ω_h=0.61). When considered as a function of test information, reliability was generally adequate across a wide range of latent trait levels, and particularly good within approximately ± 1.5 standardized units from the mean item difficulty. All of the factor analyses demonstrated evidence of both a positive manifold among items and high general factor saturation for each of the item types. In the four factor solution for the 16 item scale, the Verbal Reasoning and the Letter and Number Series factors showed particularly high *'g'* loadings (0.8).

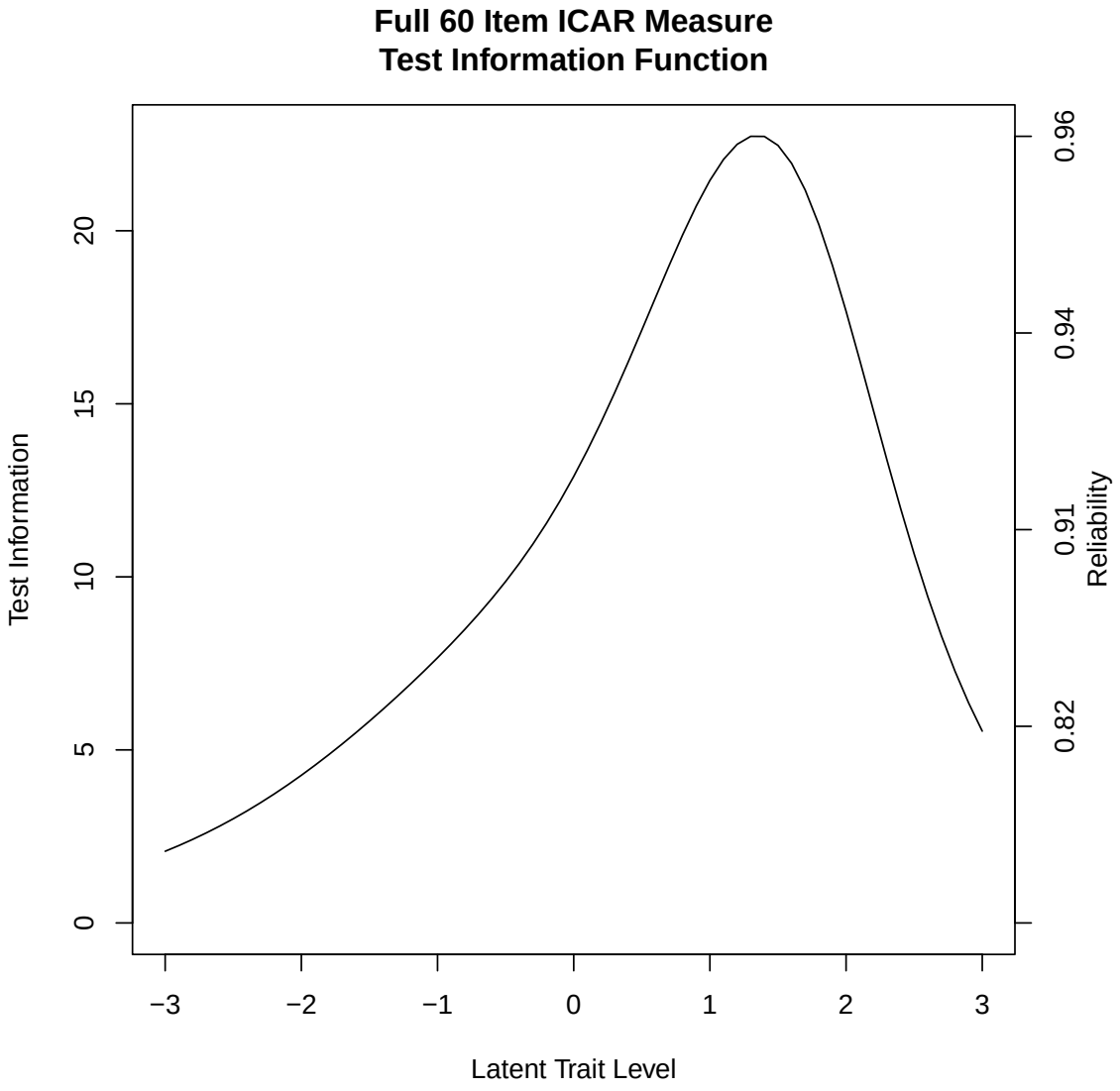


Figure 4.6: Test Information Function for the 60 ICAR items

4.4 Study 2

Following the evidence for reliable variability in ICAR scores in Study 1, it was the goal of Study 2 to evaluate the validity of these scores when using the same administration procedures. While online administration protocols precluded validation against copyrighted commercial measures, it was possible to evaluate the extent to which ICAR scores correlated with (1) self-reported achievement test scores and (2) published rank orderings of mean scores by university major. In the latter case, ICAR scores were expected to demonstrate group discriminant validity by correlating highly with the rank orderings of mean scores by university major as previously described by the Educational Testing Service (Educational Testing Service, 2010) and the College Board (College Board, 2012).

In the former case, ICAR scores were expected to reflect a similar relationship with achievement test scores as extant measures of cognitive ability. Using data from the National Longitudinal Study of Youth 1979, Frey and Detterman (2004) reported simple correlations between the SAT and the Armed Services Vocational Aptitude Battery (r = 0.82, n = 917) and several additional IQ measures (rs = 0.53 - 0.82) with smaller samples (ns = 15 - 79). In a follow-up study with a university sample, Frey and Detterman (2004) evaluated the correlation between combined SAT scores and Raven's Progressive Matrices scores, finding an uncorrected correlation of 0.48 ($p <$ .001) and a correlation after correcting for restriction of range of 0.72. Similar analyses with ACT composite scores (Koenig et al., 2008) showed a correlation of 0.77 ($p <$ .001) with the ASVAB, an uncorrected correlation with the Raven's Advanced Progressive Matrices of 0.61 ($p <$ .001), and a correlation corrected for range restriction with the Raven's APM of 0.75.

Given the breadth and duration of assessment for the ASVAB, the SAT and the ACT, positive correlations of a lesser magnitude were expected between the ICAR scores and the achievement

tests than were previously reported with the ASVAB. Correlations between the Raven's APM and the achievement test scores were expected to be more similar to the correlations between the achievement test scores and the ICAR scores, though it was not possible to estimate the extent to which the correlations would be affected by methodological differences (i.e., the un-proctored online administration of relatively few ICAR items and the use of self-reported, rather than independently verified, achievement test scores as described in the Methods section below).

4.4.1 Method

4.4.1.1. Participants

The 34,229 participants in Study 2 were a subset of those used for Study 1, chosen on the basis of age and level of educational attainment. Participants were 18 to 22 years old (m = 19.9, *s.d.* = 1.3, median = 20). Approximately 91% of participants had begun but not yet attained an undergraduate degree; the remaining 9% had attained an undergraduate degree. Among the 26,911 participants from the United States, 67.1% identified themselves as White/Caucasian, 9.8% as Hispanic-American, 8.4% as African-American, 6.0% as Asian-American, 1.0% as Native-American, and 6.3% as multi-ethnic (the remaining 1.5% did not specify).

4.4.1.2. Measures

Both the sampling method and the ICAR items used in Study 2 were identical to the procedures described in Study 1, though the total item administrations (median = 7,659) and pairwise administrations (median = 906) were notably fewer given that the participants in Study 2 were a sub-sample of those in Study 1. Study 2 also used self-report data for three additional variables collected through SAPA-project.org: (1) participants' academic major on the university level, (2)

their achievement test scores, and (3) participants' scale scores based on randomly administered items from the Intellect scale of the "100-Item Set of IPIP Big-Five Factor Markers" (Goldberg, 2014). For university major, participants were allowed to select only one option from 147 choices, including "undecided" (n = 3,460) and several categories of "other" based on academic disciplines. For the achievement test scores, participants were given the option of reporting 0, 1, or multiple types of scores, including: SAT Critical Reading (n = 7,404); SAT Mathematics (n = 7,453); and the ACT (n = 12,254). Intellect scale scores were calculated using IRT procedures, assuming unidimensionality for the Intellect items only (items assessing Openness were omitted). Based on composites of the Pearson correlations between items without imputation of missing values, the Intellect scale had an α of 0.74, an ω_h of 0.60, and an ω_{total} of 0.80. The median number of pairwise administrations for these items was 4,475.

4.4.1.3. Analyses

Two distinct methods were used to calculate the correlations between the achievement test scores and the ICAR scores in order to evaluate the effects of two different corrections. The first method used ICAR scale scores based on composites of the tetrachoric correlations between ICAR items (composites are used because each participant was administered 16 or fewer items). The correlations between these scale scores and the achievement test scores were then corrected for reliability. The α reliability coefficients reported in Study 1 were used for the ICAR scores. For the achievement test scores, the need to correct for reliability was necessitated by the use of self-reported scores. Several researchers have demonstrated the reduced reliability of self-reported scores in relation to official test records (Cassady, 2001; Cole and Gonyea, 2009; Kuncel et al., 2005; Mayer et al., 2006), citing participants' desire to misrepresent their performance and/or memory errors as the most likely causes. Despite these concerns, the reported correlations between self-reported and actual scores suggest that the rank-ordering of

scores is maintained, regardless of the magnitude of differences (Cole and Gonyea, 2009; Kuncel et al., 2005; Mayer et al., 2006). Reported correlations between self-reported and actual scores have ranged from 0.74 to 0.86 for the SAT - Critical Reading section, 0.82 to 0.88 for the SAT - Mathematics, and 0.82 to 0.89 for the SAT - Combined (Cole and Gonyea, 2009; Kuncel et al., 2005; Mayer et al., 2006). Higher correlations were found by Cole and Gonyea (2009) for the ACT Composite (0.95). The Study 2 sample approximated the samples on which these reported correlations were based in that (1) participants were reminded about the anonymity of their responses and (2) the age range of participants was limited to 18 to 22 years. The weighted mean values from these findings (SAT-CR = 0.86; SAT-M = 0.88; SAT-Combined = 0.88; ACT = 0.95) were used as reliability coefficients for the achievement test scores when correcting correlations between the achievement tests and other measures (ICAR scores and the IPIP-100 Intellect scores).

The second method for calculating correlations between ICAR scores and achievement test scores used IRT-based (2PL) scoring (Revelle, 2014). Scale scores for each item type and the full test were calculated for each participant, and these scale scores were then correlated with the achievement test scores. In this case, corrections were made to address the potential for an incidental selection effect due to optional reporting of achievement test scores (Cassady, 2001; Frucot and Cook, 1994). 52.5% of participants in Study 2 did not report any achievement test scores; 10.1% reported scores for all three (SAT - CR, SAT - M, and ACT). These circumstances would result in an incidental selection effect if the correlations between self-reported achievement test scores and the ICAR measures were affected by the influence of a third variable on one or both measures (Sackett and Yang, 2000). The so-called "third" variable in this study likely represented a composite of latent factors which are neither ergodic nor quantifiable but which resulted in group differences between those who reported their scores and those who did not. If the magnitude of differences in achievement test scores between groups were

non-trivial, the effect on the overall correlations would also be non-trivial given the proportion of participants not reporting. The need for correction procedures in this circumstance was elaborated by both Pearson (1903) and Thorndike (1949), though the methods employed here were developed in the econometrics literature and are infrequently used by psychologists (Sackett and Yang, 2000). Clark and Houle (2012) and Cuddeback et al. (2004) provide useful illustrations of these procedures. The two-step method of the "Heckman correction" (Greene, 2008; Heckman, 1976, 1979; Toomet and Henningsen, 2008) was used to evaluate and correct for selection effects where warranted using IPIP-100 Intellect scores.

In addition to these analyses of the relationship between ICAR scores and achievement test scores, the Study 2 sample was used to evaluate the correlations between the ICAR items and the published rank orderings of mean scores by university major. This was done using IRT-based ICAR scores when grouped by academic major on the university level. These were evaluated relative to similar data sets published by the Educational Testing Service (Educational Testing Service, 2010) and the College Board (College Board, 2012) for the GRE and SAT, respectively. GRE scores were based on group means for 287 "intended graduate major" choices offered to fourth-year university students and non-enrolled graduates who took the GRE between July 1, 2005 and June 30, 2008 (N = 569,000). These 287 groups were consolidated with weighting for sample size in order to match the 147 university major choices offered with the ICAR. Of these 147 majors, only the 91 with $n > 20$ were used. SAT scores were based on group means for 38 "intended college major" choices offered to college-bound seniors in the high school graduating class of 2012 (N = 1,411,595). In this case, the 147 university major choices offered with the ICAR were consolidated to match 29 of the choices offered with the SAT. The 9 incompatible major choices collectively represented only 1.3% of the SAT test-takers. The omitted majors were: Construction Trades; Mechanic and Repair Technologies/Technician; Military Technologies and Applied Sciences; Multi/Interdisciplinary Studies; Precision Production;

Security and Protective Services; Theology and Religious Vocations; Other; and Undecided.

4.4.2 Results

Descriptive statistics for the self-reported achievement test scores are shown in Table 4.18. Correlations between self-reported achievement test scores and ICAR scale scores calculated using composites of the tetrachoric correlations are shown in Table 4.19, with uncorrected correlations shown below the diagonal and the correlations corrected for reliability shown above the diagonal. Reliabilities for each measure are given on the diagonal. Correlations between composites which were not independent have been omitted. Corrected correlations between the achievement test scores and both the 16 and 60 item ICAR composites ranged from 0.52 - 0.59 ($ses \leq 0.016$).[2]

Table 4.18: Self-reported achievement test scores and national norms

	Study 2 self-reported			published norms	
	n	mean	s.d.	mean	s.d.
SAT - Critical Reading	7,404	609	120	496	114
SAT - Math	7,453	611	121	514	117
ACT	12,254	25.4	5.0	21.1	5.2

Note: SAT norms are from the 2012 *Total Group Profile Report.* ACT norms are from the 2011 *ACT Profile Report.*

Table 4.20 presents the correlations between the self-reported achievement test scores and the IRT-based ICAR scores, with the uncorrected correlations below the diagonal and the correlations corrected for incidental selection effects above the diagonal. Correlations between non-independent scores were omitted. Scores for the ICAR measures were based on a mean of 2

[2]The standard error of the composite scores are a function of both the number of items and the number of participants who took each pair of items (Revelle and Brown, 2013). Estimates of the standard errors can be identified through the use of bootstrapping procedures to derive estimates of the confidence intervals of the correlations (Revelle, 2014). In this case, the confidence intervals were estimated based on 100 sampling iterations.

Table 4.19: Correlations between self-reported achievement test scores and ICAR composite scales

					ICAR composite scale scores					
	SAT-CR	SAT-M	SAT-CR+M	ACT	ICAR60	LN	MR	R3D	VR	ICAR16
SAT-CR[1]	*0.86*	0.83		0.69	0.52	0.41	0.37	0.39	0.68	0.52
SAT-M[2]	0.72	*0.88*		0.66	0.60	0.50	0.47	0.49	0.67	0.59
SAT-CR+M[3]			*0.89*	0.71	0.59	0.48	0.44	0.47	0.72	0.59
ACT[4]	0.62	0.60	0.65	*0.95*	0.52	0.39	0.35	0.44	0.61	0.52
ICAR60[5]	0.46	0.54	0.54	0.49	*0.93*					
LN[5]	0.33	0.41	0.40	0.33		*0.77*	0.84	0.59	0.90	
MR[5]	0.28	0.36	0.34	0.28		0.61	*0.68*	0.67	0.81	
R3D[5]	0.35	0.44	0.43	0.41		0.50	0.53	*0.93*	0.58	
VR[5]	0.55	0.55	0.59	0.52		0.69	0.58	0.49	*0.76*	
ICAR16[5]	0.43	0.50	0.50	0.46						*0.81*

Note: Uncorrected correlations below the diagonal, correlations corrected for reliability above the diagonal. Reliability values shown on the diagonal.
[1] n = 7,404
[2] n = 7,453
[3] n = 7,348
[4] n = 12,254
[5] Composite scales formed based on item correlations across the full sample (n = 34,229).

to 4 responses for each of the item types (mean number of LN items administered = 3.2, sd = 1.3; MR items m = 2.8, sd = 1.1; R3D items m = 2.0, sd = 1.5; VR items m = 4.3, sd = 2.2) and 12 to 16 items for the ICAR60 scores (m = 12.4, sd = 3.8). Corrected correlations between the achievement test scores and ICAR60 ranged from 0.44 to 0.47 ($ses \leq 0.016$).

Tables 4.21 and 4.22 contain group-level correlations using mean scores for university major. Table 4.21 shows the correlations between the published norms for the SAT, the mean self-reported SAT scores for each major in the Study 2 sample, and the mean IRT-based ICAR scores for each major in the Study 2 sample. The correlation between mean ICAR scores by major and mean combined SAT scores by major in the published norms was 0.75 (se = 0.147). Table 4.22 shows the correlations between the published norms for the GRE by major and the IRT-based ICAR scores for the corresponding majors in the Study 2 sample (self-reported GRE scores were not collected). The correlation between mean ICAR scores by major and mean

Table 4.20: Correlations between self-reported achievement test scores and IRT-based ICAR scores

					ICAR IRT-based scores				
	SAT-CR	SAT-M	SAT-CR+M	ACT	ICAR60	LN	MR	R3D	VR
SAT-CR[1]					0.44	0.37	0.35	0.37	0.44
SAT-M[2]	0.72				0.44	0.33	0.29	0.35	0.39
SAT-CR+M[3]	0.93	0.93			0.47	0.37	0.33	0.38	0.45
ACT[4]	0.62	0.60	0.65		0.44	0.35	0.32	0.38	0.43
ICAR60[5]	0.36	0.42	0.42	0.39					
LN[5]	0.24	0.28	0.28	0.24					
MR[5]	0.18	0.22	0.21	0.18		0.30			
R3D[5]	0.25	0.32	0.30	0.28		0.26	0.23		
VR[5]	0.35	0.36	0.38	0.36		0.36	0.26	0.22	

Note: IRT scores for ICAR measures based on 2 to 4 responses per participant for each item type (LN, MR, R3D, VR) and 12 to 16 responses for ICAR60. Uncorrected correlations are below the diagonal, correlations corrected for incidental selection are above the diagonal.
[1] n = 7,404
[2] n = 7,453
[3] n = 7,348
[4] n = 12,254
[5] n = 34,229

combined GRE scores by major in the published norms was 0.86 (*se* = 0.092).

4.4.3 Discussion

After correcting for the "reliability" of self-reported scores, the 16 item *ICAR Sample Test* correlated 0.59 with combined SAT scores and 0.52 with the ACT composite. Correlations based on the IRT-based ICAR scores were lower though these scores were calculated using even fewer items; correlations were 0.47 and 0.44 with combined SAT scores and ACT composite scores respectively based on an average of 12.4 ICAR60 items answered per participant. As expected, these correlations were smaller than those reported for longer cognitive ability measures such as the ASVAB and the Raven's APM (Frey and Detterman, 2004; Koenig et al., 2008).

Table 4.21: Correlations between mean SAT norms, mean SAT scores in Study 2 and mean IRT-based ICAR scores when ranked by university major

	College Board Norms			Study 2 Self-Reported			Study 2 IRT-based			
	SAT-CR	SAT-M	SAT-CR+M	SAT-CR	SAT-M	SAT-CR+M	ICAR60	LN	MR	R3D
SAT-M norms	0.66									
SAT-CR+M norms	0.91	0.91								
SAT-CR study 2	0.79	0.61	0.77							
SAT-M study 2	0.56	0.80	0.74	0.81						
SAT-CR+M study 2	0.71	0.74	0.80	0.95	0.95					
ICAR60 study 2	0.53	0.84	0.75	0.60	0.77	0.72				
LN study 2	0.41	0.80	0.66	0.49	0.76	0.66	0.96			
MR study 2	0.22	0.66	0.48	0.23	0.52	0.39	0.83	0.78		
R3D study 2	0.42	0.80	0.67	0.50	0.71	0.64	0.94	0.92	0.82	
VR study 2	0.69	0.79	0.81	0.76	0.80	0.82	0.91	0.82	0.64	0.76

Note: $n = 29$.

Table 4.22: Correlations between mean GRE norms and mean IRT-based ICAR scores when ranked by university major

	ETS Norms			Study 2 IRT-based			
	GREV	GREQ	GREVQ	ICAR60	LN	MR	R3D
GREQ norms	0.23						
GREVQ norms	0.63	0.90					
ICAR60 study 2	0.54	0.78	0.86				
LN study 2	0.41	0.72	0.76	0.93			
MR study 2	0.42	0.71	0.75	0.86	0.81		
R3D study 2	0.44	0.80	0.83	0.92	0.86	0.75	
VR study 2	0.67	0.63	0.80	0.92	0.80	0.79	0.77

Note: $n = 91$.

The ICAR items demonstrated strong group discriminant validity on the basis of university majors. This indicates that the rank ordering of mean ICAR scores by major is strongly correlated with the rank ordering of mean SAT scores and mean GRE scores. Consistent with the individual-level correlations, the group-level correlations were higher between the ICAR subtests and the mathematics subtests of the SAT and the GRE relative to the verbal subtests.

4.5 Study 3

The goal of the third study was to evaluate the construct validity of the ICAR items against a commercial measure of cognitive ability. Due to the copyrights associated with commercial measures, these analyses were based on administration to an offline sample of university students rather than an online administration.

4.5.1 Method

4.5.1.1. Participants

Participants in Study 3 were 137 college students (76 female) enrolled at a selective private university in the midwestern United States. Students participated in exchange for credit in an introductory psychology course. The mean age of participants in this sample was 19.7 years (*sd* = 1.2, median = 20) with a range from 17 to 25 years. Within the sample, 67.2% reported being first-year students, 14.6% second-year students, 8.0% third-year students and the remaining 10.2% were in their fourth year or beyond. With regards to ethnicity, 56.2% identified themselves as White/Caucasian, 26.3% as Asian-American, 4.4% as African-American, 4.4% as Hispanic-American, and 7.3% as multi-ethnic (the remaining 1.5% did not specify).

4.5.1.2. Measures

Participants in the university sample were administered the 16 item *ICAR Sample Test*. The presentation order of these 16 items was randomized across participants. Participants were also administered the *Shipley-2*, which is a 2009 revision and restandardization of the *Shipley Institute of Living Scale* (Shipley et al., 2009, 2010). The *Shipley-2* is a brief measure of cognitive functioning and impairment that most participants completed in 15 to 25 minutes. While the *Shipley-2* is a timed test, the majority of participants stopped working before using all of the allotted time. The *Shipley-2* has two administration options. Composite A (n = 69) includes a vocabulary scale designed to assess crystallized skills and an abstraction scale designed to assess fluid reasoning skills (Shipley et al., 2009). Composite B (n = 68) includes the same vocabulary scale and a spatial measure of fluid reasoning called the "Block Patterns" scale (Shipley et al., 2009). All three scales included several items of low difficulty with little or no variance in this sample. After removal of items without variance, internal consistencies were low for the Abstraction scale (10 of 25 items removed, α = 0.37; ω_{total} = 0.51) and the Vocabulary scale (7 of 40 items removed, α = 0.61; ω_{total} = 0.66). The Block Patterns scale had fewer items without variance (3 of 26) and adequate consistency (α = 0.83, ω_{total} = 0.88). Internal consistencies were calculated using Pearson correlations between items.

4.5.1.3. Analyses

Correlations were evaluated between scores on the *ICAR Sample Test* and a brief commercial measure of cognitive ability, the *Shipley-2*. Two types of corrections were relevant to these correlations; one for the restriction of range among scores and a second for reliability. The prospect of range restriction was expected on the grounds that participants in the sample were students at a highly selective university. The presence of restricted range was evaluated by

looking for reduced variance in the sample relative to populations with similar characteristics. In this case, the university sample was evaluated relative to the online sample. Where present, the appropriate method for correcting this type of range restriction uses the following equation (case 2c from Sackett and Yang, 2000) (Bryant and Gokhale, 1972; Alexander, 1990):

$$\hat{\rho}_{xy} = r_{xy}(s_x/S_x)(s_y/S_y) \pm \sqrt{[1-(s_x/S_x)^2][1-(s_y/S_y)^2]} \tag{4.1}$$

where s_x and s_y are the standard deviations in the restricted sample, S_x and S_y are the standard deviations in the unrestricted sample and the $\pm$ sign is conditional on the direction of the relationship between the selection effect and each of the variables, x and y. When correcting for reliability, the published reliabilities (Shipley et al., 2010) were used for each of the *Shipley-2* composites (0.925 for Composite A and 0.93 for Composite B) instead of the reliabilities within the sample due to the large number of items with little or no variance.

4.5.2 Results

The need to correct for restriction of range was indicated by lower standard deviations of scores on all of the subtests and composites for the *Shipley-2* and the *ICAR Sample Test*. Table 4.23 shows the standard deviation of scores for the participants in Study 3 (the "restricted" sample) and the reference scores (the "unrestricted" samples).

Table 4.23: Standard deviations of scores for the unrestricted samples and Study 3

		Shipley-2				*ICAR*
Sample	Block Patterns	Abstraction	Vocab	Comp A	Comp B	*Sample Test*
Unrestricted	15.0	15.0	15.0	15.0	15.0	1.86
Study 3	11.1	9.8	6.8	6.8	8.9	1.48

Note: Unrestricted standard deviations based on the published norms for the *Shipley-2* and the Study 1 sample for the *ICAR Sample Test*.

Correlations between the ICAR scores and *Shipley-2* scores are given in Table 4.24, including the uncorrected correlations, the correlations corrected for range restriction and the correlations corrected for reliability and range restriction. The range and reliability corrected correlations between the *ICAR Sample Test* and the *Shipley-2* composites were nearly identical at 0.81 and 0.82 (*se* = 0.10).

Table 4.24: Correlations between the *ICAR Sample Test* and the *Shipley-2*

ICAR16	Block Patterns[1]	Abstraction[2]	Vocab[3]	Comp A[2]	Comp B[1]
Uncorrected	0.40	0.44	0.15	0.41	0.41
Range corrected	0.64	0.69	0.59	0.68	0.68
Range & reliability corrected				0.82	0.81

[1] n = 68
[2] n = 69
[3] n = 137

4.5.3 Discussion

Correlations between the ICAR scores and the *Shipley-2* were comparable to those between the *Shipley-2* and other measures of cognitive ability. The correlations after correcting for reliability and restricted range between the 16 item *ICAR Sample Test* and *Shipley-2* composite A and B were 0.82 and 0.81, respectively. Correlations between *Shipley-2* composite A and B were 0.64 and 0.60 with the *Wonderlic Personnel Test*, 0.77 and 0.72 with the Full-Scale IQ scores for the *Wechsler Abbreviated Scale of Intelligence* in an adult sample, and 0.86 and 0.85 with the Full-Scale IQ scores for the *Wechsler Adult Intelligence Scale* (Shipley et al., 2010).

4.6 General Discussion

Reliability and validity data from these studies suggest that a public-domain measure of cognitive ability is a viable option. More specifically, they demonstrate that brief, un-proctored, and untimed administrations of items from the International Cognitive Ability Resource are moderately-to-strongly correlated with measures of cognitive ability and achievement. While this method of administration is inherently less precise and exhaustive than many traditional assessment methods, it offers many benefits. Online assessment allows for test administration at any time of day, in any geographic location, and over any type of internet-enabled electronic device. These administrations can be conducted either with or without direct interaction with the research team. Measures constructed with public-domain item types like those described here can be easily customized for test length and content as needed to match the research topic under evaluation. All of this can be accomplished without the cost, licensing, training, and software needed to administer the various types of copyright-protected commercial measures.

These data also suggest that there are many ways in which the ICAR can be improved. With regard to the existing item types, more - and more difficult - items are needed for all of the item types except perhaps the Three-Dimensional Rotation items. While the development of additional Letter and Number Series items can be accomplished formulaically, item development procedures for the Verbal Reasoning items is complicated by the need for items to be resistant to basic internet word searches. The Matrix Reasoning items require further structural analyses before further item development as these items demonstrated less unidimensionality than the other three item types. This may be appropriate if they are to be used as a measure of general cognitive ability, but it remains important to identify the ways in which these items assess subtly different constructs. This last point relates to the additional need for analyses of differential item functioning for all of the item types and the test as a whole.

The inclusion of many more item types in the ICAR is also needed as is more extensive validation of new and existing item types. The most useful additions in the near term would include item types which assess constructs distinct from the four item types described here. Several such item types are in various stages of development and piloting by the authors and their collaborators. These item types should be augmented with extant, public-domain item types when feasible.

4.7 Conclusion

Public-domain measures of cognitive ability have considerable potential. We propose that the International Cognitive Ability Resource provides a viable foundation for collaborators who are interested in contributing extant or newly-developed public-domain tools. To the extent that these tools are well-suited for online administration, they will be particularly useful for large-scale cognitive ability assessment and/or use in research contexts beyond the confines of traditional testing environments. As more item types become available, the concurrent administration of ICAR item types will become increasingly valuable for researchers studying the structure of cognitive abilities on both the broad, higher-order levels (e.g., spatial and verbal abilities) as well as the relatively narrow (e.g., more closely related abilities such as two- and three-dimensional rotation). The extent to which a public-domain resource like the ICAR fulfills this potential ultimately depends on the researchers for whom it offers the highest utility. We entreat these potential users to consider contributing to its on-going development, improvement, validation and maintenance.

4.8 Appendix B – *ICAR Sample Test*

The following items represent the 16 item ICAR Sample Test that is referenced in the submitted manuscript and several other locations. These items represent a subset of the four item types described in the main text. The Verbal Reasoning items are denoted as VR, Letter and Number Series as LN, Matrix Reasoning as MR, and Three-Dimensional Rotation as R3D.

VR.4
What number is one fifth of one fourth of one ninth of 900?
(1) 2 (2) 3 (3) 4 (4) 5 (5) 6 (6) 7

VR.16
Zach is taller than Matt and Richard is shorter than Zach. Which of the following statements would be most accurate?
(1) Richard is taller than Matt (2) Richard is shorter than Matt (3) Richard is as tall as Matt (4) It's impossible to tell

---

VR.17
Joshua is 12 years old and his sister is three times as old as he. When Joshua is 23 years old, how old will his sister be?
(1) 35 (2) 39 (3) 44 (4) 47 (5) 53 (6) 57

VR.19
If the day after tomorrow is two days before Thursday then what day is it today?
(1) Friday (2) Monday (3) Wednesday (4) Saturday (5) Tuesday (6) Sunday

---

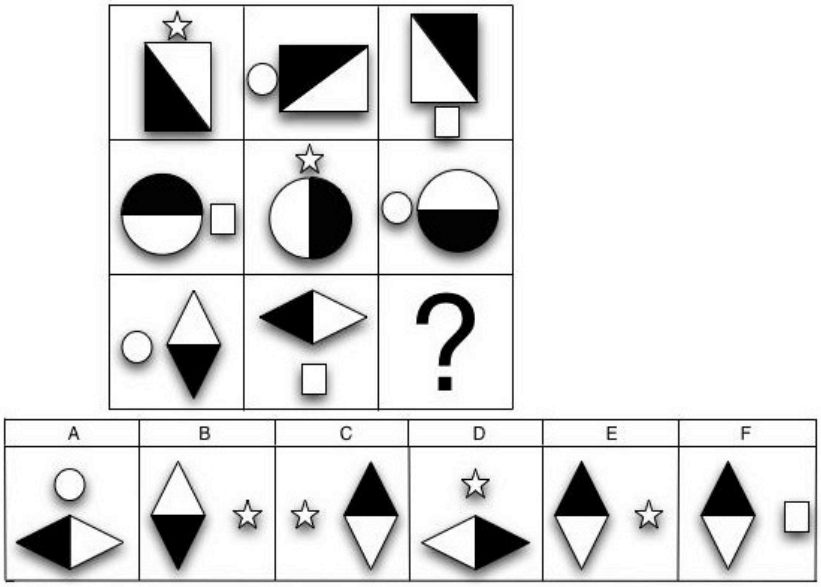
LN.7
In the following alphanumeric series, what letter comes next? K N P S U
(1) S (2) T (3) U (4) V (5) W (6) X

LN.33
In the following alphanumeric series, what letter comes next? V Q M J H
(1) E (2) F (3) G (4) H (5) I (6) J

LN.34
In the following alphanumeric series, what letter comes next? I J L O S
(1) T (2) U (3) V (4) X (5) Y (6) Z

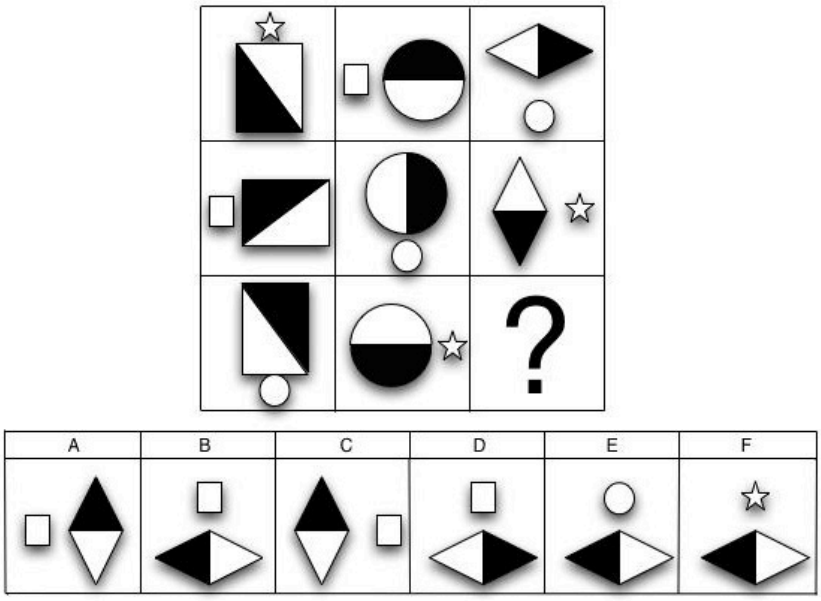
LN.58
In the following alphanumeric series, what letter comes next? Q S N P L
(1) J (2) H (3) I (4) N (5) M (6) L

MX.45 Please indicate which is the best answer to complete the figure below.



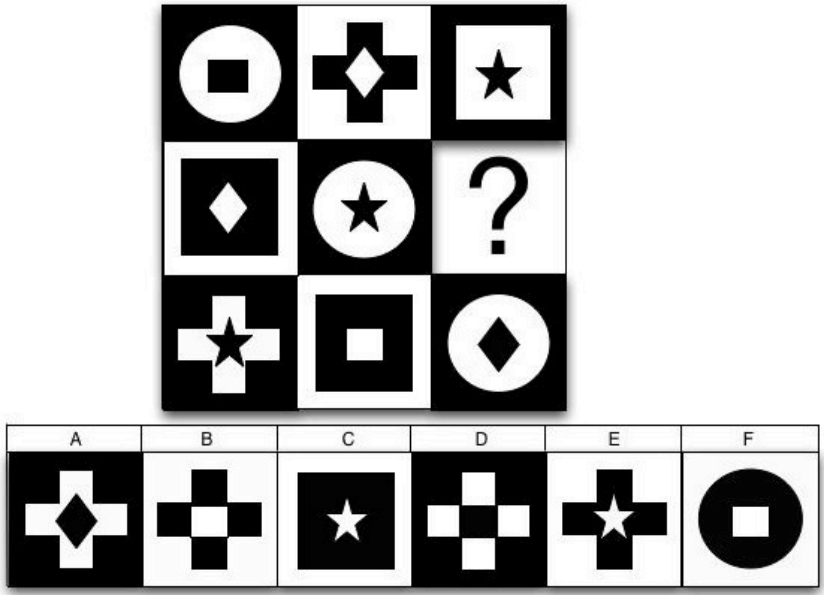
(1) A (2) B (3) C (4) D (5) E (6) F

MX.46 Please indicate which is the best answer to complete the figure below.



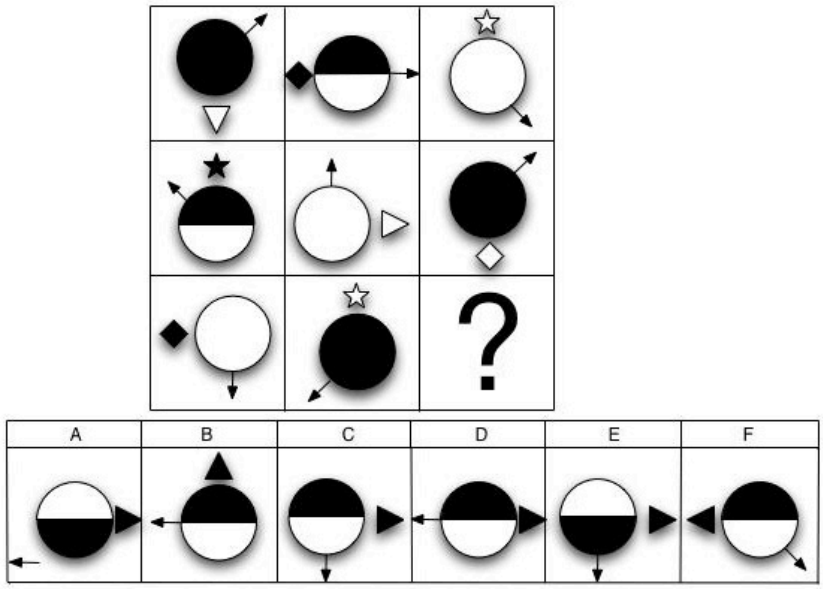
(1) A (2) B (3) C (4) D (5) E (6) F

MX.47 Please indicate which is the best answer to complete the figure below.



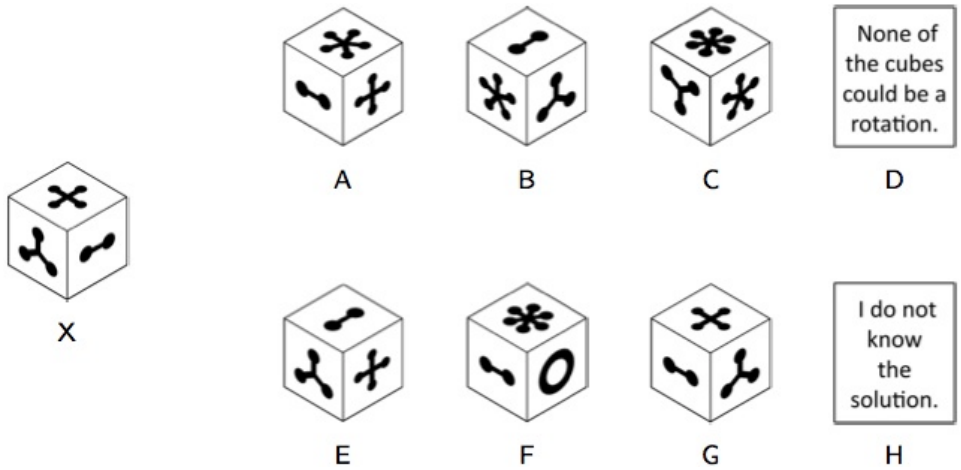
(1) A (2) B (3) C (4) D (5) E (6) F

MX.55 Please indicate which is the best answer to complete the figure below.



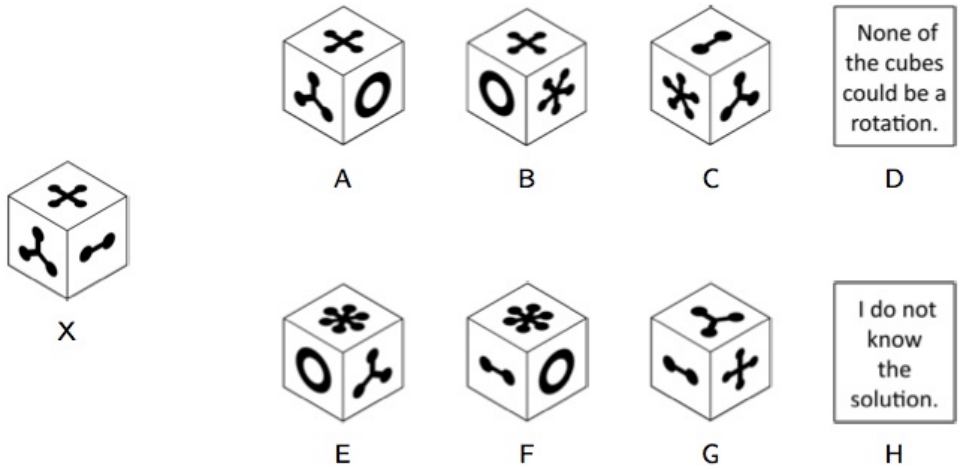
(1) A (2) B (3) C (4) D (5) E (6) F

R3D.3 All the cubes below have a different image on each side. Select the choice that could represent a rotation of the cube labeled X.



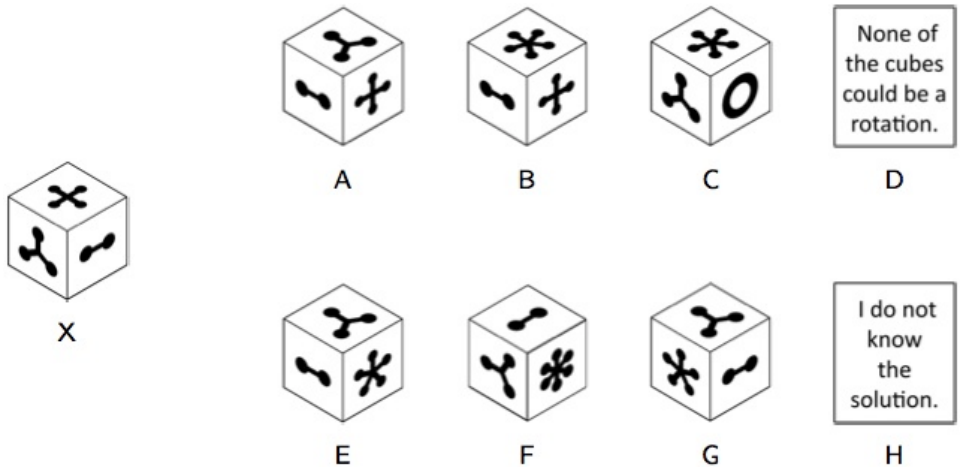
(1) A (2) B (3) C (4) D (5) E (6) F (7) G (8) H

R3D.4 All the cubes below have a different image on each side. Select the choice that could represent a rotation of the cube labeled X.



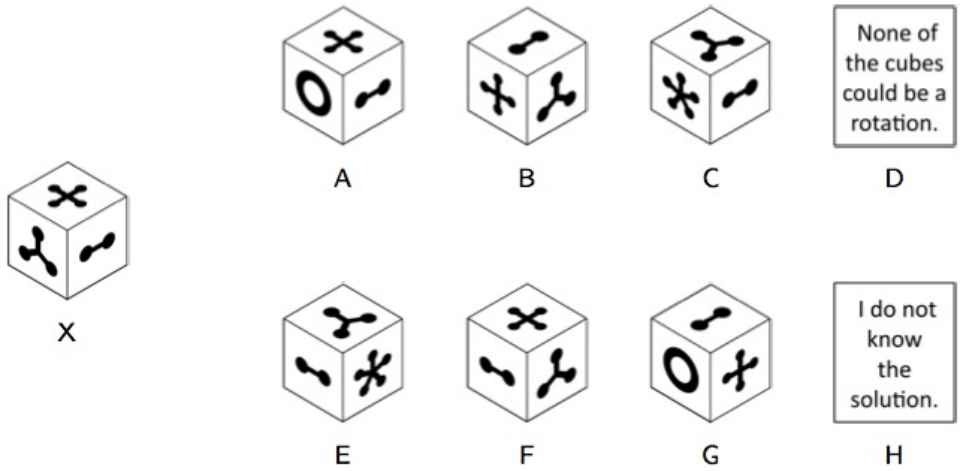
(1) A (2) B (3) C (4) D (5) E (6) F (7) G (8) H

R3D.6 All the cubes below have a different image on each side. Select the choice that could represent a rotation of the cube labeled X.



(1) A (2) B (3) C (4) D (5) E (6) F (7) G (8) H

R3D.8 All the cubes below have a different image on each side. Select the choice that could represent a rotation of the cube labeled X.



(1) A (2) B (3) C (4) D (5) E (6) F (7) G (8) H

Chapter 5

The Conative Domain: Structural evaluation and development of the SAPA Personality Inventory - Vocational Interest scales

5.1 Introduction

While the universe of conative individual differences includes a broad scope of constructs such as desires, motivations, volition and striving, research on the conative differences is most frequently conducted through the assessment of interests, especially vocational interests. The dominant interests framework, the Holland typology, has more recently come to be known as the RIASEC model of vocational interests (Holland, 1959, 1997). The RIASEC model organizes both interests and jobs according to six categories (and related scales) – Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. The framework itself allows for hierarchical organization of specific occupations which can be grouped according to shared "basic interest" categories and these in turn can be grouped at a higher level of six general interest factors (Armstrong et al., 2004). In other words, the basic interests may be seen as equivalent to the

facet level of the Big Five in the affective domain. It has also been suggested that the six factor structure can be further simplified to two dimensions which are known as "data/ideas" and "people/things" (Armstrong et al., 2008b; Prediger, 1982).

In recent years, two distinct sets of public-domain scales have been introduced for the assessment of individual differences. These include the O*NET Interest Profiler (Rounds et al., 2010) and the Oregon Vocational Interest Scales ("ORVIS", Pozzebon et al., 2010). While these measures have a considerable amount in common, they are importantly distinguished by the inclusion of two additional factors in the 8 scales of the ORVIS. It is the goal of this chapter to explore the structure of the vocational interests domain following administration of both of these sets of public-domain scales to a large international sample. Evaluations of structure shall be conducted independently for each set of scales in the first two studies before consideration is given to the full set of items together in Study 3. Following these analyses, recommendations are given for the development or improvement of the extant scales for the sake of future research based on public-domain vocational interest assessment.

5.2 Study 1

Study 1 evaluated the structure of the 60-item O*NET® Interest Profiler Short Form (Rounds et al., 2010). It was hypothesized that the extraction of six factors would demonstrate a superior fit relative to other alternatives and that the orientation of item content on these six factors would correspond to the six hypothesized Holland occupational interests types: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional ("RIASEC").

5.2.1 Methods

5.2.1.1 Participants

The data were collected from participants who voluntarily visited the SAPA-Project.org website between September 10, 2013 and March 13, 2014 in exchange for customized feedback about their personalities. This included 14,882 participants (52% female) from 138 countries. All data were self-reported. The mean age was 25.2 years (*sd* = 10.0, median = 22) with a range from 14 to 89 years.

20% of participants reported educational attainment of a high school degree or less, 56% some college, currently in college, or a college degree, 12% some graduate school or a graduate degree, and 12% chose not to report their education level. Of the 9,227 participants who reported being from the United States, 62.1% identified themselves as White/Caucasian, 5.7% as African American, 7.6% as Hispanic-American, 4.8% as Asian-American, 0.9% as Native Alaskan/Hawaiian-American, 5.5% as multi-ethnic, 1% as "Other," and 12.5% did not specify their racial/ethnic background. Participants from outside the United States were not prompted for information about their racial/ethnic background.

5.2.1.2 Measures

Participants were administered twelve item subsets of the 60-item O*NET®Interest Profiler Short Form. The Synthetic Aperture Personality Assessment ("SAPA") technique (Revelle et al., 2010b), a matrix sampling procedure, was used. The number of items to which participants responded varied by participant willingness to take more items. Of the sample, 13,215 (89%) participants responded to all twelve items. There was variability in the number of administrations for each item (median = 2,839), as well as the pairwise administrations (median

= 517, mean = 518, min = 441).

5.2.1.3 Analyses

All analyses were conducted using the *psych* package (Revelle, 2014) in R (R Core Team, 2014). Latent variable exploratory factor analysis ("EFA") was used to evaluate the structure of the 60 items in the O*NET® Interest Profiler Short Form. Factor analyses extracting from 1 to 20 factors were based on Pearson correlations between scored responses using Ordinary Least Squares ("OLS") regression models with varimax rotation (Revelle, 2014).

Goodness-of-fit was evaluated using the Minimum Average Partial criterion ("MAP", Velicer, 1976), the Standardized Root Mean Square of the Residual ("SRMR" Hu and Bentler, 1999), an empirically-derived measure of the Bayesian Information Criterion ("eBIC" Schwarz, 1978; Revelle, 2014), and an index of complexity (Hofmann, 1977, 1978). For all of these fit statistics, lower values indicate a superior fit, though the MAP and BIC will often indicate a localized minimum while the SRMR values will decrease as more factors are extracted. Good fits are typically indicated by RMSEA values of 0.05 and SRMR values of 0.08 (Kenny, 2012). The complexity reported for a given factor solution reflects the mean of the item-level complexities. It should be noted that the complexity for any single item is not sensitive to the magnitude of factor loadings but rather the degree of similarity in loading magnitudes. Given that communality is defined as the sum of the squared loadings on all factors, item complexity should be considered in conjunction with its communality. Best-fitting factor solutions were then rationally evaluated for item consistency within a given factor. The goal of these procedures was to find a factor solution that had strong empirical support as well as face validity.

5.2.2 Results

The item-level correlation matrix was not positive semi-definite, so matrix smoothing (Revelle, 2014) was performed in order to arrive at factor solutions and to calculate the empirical BIC and SRMR. Table 5.1 shows the fit statistics based on the extraction of 1 to 20 factors. SRMR suggested that more than 5 factors provided good fit. The MAP criterion and eBIC both supported a 7 factor solution. Complexity values (see also Figure 5.1) indicated similar levels of complexity for solutions extracting between 4 and 7 factors and a sharp increase in complexity at 8 factors.

Table 5.2 describes the content of each factor, where identifiable, for the factor solutions with 1 to 10 factors of extraction, with the factors in each case sorted by the magnitude of the eigenvalues. In general, the content of the factors was consistent from one level of extraction to the next though there were several notable exceptions due to occasional "re-orientation" of the factors.

The three-factor solution put items into a pattern resembling combined Holland types; each factor lumped two Holland types into one factor. The first factor was made up of Realistic and Investigative items, the second Artistic and Social items, and the third Enterprising and Conventional items. Seven of the total sixty items (12%) did not fit into the Holland types structure; for example, the item "Would like to test the quality of parts before shipment," which is a Realitistic item, was grouped into the Enterprising-Conventional factor.

Table 5.1: Fit statistics for extraction of 1 to 20 factors from the 60 O*NET items

Factors	MAP	SRMR	eBIC	complexity
1	0.0217	0.12	15,471	1.00
2	0.0181	0.10	7,095	1.24
3	0.0151	0.08	1,424	1.61
4	0.0133	0.07	-1,258	1.71
5	0.0120	0.06	-2,788	1.73
6	0.0112	0.05	-3,590	1.80
7	**0.0104**	0.05	**-4,213**	1.81
8	0.0107	0.05	-4,167	2.02
9	0.0109	0.05	-4,144	2.09
10	0.0112	0.04	-4,071	2.13
11	0.0117	0.04	-3,908	2.26
12	0.0122	0.04	-3,781	2.18
13	0.0127	0.04	-3,612	2.26
14	0.0134	0.04	-3,416	2.40
15	0.0142	0.04	-3,225	2.46
16	0.0149	0.04	-3,079	2.51
17	0.0157	0.04	-2,928	2.56
18	0.0165	0.04	-2,788	2.60
19	0.0173	0.04	-2,664	2.69
20	0.0182	0.03	-2,532	2.75

Fits are based on the “minres” factoring method and “oblimin” rotation procedures using the ‘nfactors’ function in the *psych* package in R. Minimum values are bolded for eBIC and MAP.

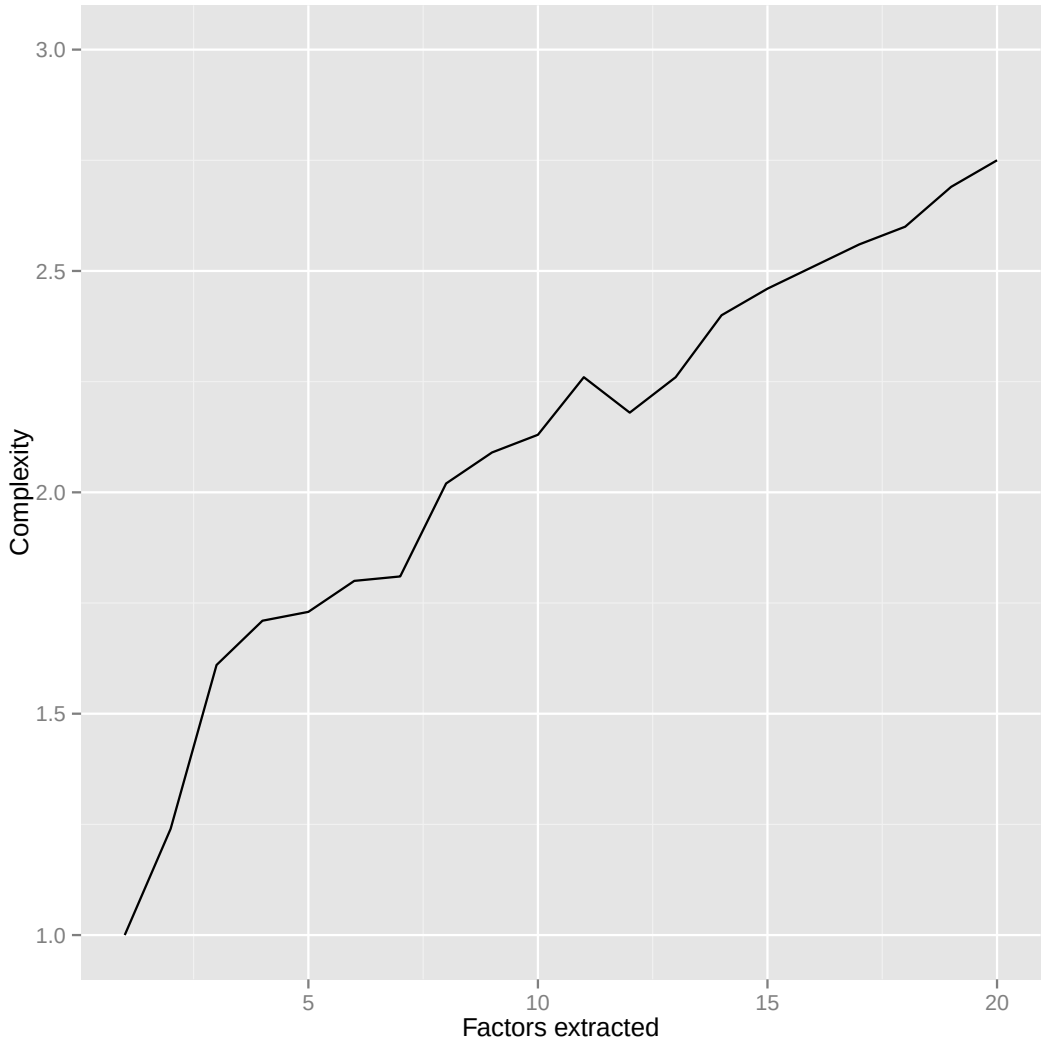


Figure 5.1: Complexity based on extraction of 1 to 20 factors from the 60 O*NET items

Table 5.2: Item content of extracted factors from 1 through 10

Extracted	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
1	Realistic/ Investigative									
eigen	11.53									
2	Realistic/ Investigative/ Conventional	Social/ Enterprising/ Artistic								
eigen	9.93	6.55								
3	Investigative/ Realistic	Conventional/ Enterprising	Social/ Artistic/ Enterprising							
eigen	8.50	6.17	5.97							
4	Conventional Enterprising Realistic	Investigative	Artistic	Social						
eigen	7.84	5.67	5.52	4.68						
5	Conventional/ Realistic	Investigative	Artistic	Enterprising	Social					
eigen	6.78	5.54	5.22	4.51	4.24					
6	Realistic	Investigative	Artistic	Conventional	Social	Enterprising				
eigen	5.69	5.04	4.82	4.31	4.28	4.25				
7	Realistic	Investigative	Artistic	Conventional	Social	Enterprising	Enterprising			
eigen	5.54	4.87	4.71	4.35	4.04	3.43	3.42			
8	Realistic	Investigative	Conventional	Social	Enterprising	Enterprising	Artistic	Artistic		
eigen	5.71	4.90	4.27	4.01	3.52	3.42	3.26	2.66		
9	Realistic	Investigative	Conventional	Social	Enterprising	Artistic	Enterprising	Artistic	unidentified	
eigen	5.05	4.85	4.40	3.72	3.55	3.36	3.03	2.89	2.24	
10	Investigative	Realistic	Conventional	Enterprising	Artistic	Social	Enterpristing	Artistic	Social/ Investigative	Conventional/ Realistic
eigen	4.77	4.61	3.99	3.46	3.35	3.21	3.04	2.81	2.62	2.49

For the six-factor solution, each factor was made up of items corresponding to one of the Holland types. Three of the sixty items (5%) did not fit into the Holland types structure (see Table 5.3 for the loadings by item for the six-factor solution). Two of these items, "Would like to investigate the cause of a fire" and "Would like to load computer software into a large computer network," were supposed to be Investigative and Conventional items, respectively, but both had primary loadings on the Realistic factor. These two items had the lowest loadings (0.36 and 0.35, respectively) on the "new" Realistic factor. The other mis-matched item, "Would like to operate a beauty salon or barber shop," was supposed to be an Enterprising item, but loaded onto Social interests. This item had the lowest loading (0.33) on the Social factor.

Table 5.3: 6 factor scales and loadings based on the 60 O*NET items

Item	Realistic	Investigative	Artistic	Social	Enterprising	Conventional	c_i	h^2	Scale
Realistic									
Would like to repair household appliances.	0.78	-0.02	-0.02	0.10	-0.06	0.07	1.06	0.62	Realistic
...repair and install locks.	0.69	0.02	0.05	0.01	-0.06	0.20	1.19	0.60	Realistic
...build kitchen cabinets.	0.69	0.02	0.07	0.12	0.02	0.02	1.09	0.54	Realistic
...assemble electronic parts.	0.69	0.07	0.11	-0.23	0.07	-0.01	1.32	0.63	Realistic
...set up and operate machines to make products.	0.68	0.10	0.01	-0.16	0.13	0.07	1.27	0.64	Realistic
...lay brick or tile.	0.64	-0.10	0.08	0.13	-0.09	0.14	1.32	0.46	Realistic
...put out forest fires.	0.48	0.19	0.08	0.23	0.06	-0.17	2.19	0.39	Realistic
...drive a truck to deliver packages to offices and homes.	0.42	-0.08	0.14	0.04	-0.06	0.25	2.04	0.31	Realistic
...raise fish in a fish hatchery.	0.42	0.08	0.14	0.17	-0.03	0.02	1.68	0.28	Realistic
...test the quality of parts before shipment.	0.41	0.08	-0.06	-0.02	0.07	0.29	2.04	0.37	Realistic
...investigate the cause of a fire.	0.36	0.28	0.03	0.06	0.09	0.08	2.23	0.34	Investigative
...load computer software into a large computer network.	0.35	0.17	0.10	-0.31	0.21	0.16	3.84	0.46	Conventional
Investigative									
...do laboratory tests to identify diseases.	-0.12	0.90	-0.08	0.07	-0.01	0.09	1.09	0.77	Investigative
...examine blood samples using a microscope.	-0.07	0.81	0.04	0.05	-0.04	0.10	1.06	0.68	Investigative
...work in a biology lab.	0.00	0.80	0.04	-0.02	-0.08	0.00	1.03	0.65	Investigative
...conduct chemical experiments.	0.13	0.78	0.04	-0.13	0.02	-0.04	1.12	0.71	Investigative
...develop a new medicine.	-0.02	0.78	0.02	0.14	0.12	-0.10	1.14	0.65	Investigative
...study the movement of planets.	0.23	0.52	0.23	-0.15	-0.05	-0.07	2.08	0.50	Investigative
...study ways to reduce water pollution.	0.28	0.45	0.13	0.20	-0.05	-0.01	2.38	0.46	Investigative
...invent a replacement for sugar.	0.16	0.40	0.07	0.08	0.14	0.02	1.81	0.31	Investigative
...develop a way to better predict the weather.	0.30	0.37	0.05	-0.01	0.11	0.09	2.29	0.39	Investigative
Artistic									
...write scripts for movies or television shows.	0.00	0.02	0.77	-0.01	0.16	-0.09	1.12	0.64	Artistic
...compose or arrange music.	0.07	0.07	0.73	-0.10	-0.02	-0.05	1.08	0.56	Artistic
...write books or plays.	-0.14	0.08	0.68	0.01	-0.05	-0.02	1.13	0.47	Artistic
...edit movies.	0.16	-0.01	0.65	-0.07	0.15	0.00	1.26	0.54	Artistic
...draw pictures.	0.05	0.00	0.61	0.05	-0.14	0.05	1.14	0.40	Artistic
...sing in a band.	-0.03	-0.03	0.60	0.10	-0.06	0.03	1.10	0.37	Artistic
...play a musical instrument.	0.00	0.09	0.59	0.01	-0.12	0.02	1.13	0.37	Artistic
...create special effects for movies.	0.30	0.03	0.57	-0.09	0.13	-0.06	1.74	0.51	Artistic
...paint sets for plays.	0.12	-0.02	0.55	0.18	-0.12	0.14	1.58	0.43	Artistic
...perform jazz or tap dance.	-0.09	0.03	0.46	0.18	0.06	0.15	1.68	0.32	Artistic

c_i = item complexity; h^2 = communality

continued on next page

Table 5.3 – 6 factor scales and loadings (continued from previous page)

Item	Realistic	Investigative	Artistic	Social	Enterprising	Conventional	c_i	h^2	Scale
Social									
...help conduct a group therapy session.	-0.06	0.07	0.07	0.68	0.08	-0.03	1.10	0.51	Social
...perform rehabilitation therapy.	-0.01	0.17	-0.02	0.63	-0.05	0.09	1.20	0.45	Social
...help people with personal or emotional problems.	-0.05	-0.04	0.07	0.62	-0.08	-0.01	1.08	0.40	Social
...teach sign language to people with hearing disabilities.	0.08	0.14	0.10	0.60	-0.11	0.09	1.34	0.45	Social
...take care of children at a day-care center.	-0.04	-0.06	-0.01	0.59	-0.06	0.13	1.15	0.37	Social
...do volunteer work at a non-profit organization.	-0.04	0.14	0.01	0.59	-0.13	0.08	1.25	0.38	Social
...teach children how to play sports.	0.26	-0.05	-0.17	0.55	0.29	-0.14	2.45	0.44	Social
...give career guidance to people.	0.13	-0.02	-0.03	0.51	0.31	-0.01	1.83	0.41	Social
...teach an individual an exercise routine.	0.27	0.05	-0.10	0.44	0.30	-0.19	3.11	0.37	Social
...teach a high-school class.	0.18	0.02	0.16	0.42	0.10	-0.08	1.94	0.28	Social
...operate a beauty salon or barber shop.	-0.05	-0.09	0.18	0.33	0.19	0.17	3.10	0.27	Enterprising
Enterprising									
...negotiate business contracts.	0.02	0.05	-0.04	-0.07	0.78	0.06	1.04	0.63	Enterprising
...manage a department within a large company.	0.00	0.01	-0.05	-0.01	0.65	0.11	1.07	0.46	Enterprising
...start your own business.	0.11	-0.06	0.09	0.00	0.64	-0.15	1.24	0.44	Enterprising
...buy and sell stocks and bonds.	0.22	0.11	-0.06	-0.21	0.58	0.03	1.69	0.50	Enterprising
...represent a client in a lawsuit.	-0.09	0.20	0.06	0.08	0.57	-0.11	1.48	0.38	Enterprising
...manage a retail store.	0.00	-0.09	0.10	0.17	0.54	0.23	1.70	0.48	Enterprising
...market a new line of clothing.	-0.20	-0.07	0.34	0.15	0.53	0.07	2.34	0.49	Enterprising
...manage a clothing store.	-0.21	-0.16	0.23	0.21	0.50	0.27	3.19	0.54	Enterprising
...sell merchandise at a department store.	-0.13	-0.12	0.10	0.19	0.40	0.28	2.96	0.35	Enterprising
Conventional									
...record rent payments.	-0.06	0.09	-0.12	0.01	0.10	0.70	1.15	0.54	Conventional
...stamp, sort, and distribute mail for an organization.	0.14	-0.07	0.09	0.11	-0.16	0.69	1.31	0.54	Conventional
...keep inventory records.	0.17	0.00	-0.06	-0.02	0.04	0.68	1.15	0.57	Conventional
...keep shipping and receiving records.	0.09	0.01	0.02	0.02	0.05	0.67	1.05	0.52	Conventional
...inventory supplies using a hand-held computer.	0.18	0.06	0.00	-0.06	0.06	0.61	1.24	0.50	Conventional
...proofread records or forms.	-0.13	0.14	0.16	0.04	0.01	0.59	1.40	0.40	Conventional
...calculate the wages of employees.	0.10	0.10	-0.19	-0.03	0.38	0.47	2.52	0.53	Conventional
...develop a spreadsheet using computer software.	0.23	0.15	0.00	-0.24	0.20	0.37	3.51	0.42	Conventional
...operate a calculator.	0.29	0.21	-0.14	-0.09	0.10	0.37	3.23	0.41	Conventional

c_i = item complexity; h^2 = communality

The seven-factor solution closely matched the structure of the six-factor solution, except the Enterprising factor was bifurcated into two factors. The ten Enterprising items were evenly split 5-5, creating Enterprising sub-factors that could be summarized as "corporate/law" and "retail/marketing." Accounting for the seven-factor solution bifurcating the Enterprising category into two factors, two of the sixty items (3%) did not fit into the Holland types structure.

5.2.3 Discussion

While both the MAP and eBIC suggested that the seven factor solution provide the best fit, the mean item complexity of the seven factor solution was essentially identical to the six-factor solution and the SRMR suggested that the six-factor solution was also good. The primary difference between the seven and six factor solutions was the splitting of the Enterprising factor into corporate/law and retail/marketing sub-scales. The critical question seems to be whether these two Enterprising factors have the same level of categorical distinction as the other five factors. If not, the seven-factor solution can be seen as the point at which the six primary factors begin to bifurcate into sub-factors. The fact that the Enterprising factor split to create these new factors points to the suggestion that seven factors is over-extraction, and the first sub-factor has appeared.

It is also worth noting that the two and three factor solutions are largely consistent with hierarchical grouping of the the six-factor solution as has been hypothesized previously. When considered in conjunction with the splitting of the Enterprising factor, it is possible that both lower and higher order structures can be identified within the six O*NET® Interest Profiler Short Form scales. This prospect, however, lies beyond the scope of the current study and is left for future research.

5.3 Study 2

Study 2 evaluated the structural properties of the 92-item Oregon Vocational Interest Scale ("ORVIS") (Pozzebon et al., 2010). It was hypothesized that the extraction of six factors would

result in factors with item content which largely correspond to the six Holland occupational interests types (see Study 1) and that the eight-factor solution would largely match the prescribed factors of the eight ORVIS scales: Leadership, Organization, Altruism, Creativity, Analysis, Producing, Adventuring, and Erudition. Finally, it was also hypothesized that the eight factor solution would demonstrate a superior fit to the data relative to the factor solution with six (and any other number of) factors.

5.3.1 Methods

5.3.1.1 Participants

The data were collected from participants who voluntarily visited the SAPA-Project.org website between May 20, 2013 and March 13, 2014 in exchange for customized feedback about their personalities (note that this sample overlapped with the sample used in Study 1). This included 35,856 participants (53% female) from 170 countries. All data were self-reported. The mean age was 25.4 years (*sd* = 9.8, median = 22) with a range from 14 to 90 years.

19% of participants reported educational attainment of a high school degree or less, 59% some college, currently in college, or a college degree, 12% some graduate school or a graduate degree, and 10% chose not to report their education level. Of the 23,190 participants who reported being from the United States, 63.5% identified themselves as White/Caucasian, 6.8% as African American, 7.7% as Hispanic-American, 4.4% as Asian-American, 0.9% as Native Alaskan/Hawaiian-American, 5.7% as multi-ethnic, 1.1% as "Other," and 9.9% did not specify their racial/ethnic background. Participants from outside the United States were not prompted for information about their racial/ethnic background.

5.3.1.2 Analyses and Measures

The same analyses and fit statistics were used in Study 2 as those in Study 1. The Synthetic Aperture Personality Assessment ("SAPA") technique (Revelle et al., 2010b), a matrix sampling procedure, was used for item administration. Participants were given 24 item subsets of the 92-item ORVIS item pool by default but all participants had the option of completing more items (up to a maximum of 65); the number of items to which participants responded varied by participant willingness to take more items. Approximately 60% of the sample responded to exactly 24 ORVIS items. The items were administered a mean of 9,585 times and the number of pairwise administrations (median = 3,050, mean = 2,999, min = 1,883) provided sufficiently high stability in the covariance matrix for the structural analyses described below (Kenny, 2012)

5.3.2 Results

Table 5.4 shows the fit statistics based on the extraction of 1 to 20 factors. SRMR suggested that more than 4 factors provided good fit while the MAP criterion demonstrated a local minimum at 11 factors. The empirical BIC statistic did not demonstrate a local minimum at less than the maximum number of extracted factors. Complexity values (see also Figure 5.2) indicated a sharp decrease in mean item complexity at 6 factors before increasing steadily from 7 to 9 factors and then decreasing again at 10 factors.

Table 5.5 describes the content of each factor for the factor solutions with 1 to 10 factors of extraction, with the factors in each case sorted by the magnitude of the eigenvalues. In several cases, the items from an ORVIS scale were split among factors. For example, in the seven and eight factor solutions, the Leadership and Erudition items split into two mixed scales rather than into scales made up exclusively of items from the same scale. The six factor solution did not demonstrate mixing of items from one of the original scales across two factors, but it did

Table 5.4: Fit statistics for extraction of 1 to 20 factors from the 92 ORVIS items

Factors	MAP	SRMR	eBIC	complexity
1	0.0162	0.11	261,434	1.00
2	0.0133	0.09	168,195	1.33
3	0.0099	0.07	93,581	1.58
4	0.0086	0.06	62,480	1.74
5	0.0072	0.05	36,271	1.96
6	0.0061	0.04	19,525	1.90
7	0.0054	0.04	8,286	2.11
8	0.0049	0.03	1,022	2.29
9	0.0048	0.03	-2,470	2.36
10	0.0047	0.03	-4,796	2.31
11	**0.0047**	0.03	-6,749	2.36
12	0.0047	0.03	-8,067	2.39
13	0.0048	0.02	-9,458	2.38
14	0.0048	0.02	-10,285	2.45
15	0.0048	0.02	-10,939	2.50
16	0.0049	0.02	-11,531	2.59
17	0.0050	0.02	-11,799	2.65
18	0.0051	0.02	-12,109	2.70
19	0.0053	0.02	-12,246	2.73
20	0.0054	0.02	**-12,299**	2.70

Fits are based on the "minres" factoring method and "oblimin" rotation procedures using the 'nfactors' function in the *psych* package in R. Minimum values are bolded for eBIC and MAP.

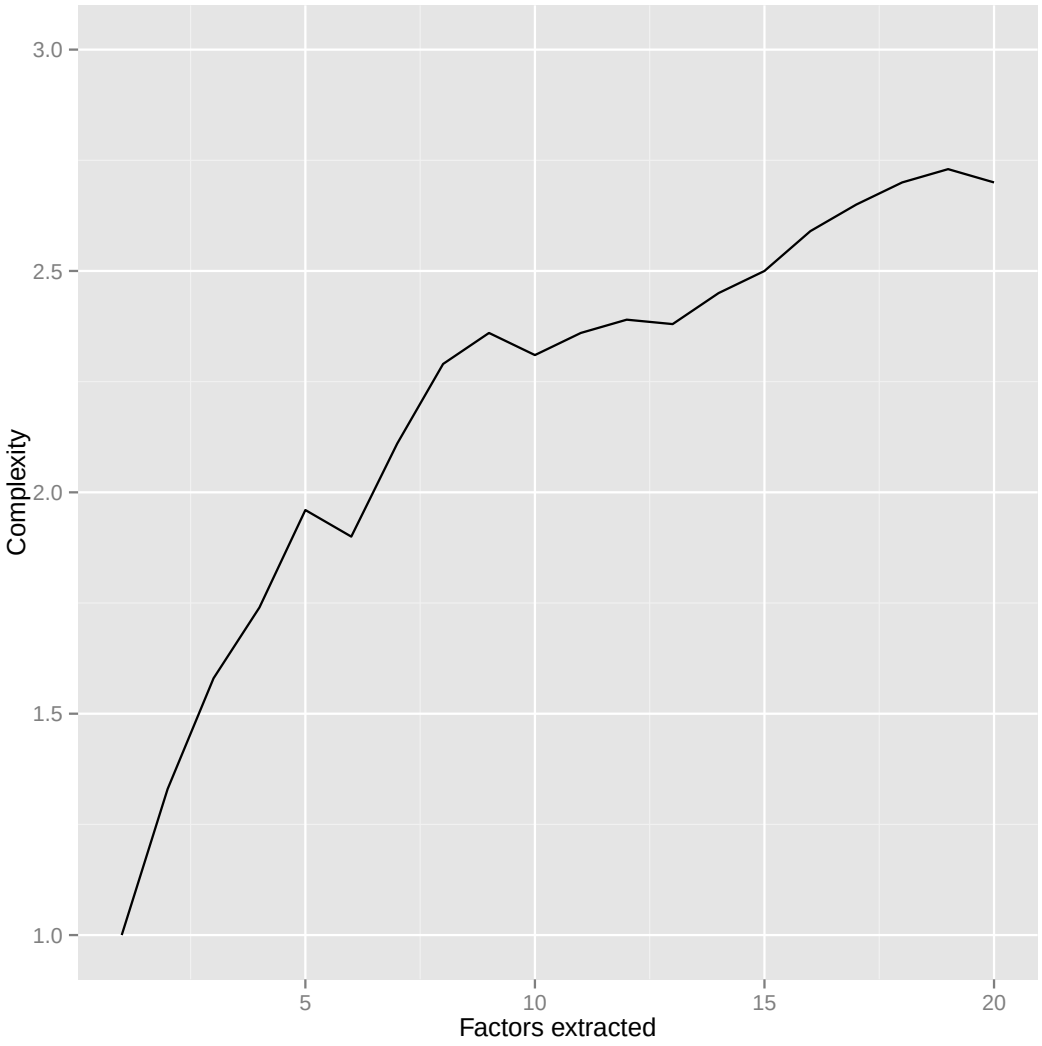


Figure 5.2: Complexity based on extraction of 1 to 20 factors

demonstrate blends of scales into single factors (Creativity/Erudition and Adventure/Production).

Table 5.5: Item content of extracted factors from 1 through 10

Extracted	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
1	Leadership/ Analysis									
eigen	12.43									
2	Analysis	Altruism/ Creativity								
eigen	9.99	8.18								
3	Leadership/ Organization	Creativity/ Erudition/	Analysis/ Production/ Altruism							
eigen	7.98	7.77	7.55							
4	Leadership/ Organization	Erudition/ Creativity	Adventure/ Production	Altruism/ Analysis						
eigen	8.14	7.15	6.23	5.17						
5	Creativity/ Erudition	Adventure/ Production	Leadership	Organization	Altruism					
eigen	6.57	6.29	5.93	5.78	5.27					
6	Creativity/ Erudition	Leadership	Analysis	Organization	Altruism	Adventure/ Production				
eigen	5.97	5.51	5.47	5.44	5.05	4.98				
7	Organization	Analysis	Leadership/ Erudition	Altruism	Production/ Adventure	Creativity	Erudtion/ Leadership			
eigen	5.57	5.52	5.40	5.24	4.90	4.80	3.16			
8	Analysis	Organization	Creativity	Altruism	Leadership/ Erudition	Production/ Adventure	Leadership/ -Erudition	Adventure		
eigen	5.64	5.60	5.20	5.09	4.71	4.26	3.01	2.83		
9	Organization	Analysis	Creativity	Production/ Adventure	Altruism	Leadership	Erudition/ Leadership	Creativity/ Adventure/	Altruism/ Analysis/ Creativity	
eigen	5.51	5.43	4.89	4.84	4.66	4.66	3.17	2.40	2.18	
10	Organization	Production/ Adventure	Altruism	Leadership	Analysis	Creativity	Erudition	Creativity	Analysis/ -Creativity	Leadership
eigen	5.47	4.97	4.21	4.04	3.77	3.64	3.44	3.41	3.35	2.58

The organization of item content and factor loadings for the six and eight factor solutions are given in Tables 5.6 and 5.7, respectively. The six-factor solution organized items into a pattern resembling the Holland types. Four of the six factors closely matched a Holland type; Analysis with Investigative, Organizational with Conventional, Altruism with Social, and Leadership with Enterprising. In the fifth factor, the combination of Creativity and Erudition closely matched Artistic. The last factor was composed of Production and Adventure item sets, which is consistent with ORVIS theory that those two item sets would combine to become the Holland Realistic type (Pozzebon et al., 2010). Assuming consistency among the scales in this way, only seven of the ninety-two ORVIS items (8%) did not "fit" into a six factor structure matching the Holland types.

The eight-factor solution roughly matched the structure of the six-factor solution. The seventh factor was comprised of 2 Leadership items, 1 Adventure item, and 4 negatively keyed Erudition items. The eighth factor contained 5 Adventure items and 1 negatively keyed Erudition item.

Table 5.6: 6 factor scales and loadings based on the 92 ORVIS items

Item	Creativity/ Erudition	Organization	Leadership	Analysis	Altruism	Adventure/ Production	c_i	h^2	Scale
Creativity/Erudition									
Would like to be an artist or architect.	0.68	-0.05	0.02	0.01	-0.13	0.20	1.26	0.53	Creativity
Create works of art.	0.67	-0.03	-0.04	-0.12	0.01	0.12	1.15	0.45	Creativity
Would like to select art works for a museum.	0.66	0.01	-0.01	0.01	0.05	0.04	1.02	0.46	Creativity
Would like to write short stories or novels.	0.58	-0.12	0.15	0.11	0.03	-0.11	1.40	0.40	Creativity
Paint or draw.	0.54	-0.01	-0.15	-0.13	0.01	0.19	1.52	0.34	Creativity
Would like to be a professor of English.	0.52	-0.07	0.17	0.16	0.12	-0.17	1.83	0.39	Erudition
Would like to edit a newspaper.	0.50	0.07	0.19	0.18	0.04	-0.16	1.85	0.38	Erudition
Would like to be a librarian.	0.47	0.04	-0.18	0.30	0.03	-0.11	2.20	0.36	Erudition
Like to create new fashion designs.	0.44	0.13	0.04	-0.17	0.15	0.04	1.83	0.26	Creativity
Would like to act in a play.	0.42	-0.08	0.34	-0.02	0.18	-0.02	2.40	0.37	Creativity
Write songs.	0.41	-0.10	0.15	0.03	0.05	0.07	1.50	0.23	Creativity
Would like to be an actor or actress.	0.39	-0.09	0.39	-0.10	0.09	0.03	2.35	0.34	Creativity
Would like to sing professionally.	0.39	-0.07	0.22	-0.08	0.11	0.03	2.00	0.24	Creativity
Would like to design Internet web pages.	0.39	0.24	0.04	0.22	-0.19	0.00	2.94	0.35	Creativity
Would like to play an instrument in a symphony.	0.39	-0.04	0.03	0.20	0.07	0.08	1.74	0.26	Creativity
Keep a diary or journal.	0.34	0.04	-0.07	-0.06	0.17	-0.06	1.72	0.16	Erudition
Would like to be a translator or interpreter.	0.34	0.04	0.06	0.18	0.15	0.03	2.14	0.23	Erudition
Read many books.	0.30	-0.07	-0.02	0.21	0.12	-0.13	2.79	0.17	Erudition
Make up word puzzles.	0.29	0.10	0.03	0.25	0.04	-0.02	2.29	0.21	Erudition
Would like to be a professional dancer.	0.28	0.02	0.05	-0.04	0.25	0.07	2.26	0.18	Creativity
Would like to know many languages.	0.26	-0.04	0.11	0.14	0.09	0.04	2.35	0.14	Erudition
Would like to be a minister, priest, rabbi or other religious teacher.	0.17	0.10	0.10	0.00	0.14	0.06	3.60	0.10	Altruism
Organization									
Like to monitor business expenses.	-0.05	0.71	0.04	0.01	-0.01	0.04	1.02	0.54	Organization
Would like to keep track of a companys inventory.	0.04	0.65	-0.13	0.04	0.01	0.05	1.11	0.42	Organization
Like to prepare financial contracts.	-0.06	0.65	0.10	0.06	-0.03	0.05	1.10	0.50	Organization
Would like to be the financial officer for a company.	-0.08	0.62	0.20	0.06	-0.13	0.03	1.35	0.53	Organization
Like to plan budgets.	-0.06	0.62	-0.11	-0.02	0.14	0.00	1.19	0.38	Organization
Would like to be an office manager.	-0.04	0.60	0.17	-0.09	0.11	-0.04	1.29	0.44	Organization
Would like to develop an office filing system.	0.15	0.58	-0.11	0.11	0.09	-0.09	1.39	0.37	Organization
Like to plan investment strategies.	-0.10	0.55	0.15	0.07	-0.03	0.11	1.36	0.43	Organization
Keep detailed records.	0.01	0.46	-0.16	0.01	0.21	-0.09	1.74	0.24	Organization

c_i = item complexity; h^2 = communality

continued on next page

Table 5.6 – 6 factor scales and loadings (continued from previous page)

Item	Creativity/ Erudition	Organization	Leadership	Analysis	Altruism	Adventure/ Production	c_i	h^2	Scale
Would like to be a purchasing agent.	0.11	0.46	0.16	-0.03	0.01	0.08	1.45	0.31	Organization
Like to establish time schedules.	-0.06	0.40	-0.11	-0.04	0.27	-0.09	2.16	0.22	Organization
Would like to be a sales or marketing director.	0.12	0.39	0.37	-0.15	-0.01	0.04	2.52	0.37	Leadership
Like to supervise the work of others.	-0.10	0.31	0.29	-0.07	0.22	0.02	3.21	0.29	Organization
Leadership									
Would like to run for political office.	0.07	0.00	0.70	0.10	-0.06	0.05	1.09	0.55	Leadership
Would like to be a state governor or senator.	0.04	0.05	0.70	0.14	-0.04	0.04	1.11	0.58	Leadership
Would like to organize a political campaign.	0.13	0.04	0.63	0.13	-0.01	-0.03	1.19	0.48	Leadership
Would like to make decisions that affect a lot of people.	-0.10	0.10	0.54	0.02	0.14	0.06	1.32	0.38	Leadership
Like to debate topics in a public meeting.	0.05	-0.02	0.53	0.12	0.06	0.03	1.16	0.33	Leadership
Would like to be the chief executive of a large company.	-0.08	0.38	0.51	-0.04	-0.06	0.08	2.01	0.52	Leadership
Like to lead other people.	-0.14	0.11	0.49	-0.11	0.27	0.12	2.14	0.41	Leadership
Would like to be the master of ceremonies at a meeting.	0.11	0.16	0.46	-0.03	0.19	0.01	1.76	0.37	Leadership
Persuade others to change their views.	-0.01	-0.03	0.43	0.10	0.04	0.03	1.14	0.21	Leadership
Would like to plan an advertising campaign.	0.35	0.25	0.41	-0.08	-0.03	-0.03	2.76	0.42	Leadership
Would like to be a foreign correspondent.	0.30	0.00	0.35	0.15	0.08	0.08	2.57	0.33	Erudition
Like to make important things happen.	-0.09	0.09	0.31	-0.01	0.29	0.08	2.52	0.23	Leadership
Speak fluently on any subject.	0.07	0.00	0.31	0.18	0.07	0.02	1.86	0.17	Erudition
Analysis									
Would like to be a scientific reporter.	0.11	-0.03	0.09	0.73	0.03	-0.02	1.09	0.59	Analysis
Would like to design a laboratory experiment.	0.02	-0.02	0.04	0.72	0.02	0.10	1.05	0.57	Analysis
Would like to be a chemist.	-0.01	0.01	-0.01	0.69	-0.01	0.11	1.05	0.53	Analysis
Like to explain scientific concepts to others.	-0.03	-0.05	0.13	0.65	0.02	0.06	1.11	0.47	Analysis
Would like to be a physicist.	0.00	0.03	0.07	0.65	-0.07	0.10	1.10	0.51	Analysis
Would like to carry out medical research.	-0.14	-0.04	-0.01	0.63	0.34	0.05	1.66	0.47	Analysis
Would like to be a mathematician.	-0.02	0.24	-0.04	0.55	-0.12	0.03	1.51	0.44	Analysis
Would like to be a statistician.	0.00	0.35	0.01	0.45	-0.09	-0.03	1.98	0.40	Analysis
Like to solve complex puzzles.	0.07	0.10	0.01	0.43	-0.02	0.07	1.22	0.26	Analysis
Would like to manage a computer data base.	0.06	0.34	-0.02	0.43	-0.25	-0.01	2.62	0.43	Organization
Like to develop computer programs.	0.12	0.19	0.02	0.42	-0.32	0.06	2.54	0.41	Analysis
Altruism									
Like to care for sick people.	-0.10	0.01	-0.11	0.02	0.65	0.06	1.13	0.41	Altruism
Like to provide comfort and support to others.	0.01	0.01	-0.02	-0.10	0.65	-0.03	1.06	0.43	Altruism
Like to counsel persons who need help.	0.04	0.00	0.14	0.01	0.61	-0.08	1.14	0.42	Altruism

c_i = item complexity; h^2 = communality

continued on next page

Table 5.6 – 6 factor scales and loadings (continued from previous page)

Item	Creativity/ Erudition	Organization	Leadership	Analysis	Altruism	Adventure/ Production	c_i	h^2	Scale
Would like to be a counselor or therapist.	0.16	-0.04	0.08	0.06	0.58	-0.11	1.31	0.40	Altruism
Would like to be a social worker.	0.14	0.03	0.01	0.01	0.57	-0.05	1.15	0.37	Altruism
Like to participate in charity events.	0.06	0.09	0.05	-0.09	0.52	0.03	1.18	0.32	Altruism
Would like to instruct parents on child care.	0.06	0.09	0.03	0.03	0.51	0.01	1.10	0.30	Altruism
Would like to be a doctor or nurse.	-0.19	-0.06	0.00	0.36	0.46	0.10	2.42	0.32	Altruism
Would like to help people make career decisions.	0.06	0.27	0.23	0.03	0.40	-0.03	2.49	0.37	Altruism
Like to help others learn new ideas.	0.06	0.01	0.11	0.18	0.40	0.00	1.56	0.23	Altruism
Would like to be a physical therapist.	-0.02	0.04	-0.02	0.11	0.40	0.20	1.66	0.22	Altruism
Would like to be an elementary-school teacher.	0.22	0.05	0.00	0.02	0.36	0.00	1.69	0.21	Altruism
Like to redecorate my house.	0.23	0.19	-0.09	-0.23	0.33	0.04	3.56	0.25	Creativity
Like to grow flowers.	0.30	0.11	-0.25	0.01	0.31	0.21	4.01	0.30	Production
Adventure/Production									
Like to survive in the wilderness.	0.03	-0.08	0.03	0.08	0.06	0.63	1.09	0.44	Adventure
Like to work with tools and machinery.	0.01	0.12	-0.08	0.17	-0.13	0.58	1.41	0.47	Production
Like woodworking.	0.19	0.08	-0.11	0.11	-0.05	0.57	1.44	0.44	Production
Like to face physical danger.	-0.10	-0.12	0.20	-0.02	-0.01	0.56	1.41	0.35	Adventure
Like to repair cars or trucks.	-0.06	0.11	0.01	0.13	-0.07	0.56	1.24	0.40	Production
Would like to be a forest ranger.	0.20	-0.10	-0.07	0.16	0.00	0.54	1.60	0.42	Production
Would like to be a farmer.	0.21	0.04	-0.17	0.06	0.09	0.50	1.71	0.35	Production
Would like to be a racing car driver.	-0.01	0.07	0.19	0.00	-0.08	0.48	1.41	0.32	Adventure
Like to compete in athletic events.	-0.24	0.04	0.20	-0.11	0.14	0.44	2.38	0.30	Adventure
Would like to be a bounty hunter.	-0.01	-0.02	0.25	0.03	-0.11	0.41	1.84	0.27	Adventure
Would like to be a military officer.	-0.14	0.06	0.26	0.04	-0.05	0.41	2.08	0.30	Adventure
Would like to be a professional athlete.	-0.12	0.02	0.24	-0.02	0.02	0.40	1.84	0.25	Adventure
Would like to construct new buildings.	0.18	0.18	0.09	0.18	-0.15	0.40	2.89	0.42	Production
Would like to care for cattle or horses.	0.17	0.02	-0.15	0.00	0.20	0.39	2.27	0.25	Production
Engage in exciting adventures.	-0.04	-0.02	0.17	-0.15	0.27	0.39	2.60	0.27	Adventure
Would like to be a long-distance bicycle rider.	0.13	0.03	0.03	0.13	0.04	0.36	1.59	0.23	Adventure
Go on nature walks.	0.20	-0.02	-0.15	0.02	0.23	0.34	2.94	0.25	Production
Would like to be a police officer.	-0.07	0.01	0.20	0.00	0.04	0.34	1.76	0.17	Adventure
Cultivate plants.	0.23	0.12	-0.20	0.04	0.19	0.31	3.73	0.25	Production

c_i = item complexity; h^2 = communality

Table 5.7: 8 factor scales and loadings based on the 92 O*NET items

Item	Organiz.	Analysis	Altruism	Creativity	Leader./ Erudtion	Production/ Adventure	Leader./ -Erudition	Adventure	c_i	h^2	Scale
Organization											
Like to monitor business expenses.	0.72	0.01	-0.03	-0.05	0.03	0.04	0.05	-0.04	1.04	0.54	Organization
Would like to keep track of a companys inventory.	0.66	0.02	0.02	-0.02	-0.10	0.10	-0.12	0.00	1.16	0.44	Organization
Like to prepare financial contracts.	0.66	0.06	-0.03	-0.06	0.08	0.03	0.01	0.03	1.08	0.50	Organization
Would like to be the financial officer for a company.	0.64	0.05	-0.10	-0.08	0.14	-0.04	-0.03	0.14	1.31	0.54	Organization
Would like to be an office manager.	0.62	-0.09	0.14	-0.03	0.07	-0.08	0.02	0.08	1.25	0.45	Organization
Like to plan budgets.	0.61	-0.03	0.08	-0.06	-0.05	0.09	0.08	-0.19	1.36	0.40	Organization
Would like to develop an office filing system.	0.57	0.08	0.09	0.06	-0.03	0.03	-0.13	-0.15	1.38	0.38	Organization
Like to plan investment strategies.	0.55	0.10	-0.06	-0.05	0.10	0.04	0.15	0.02	1.36	0.43	Organization
Would like to be a purchasing agent.	0.49	-0.03	0.06	0.09	0.08	0.02	-0.09	0.20	1.58	0.35	Organization
Keep detailed records.	0.44	0.01	0.13	0.01	-0.09	0.02	0.11	-0.29	2.18	0.29	Organization
Would like to be a sales or marketing director.	0.42	-0.11	0.04	0.16	0.21	-0.10	0.04	0.22	2.84	0.39	Leadership
Would like to be the chief executive of a large company.	0.40	0.03	-0.06	0.06	0.30	-0.13	0.27	0.21	3.68	0.54	Leadership
Like to establish time schedules.	0.38	-0.04	0.21	-0.06	-0.08	0.00	0.12	-0.22	2.67	0.24	Organization
Like to supervise the work of others.	0.31	-0.02	0.17	-0.01	0.17	-0.05	0.31	-0.04	3.23	0.32	Organization
Analysis											
Would like to design a laboratory experiment.	-0.03	0.76	0.01	0.03	0.02	0.05	0.05	-0.02	1.03	0.60	Analysis
Like to explain scientific concepts to others.	-0.07	0.72	-0.03	0.03	0.10	0.01	0.19	-0.12	1.28	0.56	Analysis

c_i = item complexity; h^2 = communality

continued on next page

Table 5.7 – 8 factor scales and loadings (continued from previous page)

Item	Organization	Analysis	Altruism	Creativity	Leadership/ Erudtion	Production/ Adventure	Leadership/ -Erudition	Adventure	c_i	h^2	Scale
Would like to be a scientific reporter.	-0.04	0.70	0.07	0.02	0.14	0.03	-0.13	-0.04	1.19	0.58	Analysis
Would like to be a chemist.	0.00	0.68	0.02	-0.05	0.01	0.09	-0.09	0.05	1.09	0.53	Analysis
Would like to be a physicist.	0.03	0.66	-0.03	-0.02	0.04	0.04	-0.07	0.11	1.11	0.52	Analysis
Would like to carry out medical research.	-0.05	0.63	0.35	-0.14	-0.05	0.03	0.02	0.06	1.75	0.47	Analysis
Would like to be a mathematician.	0.23	0.58	-0.11	-0.01	-0.05	-0.01	-0.05	0.02	1.44	0.44	Analysis
Like to solve complex puzzles.	0.08	0.51	-0.07	0.15	-0.03	0.03	0.18	-0.13	1.68	0.35	Analysis
Like to develop computer programs.	0.18	0.49	-0.29	0.15	-0.03	-0.03	-0.05	0.08	2.32	0.43	Analysis
Would like to manage a computer data base.	0.35	0.46	-0.20	0.05	-0.05	-0.06	-0.13	0.06	2.63	0.45	Organization
Would like to be a statistician.	0.36	0.44	-0.04	-0.05	0.01	-0.03	-0.15	0.04	2.26	0.41	Analysis
Make up word puzzles.	0.09	0.27	0.05	0.26	0.05	0.02	-0.04	-0.11	2.73	0.22	Erudition
Altruism											
Would like to be a counselor or therapist.	-0.03	0.03	0.65	0.08	0.02	-0.06	-0.11	0.06	1.13	0.46	Altruism
Would like to be a social worker.	0.04	-0.07	0.64	-0.01	0.05	0.07	-0.18	0.02	1.23	0.45	Altruism
Like to care for sick people.	0.01	0.01	0.61	-0.09	-0.13	0.10	0.12	-0.03	1.29	0.39	Altruism
Like to provide comfort and support to others.	0.00	-0.08	0.61	0.06	-0.10	-0.02	0.20	-0.07	1.36	0.44	Altruism
Like to counsel persons who need help.	-0.01	0.03	0.59	0.06	0.06	-0.07	0.19	-0.09	1.33	0.43	Altruism
Would like to instruct parents on child care.	0.09	0.01	0.53	0.01	0.01	0.05	0.00	0.01	1.09	0.31	Altruism
Would like to be a doctor or nurse.	-0.05	0.33	0.49	-0.20	-0.06	0.06	0.02	0.15	2.51	0.34	Altruism
Would like to be a physical therapist.	0.07	0.10	0.48	-0.06	-0.11	0.13	-0.10	0.31	2.30	0.35	Altruism

c_i = item complexity; h^2 = communality

continued on next page

Table 5.7 – 8 factor scales and loadings (continued from previous page)

Item	Organization	Analysis	Altruism	Creativity	Leadership/ Erudtion	Production/ Adventure	Leadership/ -Erudition	Adventure	c_i	h^2	Scale
Like to participate in charity events.	0.09	-0.10	0.47	0.06	0.05	0.09	0.16	-0.11	1.69	0.33	Altruism
Would like to be an elementary-school teacher.	0.06	-0.05	0.42	0.08	0.06	0.10	-0.20	0.02	1.77	0.26	Altruism
Would like to help people make career decisions.	0.28	0.03	0.40	0.05	0.16	-0.03	0.12	-0.02	2.42	0.37	Altruism
Like to help others learn new ideas.	-0.01	0.23	0.33	0.12	0.05	0.01	0.30	-0.20	3.91	0.33	Altruism
Would like to be a professional dancer.	0.04	-0.02	0.32	0.28	-0.07	-0.01	-0.08	0.22	3.08	0.25	Creativity
Like to redecorate my house.	0.18	-0.21	0.28	0.25	-0.09	0.08	0.10	-0.10	4.78	0.26	Creativity
Would like to be a minister, priest, rabbi or other religious teacher.	0.11	-0.05	0.16	0.08	0.15	0.11	-0.10	0.01	5.19	0.12	Altruism
Creativity											
Create works of art.	-0.05	-0.02	-0.06	0.80	-0.11	0.08	0.12	-0.09	1.14	0.64	Creativity
Would like to be an artist or architect.	-0.06	0.08	-0.12	0.70	0.00	0.15	-0.06	0.06	1.22	0.57	Creativity
Paint or draw.	-0.03	-0.05	-0.07	0.65	-0.15	0.17	0.11	-0.10	1.42	0.46	Creativity
Would like to select art works for a museum.	0.00	0.01	0.07	0.57	0.06	0.11	-0.17	-0.08	1.35	0.46	Creativity
Like to create new fashion designs.	0.14	-0.10	0.15	0.50	-0.06	-0.03	0.05	0.06	1.53	0.31	Creativity
Write songs.	-0.10	0.09	0.07	0.46	0.05	-0.02	0.01	0.10	1.36	0.27	Creativity
Would like to be an actor or actress.	-0.05	-0.04	0.20	0.46	0.16	-0.18	-0.05	0.38	3.10	0.48	Creativity
Would like to write short stories or novels.	-0.12	0.09	0.07	0.45	0.22	-0.01	-0.20	-0.12	2.44	0.40	Creativity
Would like to act in a play.	-0.06	0.02	0.26	0.44	0.19	-0.13	-0.04	0.21	2.90	0.43	Creativity
Would like to sing professionally.	-0.05	-0.03	0.20	0.44	0.05	-0.11	-0.06	0.27	2.47	0.34	Creativity
Would like to design Internet web pages.	0.24	0.28	-0.15	0.38	-0.01	-0.05	-0.13	0.05	3.33	0.37	Creativity

c_i = item complexity; h^2 = communality

continued on next page

Table 5.7 – 8 factor scales and loadings (continued from previous page)

Item	Organization	Analysis	Altruism	Creativity	Leadership/ Erudtion	Production/ Adventure	Leadership/ -Erudition	Adventure	c_i	h^2	Scale
Would like to plan an advertising campaign.	0.27	-0.05	0.01	0.36	0.31	-0.11	0.00	0.10	3.30	0.41	Leadership
Would like to play an instrument in a symphony.	-0.04	0.23	0.11	0.35	0.01	0.06	-0.12	0.08	2.44	0.27	Creativity
Keep a diary or journal.	0.03	-0.09	0.15	0.26	0.03	0.08	-0.09	-0.18	3.47	0.17	Erudition
Would like to know many languages.	-0.04	0.14	0.10	0.22	0.11	0.05	-0.03	-0.02	3.09	0.14	Erudition
Leadership/Erudition											
Would like to run for political office.	0.00	-0.01	-0.07	-0.06	0.87	0.07	0.01	0.01	1.04	0.74	Leadership
Would like to be a state governor or senator.	0.05	0.03	-0.04	-0.09	0.84	0.05	-0.01	0.04	1.05	0.75	Leadership
Would like to organize a political campaign.	0.04	0.02	-0.01	0.00	0.77	0.03	-0.02	-0.06	1.03	0.60	Leadership
Like to debate topics in a public meeting.	-0.02	0.15	0.02	0.11	0.45	-0.06	0.27	-0.02	2.11	0.36	Leadership
Would like to make decisions that affect a lot of people.	0.10	0.05	0.09	0.01	0.42	-0.05	0.37	0.00	2.27	0.42	Leadership
Would like to be a foreign correspondent.	0.02	0.09	0.13	0.17	0.40	0.11	-0.13	0.07	2.30	0.36	Erudition
Persuade others to change their views.	-0.03	0.14	0.00	0.07	0.35	-0.06	0.28	-0.03	2.50	0.25	Leadership
Would like to be the master of ceremonies at a meeting.	0.18	0.01	0.20	0.18	0.31	-0.10	0.18	0.10	4.60	0.37	Leadership
Speak fluently on any subject.	-0.01	0.22	0.04	0.13	0.24	-0.04	0.21	-0.05	3.76	0.21	Erudition
Production/Adventure											
Would like to be a farmer.	0.05	-0.03	0.05	0.02	0.07	0.66	-0.15	-0.02	1.16	0.48	Production
Would like to be a forest ranger.	-0.09	0.08	-0.01	0.02	0.12	0.64	-0.17	0.09	1.34	0.52	Production
Like to survive in the wilderness.	-0.07	0.09	-0.01	0.03	0.08	0.58	0.18	0.13	1.43	0.46	Adventure
Like woodworking.	0.07	0.15	-0.12	0.19	-0.03	0.53	0.10	0.06	1.70	0.45	Production

c_i = item complexity; h^2 = communality

continued on next page

Table 5.7 – 8 factor scales and loadings (continued from previous page)

Item	Organization	Analysis	Altruism	Creativity	Leadership/ Erudtion	Production/ Adventure	Leadership/ -Erudition	Adventure	c_i	h^2	Scale
Cultivate plants.	0.10	-0.03	0.10	0.09	0.05	0.51	-0.02	-0.22	1.66	0.35	Production
Would like to care for cattle or horses.	0.02	-0.07	0.18	0.03	0.02	0.51	-0.11	0.02	1.41	0.32	Production
Go on nature walks.	-0.03	-0.02	0.15	0.12	0.02	0.49	0.05	-0.16	1.61	0.32	Production
Like to grow flowers.	0.10	-0.06	0.24	0.15	-0.02	0.44	-0.07	-0.26	2.80	0.38	Production
Like to work with tools and machinery.	0.12	0.25	-0.19	0.09	-0.10	0.44	0.18	0.17	3.36	0.48	Production
Like to repair cars or trucks.	0.12	0.17	-0.10	-0.01	-0.03	0.41	0.12	0.26	2.76	0.40	Production
Like to face physical danger.	-0.09	0.05	-0.04	0.01	0.08	0.35	0.25	0.33	3.14	0.35	Adventure
Would like to be a long-distance bicycle rider.	0.05	0.12	0.05	0.07	0.06	0.34	-0.05	0.16	2.03	0.24	Adventure
Would like to be a military officer.	0.08	0.02	-0.05	-0.15	0.25	0.29	0.06	0.29	3.78	0.33	Adventure
Would like to construct new buildings.	0.19	0.23	-0.17	0.20	0.08	0.29	0.06	0.16	5.36	0.42	Production
Leadership/-Erudition											
Like to lead other people.	0.12	-0.02	0.19	0.05	0.30	-0.05	0.53	0.03	2.06	0.51	Leadership
Would like to be a librarian.	0.05	0.18	0.12	0.18	0.07	0.16	-0.51	-0.19	2.27	0.48	Erudition
Like to make important things happen.	0.08	0.08	0.20	0.08	0.15	-0.04	0.46	-0.05	1.89	0.35	Leadership
Engage in exciting adventures.	-0.01	-0.05	0.18	0.13	0.01	0.22	0.43	0.14	2.45	0.33	Adventure
Would like to be a professor of English.	-0.06	0.08	0.21	0.30	0.28	-0.01	-0.36	-0.08	3.92	0.44	Erudition
Would like to edit a newspaper.	0.08	0.10	0.12	0.29	0.31	0.00	-0.33	-0.10	3.80	0.42	Erudition
Would like to be a translator or interpreter.	0.06	0.14	0.22	0.20	0.10	0.09	-0.23	0.04	4.74	0.26	Erudition
Adventure											
Would like to be a professional athlete.	0.06	0.03	0.07	-0.03	0.04	0.15	0.10	0.50	1.38	0.35	Adventure

c_i = item complexity; h^2 = communality

continued on next page

Table 5.7 – 8 factor scales and loadings (continued from previous page)

Item	Organization	Analysis	Altruism	Creativity	Leadership/ Erudtion	Production/ Adventure	Leadership/ -Erudition	Adventure	c_i	h^2	Scale
Would like to be a racing car driver.	0.11	0.05	-0.04	0.04	0.06	0.27	0.07	0.43	2.01	0.37	Adventure
Like to compete in athletic events.	0.07	-0.02	0.12	-0.07	-0.02	0.19	0.32	0.39	2.81	0.35	Adventure
Would like to be a bounty hunter.	0.02	0.05	-0.07	0.02	0.16	0.23	0.04	0.37	2.28	0.30	Adventure
Read many books.	-0.09	0.15	0.10	0.16	0.16	0.07	-0.11	-0.33	3.40	0.23	Erudition
Would like to be a police officer.	0.04	-0.03	0.09	-0.11	0.17	0.23	-0.03	0.33	2.91	0.24	Adventure

c_i = item complexity; h^2 = communality

continued on next page

5.3.3 Discussion

The fit statistics for factor analysis of the ORVIS items did not converge on a common solution; the MAP and eBIC both suggested the presence of more factors than the 8 scales in the measure. Mean item complexity, on the other hand, pointed towards six factors as a superior solution to eight. Examination of the item content for the six factors confirmed the hypothesis that these items would generally correspond to the six factor RIASEC model. The item content for the eight factor solution, however, did not closely match those which were prescribed by the eight ORVIS scales, and this was mainly due to primary loading of several Leadership, Erudition, and Adventure items onto alternate scales. While the possibility of an eight-factor solution using different items or in a different sample may still exist, these analyses provide strong evidence that the Holland types structure provides a superior fit, even among items which were designed to describe an eight-factor structure.

5.4 Study 3

In Study 1 and Study 2, it was concluded that the 6-factor Holland types structure fit the 60-item O*NET® Interest Profiler Short Form, as well as the 92-item ORVIS subset. Additionally, an empirical six-factor solution of the O*NET® item pool had less item-structure disagreement, as well as lower average item complexity than the empirical six-factor ORVIS solution (1.80 vs. 1.90). Based on these conclusions, the O*NET® empirical six-factor solution (which differed from the theoretical six-factor solution by three items) was used as a starting point in an attempt to identify an improved Holland types item pool. This was done by extending the factors of the O*NET® empirical six-factor solution onto the 92 ORVIS items. It was hypothesized that some of the ORVIS items would have higher factor loadings on this structure than the O*NET® items, and to the extent that this was the case, new scales could be developed with lower mean item complexity.

5.4.1 Methods

The participants in Study 3 were the same as those in Study 1 (and overlapped with the sample used in Study 2); the same 14,882 participants who answered subsets of the O*NET items were also administered subsets of the ORVIS items. The pairwise administration count between all of the O*NET and ORVIS items had a mean of 701 (median = 634, minimum = 405).

The factor extension procedures used are based on those available in the *psych* package in R (Revelle, 2014) and are essentially based on the extension of exploratory factor analysis to new variables which were not part of the original set of items factored. More specifically, the six factor solution based on the 60 items of the O*NET scales was extended onto the 92 items of the ORVIS using the underlying correlations among all items. The top 10 items for each of the six factors were then identified; items from the O*NET scales which were missing from the top 10 items in the factor extension output were dropped and items which were not in the original O*NET scales but which were present among the top 10 items were added. In order to compare the revised scales to the factor analytic output from Studies 1 and 2, the last step was to extract 6 factors from the items identified as the new revised set by themselves (rather than with all 152 items together).

5.4.2 Results

Table 5.10 shows the items and factor loadings for the top sixty items based on factor extension of the empirical six factor solution for the O*NET items onto the full set of O*NET and ORVIS items together. These sixty items, representing the top 10 items for each of the six extended factors, are referred to below as the SAPA Personality Inventory - Vocational Interest ("SPI-VI") scales. Comparison of the psychometric properties for the SPI-VI scales and the empirical O*NET scales described in Study 1 is provided in Table 5.8. The properties are improved for all

Table 5.8: Comparison of scale properties: O*NET 6 Factor and SPI-VI 6 Factor

	Realistic	Investigative	Artistic	Social	Enterprising	Conventional
(Standardized) Alpha:						
O*NET	0.86	0.88	0.84	0.80	0.82	0.85
SPI-VI	0.90	0.93	0.87	0.87	0.86	0.87
Average item correlation:						
O*NET	0.38	0.42	0.35	0.29	0.31	0.36
SPI-VI	0.48	0.56	0.40	0.41	0.38	0.41
Signal to Noise ratio based upon average r and n:						
O*NET	6.12	7.12	5.43	4.10	4.48	5.54
SPI-VI	9.24	12.55	6.67	6.85	6.20	6.93
Number of items:						
O*NET	10	10	10	10	10	10
SPI-VI	10 (4 new)	10 (4 new)	10 (2 new)	10 (4 new)	10 (3 new)	10 (2 new)

six scales and the mean item complexity of the six factor solution was improved considerably, from 1.80 for the six-factor O*NET solution to 1.25 for the six-factor SPI-VI solution.

Table 5.9 shows the scale intercorrelations for the O*NET and SPI-VI scales after correcting for item overlap (below the diagonal) and after correcting for item overlap and attenuation (above the diagonal). The correlations between constructs across the scales is uniformly high, ranging from 0.95 to 0.99 after correcting for attenuation. While the correlations between the sets of scales was very high, the SPI-VI scales contained 20 ORVIS items (and 20 fewer O*NET items), as detailed in Table 5.8.

Table 5.9: O*NET and SPI-VI scale intercorrelations corrected for item overlap and attenuation

	O*NET						SPI-VI					
	Realistic	Investig.	Artistic	Social	Enterpris.	Convent.	Realistic	Investig.	Artistic	Social	Enterpris.	Convent.
O*NET Realistic	*0.86*	0.62	0.41	0.21	0.35	0.63	0.98	0.49	0.32	0.05	0.36	0.55
O*NET Investigative	0.54	*0.88*	0.38	0.24	0.23	0.47	0.58	0.97	0.32	0.16	0.27	0.38
O*NET Artistic	0.35	0.33	*0.84*	0.35	0.32	0.21	0.37	0.33	0.99	0.30	0.25	0.18
O*NET Social	0.17	0.20	0.29	*0.80*	0.42	0.20	0.13	0.17	0.30	0.96	0.29	0.24
O*NET Enterprising	0.29	0.20	0.27	0.34	*0.82*	0.45	0.30	0.14	0.26	0.29	0.95	0.47
O*NET Conventional	0.54	0.40	0.18	0.17	0.38	*0.85*	0.52	0.34	0.12	0.09	0.48	0.99
SPI-VI Realistic	0.86	0.51	0.33	0.11	0.25	0.45	*0.90*	0.47	0.32	-0.02	0.34	0.43
SPI-VI Investigative	0.44	0.87	0.29	0.15	0.12	0.31	0.43	*0.93*	0.28	0.11	0.19	0.27
SPI-VI Artistic	0.28	0.28	0.85	0.25	0.22	0.10	0.28	0.25	*0.87*	0.27	0.18	0.11
SPI-VI Social	0.04	0.14	0.26	0.8	0.25	0.08	-0.02	0.10	0.24	*0.87*	0.13	0.14
SPI-VI Enterprising	0.31	0.23	0.21	0.24	0.79	0.41	0.30	0.17	0.16	0.11	*0.86*	0.48
SPI-VI Conventional	0.48	0.34	0.16	0.20	0.40	0.85	0.38	0.24	0.10	0.12	0.41	*0.87*

All correlations have been corrected for item overlap. The alpha for each scale is reported on the diagonal and the alpha corrected correlations are reported above the diagonal.

Table 5.10: SPI-VI 6 factor scales and loadings

Item	Investigative	Realistic	Social	Artistic	Conventional	Enterprising	c_i	h^2	New item
Investigative									
Would like to do laboratory tests to identify diseases.	0.88	-0.08	0.08	-0.09	0.07	-0.05	1.08	0.74	N
Would like to work in a biology lab.	0.80	0.02	-0.01	0.00	0.00	-0.07	1.01	0.64	N
Would like to carry out medical research.	0.79	-0.07	0.15	-0.09	-0.04	0.05	1.13	0.60	Y
Would like to examine blood samples using a microscope.	0.77	-0.01	0.02	0.02	0.11	-0.07	1.06	0.62	N
Would like to conduct chemical experiments.	0.77	0.15	-0.10	0.02	-0.03	-0.03	1.12	0.69	N
Would like to develop a new medicine.	0.74	-0.02	0.11	0.01	-0.06	0.16	1.16	0.59	N
Would like to be a chemist.	0.71	0.13	-0.10	0.05	-0.01	-0.01	1.12	0.60	Y
Would like to design a laboratory experiment.	0.70	0.09	-0.09	0.12	-0.02	0.07	1.15	0.61	Y
Would like to be a scientific reporter.	0.63	0.05	-0.03	0.18	0.00	0.03	1.17	0.51	Y
Would like to study the movement of planets.	0.44	0.20	-0.14	0.22	-0.02	-0.03	2.19	0.40	N
Realistic									
Would like to repair household appliances.	-0.01	0.77	0.09	-0.06	0.09	-0.06	1.08	0.60	N
Like to work with tools and machinery.	0.05	0.77	-0.04	-0.01	-0.11	0.04	1.06	0.60	Y
Would like to build kitchen cabinets.	-0.01	0.70	0.12	0.08	0.06	0.01	1.10	0.55	N
Like to repair cars or trucks.	0.05	0.69	0.04	-0.10	-0.14	0.08	1.17	0.47	Y
Would like to assemble electronic parts.	0.08	0.65	-0.15	0.06	0.04	0.06	1.19	0.56	N
Like woodworking.	0.00	0.65	0.04	0.12	-0.07	0.00	1.11	0.44	Y
Would like to set up and operate machines to make products.	0.11	0.64	-0.11	-0.04	0.12	0.10	1.26	0.60	N
Would like to repair and install locks.	0.04	0.63	0.01	0.00	0.22	-0.05	1.27	0.53	N
Would like to lay brick or tile.	-0.05	0.56	0.07	0.06	0.17	-0.07	1.29	0.38	N
Would like to construct new buildings.	0.08	0.50	-0.02	0.14	0.01	0.23	1.65	0.45	Y
Social									
Would like to help conduct a group therapy session.	0.04	0.02	0.75	0.04	-0.03	0.08	1.04	0.59	N
Would like to be a counselor or therapist.	0.06	-0.07	0.75	0.10	0.01	0.00	1.06	0.63	Y
Would like to help people with personal or emotional problems.	-0.07	0.05	0.74	0.01	-0.05	-0.03	1.05	0.54	N
Like to counsel persons who need help.	-0.05	0.03	0.73	0.03	-0.09	0.09	1.08	0.55	Y
Would like to be a social worker.	-0.02	0.02	0.71	-0.02	0.02	-0.02	1.01	0.50	Y
Would like to instruct parents on child care.	0.04	0.06	0.61	-0.06	0.03	0.04	1.07	0.37	Y
Would like to perform rehabilitation therapy.	0.13	0.02	0.61	-0.04	0.10	-0.05	1.18	0.41	N
Would like to take care of children at a day-care center.	-0.02	0.02	0.48	-0.04	0.09	-0.07	1.14	0.23	N

c_i = item complexity; h^2 = communality

continued on next page

Table 5.10 – SPI-VI 6 factor scales and loadings (continued from previous page)

Item	Investigative	Realistic	Social	Artistic	Conventional	Enterprising	c_i	h^2	New item
Would like to teach sign language to people with hearing disabilities.	0.11	0.02	0.46	0.10	0.15	-0.09	1.52	0.30	N
Would like to do volunteer work at a non-profit organization.	0.08	-0.05	0.43	0.04	0.12	-0.11	1.43	0.23	N
Artistic									
Would like to write short stories or novels.	0.00	-0.13	0.02	0.76	0.05	-0.04	1.07	0.57	Y
Would like to write scripts for movies or television shows.	0.02	-0.05	0.01	0.74	-0.03	0.17	1.13	0.58	N
Would like to write books or plays.	0.06	-0.11	0.05	0.72	-0.02	0.00	1.07	0.53	N
Would like to be an artist or architect.	-0.03	0.17	0.01	0.66	0.01	-0.05	1.15	0.50	Y
Create works of art.	-0.06	0.05	0.07	0.60	0.01	-0.05	1.08	0.39	Y
Would like to draw pictures.	-0.01	0.10	0.03	0.60	0.03	-0.14	1.18	0.41	N
Would like to edit movies.	0.02	0.11	-0.05	0.60	0.08	0.10	1.18	0.44	N
Would like to compose or arrange music.	0.08	0.12	-0.04	0.56	-0.07	0.03	1.18	0.38	N
Would like to play a musical instrument.	0.07	0.07	0.05	0.45	0.01	-0.07	1.18	0.25	N
Would like to sing in a band.	-0.02	0.03	0.12	0.43	-0.02	0.01	1.19	0.22	N
Conventional									
Would like to keep inventory records.	-0.04	0.08	0.00	0.03	0.75	-0.01	1.03	0.58	N
Would like to keep track of a companys inventory.	0.01	-0.03	-0.04	-0.04	0.71	0.18	1.15	0.60	Y
Would like to keep shipping and receiving records.	0.01	0.05	0.00	0.04	0.64	0.04	1.03	0.45	N
Would like to stamp, sort, and distribute mail for an organization.	-0.04	0.10	0.06	0.04	0.63	-0.22	1.34	0.41	N
Would like to record rent payments.	0.04	-0.05	0.03	-0.10	0.63	0.07	1.10	0.42	N
Would like to inventory supplies using a hand-held computer.	0.05	0.09	-0.04	0.00	0.62	0.01	1.07	0.44	N
Would like to develop an office filing system.	0.01	0.03	0.08	0.06	0.60	0.06	1.08	0.43	Y
Would like to proofread records or forms.	0.09	-0.14	0.05	0.18	0.54	-0.01	1.45	0.33	N
Would like to calculate the wages of employees.	0.06	0.01	0.02	-0.12	0.52	0.32	1.82	0.50	N
Would like to develop a spreadsheet using computer software.	0.11	0.13	-0.18	0.06	0.44	0.15	2.01	0.38	N
Enterprising									
Would like to be the chief executive of a large company.	0.04	0.01	-0.05	0.01	-0.08	0.81	1.03	0.65	Y
Would like to negotiate business contracts.	-0.02	0.00	0.02	0.03	0.09	0.74	1.04	0.60	N
Would like to be a sales or marketing director.	-0.10	-0.01	0.07	0.07	0.05	0.65	1.11	0.45	Y

c_i = item complexity; h^2 = communality

continued on next page

Table 5.10 – SPI-VI 6 factor scales and loadings (continued from previous page)

Item	Investigative	Realistic	Social	Artistic	Conventional	Enterprising	c_i	h^2	New item
Would like to be the financial officer for a company.	0.03	0.04	-0.08	-0.08	0.24	0.63	1.38	0.57	Y
Would like to manage a department within a large company.	-0.02	0.03	0.09	-0.08	0.12	0.63	1.16	0.47	N
Would like to buy and sell stocks and bonds.	0.07	0.19	-0.13	-0.04	0.07	0.56	1.43	0.46	N
Would like to start your own business.	-0.06	0.12	0.07	0.09	-0.08	0.54	1.27	0.33	N
Would like to represent a client in a lawsuit.	0.16	-0.07	0.13	0.07	-0.05	0.51	1.43	0.31	N
Would like to manage a retail store.	-0.07	0.06	0.16	0.05	0.22	0.39	2.16	0.30	N
Would like to market a new line of clothing.	-0.07	-0.09	0.15	0.25	0.04	0.38	2.35	0.25	N

c_i = item complexity; h^2 = communality

5.4.3 Discussion

After factor-extending the O*NET®structure onto the ORVIS items, 20 ORVIS items had higher loadings on the six RIASEC factors than the 60-item O*NET® Interest Profiler Short Form. Replacing these 20 items reduced item complexity and increased the internal consistency of the scales, and the resulting 60 items are available for use in the public-domain as the SPI-VI 6 factor scales.

5.5 Conclusion

It was determined in Study 1 that the six-factor solution of the O*NET Interest Profiler corresponded well with the six factors of the RIASEC Holland types. By contrast, the evidence for the proposed eight-factor structure of the Oregon Vocational Interest Scales was not supported in Study 2, though there was again support for six-factors similar to the Holland types in the ORVIS items. On the basis of these first two studies, new scales were developed on the basis of the six-factor structure in Study 3. The primary benefit of these six new scales is reduced complexity though this is supplemented by the fact that the new scales are also available for unrestricted use in the public domain.

Chapter 6

Integration and conclusion

In the first chapter, it was ascertained that individual differences research as a whole suffered from the absence of a scientific paradigm, despite the fact that interpersonal recognition of the *importance* of differences among individuals has existed for millennia. Many of the potential reasons for this absence have been addressed, some of them theoretical and others more practical.

Among the theoretical has been the need to explicitly circumscribe the scope of "differential" psychology to include only those differences which are amenable to scientific inquiry by virtue of being quantifiable and generalizable. This circumscription, in turn, precipitates the recognition that differential psychology need not (and perhaps *should* not) be viewed as synonymous with the vague, overly-broad, and occasionally phenomenological field of "personality" psychology. Certainly, the two fields are overlapping but it seems that considerable benefit to the prospects for a scientific paradigm would be achieved by this distinction.

The circumscribed scope of differential psychology nevertheless remains broad and this breadth underlies some of the practical reasons which contribute to the absence of an existing paradigm. The field should, at least, take into account the findings reported from the affective, cognitive

and conative domains of research over the last 125 years, yet the development of an assessment model which integrates this research has proven formidable. Only in recent years has this task become feasible, thanks in large part to the development of public-domain item sets and telemetric assessment methods.

The challenges unique to the development of a public-domain measure for each of the cognitive, conative, and affective disciplines has been described here in a separate section (mainly Chapters 3 through 5), but the ultimate goal has been to advocate for the concomitant administration of these measures. Assuming administration of the 150 SPI3/5/15 items, the 60 SPI-VI items, and a minimum of 16 ICAR items, the total number of items to be administered (226) is below that used for several traditional "personality" measures. The NEO-PI-R, for example, contains 240 items; the HEXACO contains 300. It remains possible to administer the SPI3/5/15, ICAR, and SPI-VI measures independently of course but joint administration offers the best hope for iterative testing, refinement, and advancement of differential psychology research.

6.1 Future research and limitations

Above all, joint administration of the measures described herein will allow for integration across domains and specification of an empirically-testable paradigm for individual differences. Relatively little research has previously been reported about the ways in which the affective, cognitive and conative domains offer overlapping and/or unique predictive validity; the measures described here offer considerable utility for the exploration of these topics, especially to the extent that they can be included in longitudinal research. Because the samples described in previous chapters were overlapping to some extent, it would have been feasible to begin evaluating these topics on the basis of the data reported herein. The results of these analyses were omitted from this project however, largely on the grounds that it was inappropriate to

report them prior to administration of the proposed measures in independent samples. This remains the top priority for future work and data collection is already underway.

Of course, there are many contexts in which administration of the full 226 item set will not be an option due to time constraints or concerns about participant fatigue. Assessment using SAPA-like sampling techniques are one solution to this problem; the use of psychometric test reduction procedures and computer-adaptive testing are another. There are several opportunities to improve upon the administration of the measures proposed here through the use of item response theory-based psychometric techniques which serve to reduce the burden on participants without compromising the reliability of their scores.

The possibility of more efficient administration might also be desirable for the sake of being able to expand the number of domains being assessed. While the affective, cognitive, and conative domains have long research traditions, several other domains of individual differences are also likely to offer incremental predictive utility, including avocational interests (pasttimes and hobbies), values (especially those related to work and interpersonal relations), biophysiological characteristics, and health behaviors. All of these areas offer at least one public-domain measure which might be used as a basis for integration into the assessment model described here.

It should also be noted that considerable research is needed to confirm the merits of these scales through replication. While the item sets used here were generally large relative to those which have been administered historically (especially in the affective domain), they were still limited relative to the very large universe of items available. For example, the samples described in Chapters 2 and 3 used fewer than one-sixth of the affective items available in the public-domain. In the cognitive ability domain, additional research is needed to develop more item types and ensure that the items recommended for administration conform to modern theory regarding the structure of cognitive abilities generally.

It should be similarly noted that, while the samples used here were generally larger and more internationally representative than those which have been collected historically, they are by no means "representative." Differential psychologists should be striving for the largest and most representative samples possible by virtue of the nature of individual differences. This includes samples with a much greater proportion of participants from outside the United States, and should particularly include greater representation from non-English-speaking countries. This implies translation of the items administered, a topic which represents a massive undertaking for future research.

6.2 Conclusion

The time is ripe for individual differences research to move forward as a paradigmatic science. In some sense, this process has already begun with the acceptance of the Big Five as the consensual theory of temperament in the late 20th century, but this addresses only a fraction of the individual differences which offer predictive utility. If the most difficult aspect of developing a broader paradigm is the creation of an empirically-informed assessment model which can be widely re-administered, tested, and improved, then the scales put forth in these chapters may serve as a starting point from which future research will proceed. Perhaps in time, with refinement of the potential redundancies and oversights of these scales, the field of differential psychology can proceed with an integrated paradigm that is predictive of a wide-range of real-world outcomes.

References

Ackerman, P. L. (1997). Personality, self-concept, interests, and intelligence: Which construct doesn't fit? *Journal of Personality*, 65(2):171–204.

Ackerman, P. L. and Beier, M. E. (2003). Intelligence, personality, and interests in the career choice process. *Journal of Career Assessment*, 11(2):205–218.

Ackerman, P. L. and Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121(2):219–245.

Alexander, R. A. (1990). Correction formulas for correlations restricted by selection on an unmeasured variable. *Journal of Educational Measurement*, 27(2):187–189.

Allport, G. W. (1921). Personality and character. *Psychological Bulletin*, 18(9):441–455.

Allport, G. W. (1940). The psychologist's frame of reference. *Psychological Bulletin*, 37(1):1–26.

Allport, G. W. and Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs: General and Applied*, 47(1):1–170.

Allport, G. W. and Vernon, P. E. (1930). The field of personality. *Psychological Bulletin*, 27(10):677–730.

Anastasi, A. (1948). The nature of psychological "traits". *Psychological Review*, 55(3):127–138.

Anderson, J. O., Lin, H.-S., Treagust, D. F., Ross, S. P., and Yore, L. D. (2007). Using large-scale assessment datasets for research in science and mathematics education: Programme for International Student Assessment (PISA). *International Journal of Science and Mathematics Education*, 5(4):591–614.

Arendasy, M., Sommer, M., Gittler, G., and Hergovich, A. (2006). Automatic generation of quantitative reasoning items. *Journal of Individual Differences*, 27(1):2–14.

Aristotle (2000). *Nicomachean Ethics*. Crisp, R. (Ed.). Cambridge Univ Press., Cambridge, UK. (Original work published c. 350 BCE).

Armstrong, P., Smith, T., Donnay, D., and Rounds, J. (2004). The Strong Ring: A Basic Interest Model of Occupational Structure. *Journal of Counseling Psychology*, 51(3):299–313.

Armstrong, P. I., Allison, W., and Rounds, J. (2008a). Development and initial validation of brief public domain RIASEC marker scales. *Journal of Vocational Behavior*, 73(2):287–299.

Armstrong, P. I. and Anthoney, S. F. (2009). Personality facets and RIASEC interests: An integrated model. *Journal of Vocational Behavior*, 75(3):346–359.

Armstrong, P. I., Day, S. X., McVay, J. P., and Rounds, J. (2008b). Holland's RIASEC model as an integrative framework for individual differences. *Journal of Counseling Psychology*, 55(1):1–18.

Ashton, M. C. and Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Bulletin*, 11(2):150–166.

Ashton, M. C., Lee, K., and Goldberg, L. R. (2007). The IPIP–HEXACO scales: An alternative, public-domain measure of the personality constructs in the HEXACO model. *Personality and Individual Differences*, 42(8):1515–1526.

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., Boies, K., and De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, 86(2):356–366.

Atkinson, J. and Birch, D. (1970). *The dynamics of action.* John Wiley, Oxford.

Baker, F. B. (1985). *The basics of item response theory.* Heinemann Educational Books, Portsmouth, NH.

Barrick, M. R. and Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44(1):1–26.

Barrick, M. R., Mount, M. K., and Gupta, R. (2003). Meta-analysis of the relationship between the five-factor model of personality and Holland's occupational types. *Personnel Psychology*, 56(1):45–74.

Bashaw, W. L. and Anderson, Jr, H. E. (1967). A correction for replicated error in correlation coefficients. *Psychometrika*, 32(4):435–441.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117(2):187–215.

Boegehold, A. L. (1959). The Date of Theophrastus' Characters. *Transactions and Proceedings of the American Philological Association*, 90:15–19.

Boring, E. G. (1923). Intelligence as the tests test it. *The New Republic*, 36:35–37.

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., and Tellegen, A. (1990). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science*, 250(4978):223–228.

Brett, G. S. (1921). *A history of psychology*. The Macmillan Company, New York.

Briggs, S. R. (1992). Assessing the five-factor model of personality description. *Journal of Personality*, 60(2):253–293.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36(1):111–150.

Bryant, N. D. and Gokhale, S. (1972). Correcting correlations for restrictions in range due to selection on an unmeasured variable. *Educational and Psychological Measurement*, 32(2):305–310.

Bureau, U. S. C. (2012). Educational Attainment in the United States: 2009. *2009 American Community Survey 5-Year Estimates*.

Burnham, W. H. (1929). Personality differences and mental health. *The Pedagogical Seminary and Journal of Genetic Psychology*, 36(3):361–389.

Buss, D. M. (1984). Evolutionary biology and personality psychology: Toward a conception of human nature and individual differences. *American Psychologist*, 39(10):1135.

Buss, D. M. (1991). Evolutionary personality psychology. *Annual Review of Psychology*, 42(1):459–491.

Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., and Dahlstrom, W. G. (2003). *MMPI-2: Minnesota Multiphasic Personality Inventory-2*. University of Minnesota Press.

Camara, W. J., Nathan, J. S., and Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice*, 31(2):141–154.

Canli, T. (2008). Toward a "molecular psychology" of personality. In John, O. P., Robins, R. W., and Pervin, L. A., editors, *Handbook of personality: Theory and research*, pages 1–17. Guilford Press, New York.

Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., and Gabrieli, J. D. E. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience*, 115(1):33–42.

Carroll, J. B. (1953). An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, 18(1):23–38.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press, Cambridge, UK.

Carver, C. S. and Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological Bulletin*, 92(1):111.

Carver, C. S. and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67(2):319–333.

Cassady, J. C. (2001). Self-reported GPA and SAT: A methodological note. *Practical Assessment, Research & Evaluation*, 7(12).

Cattell, R. B. (1940). Sentiment or attitude? The core of a terminology problem in personality research. *Journal of Personality*, 9(1):6–17.

Cattell, R. B. (1943a). The measurement of adult intelligence. *Psychological Bulletin*, 40(3):153–193.

Cattell, R. B. (1943b). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, 38(4):476.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, 12(3):197–220.

Cattell, R. B. (1952). *Factor analysis: An introduction and manual for the psychologist and social scientist*. Harper, New York.

Cattell, R. B. (1966). *Handbook of multivariate experimental psychology*. Rand McNally & Company, Chicago.

Cattell, R. B., Eber, H. W., and Tatsuoka, M. M. (1970). *Handbook for the sixteen personality factor questionnaire (16 PF)*. Institute for Personality and Ability Testing, Champaign, IL.

Cattell, R. B., Young, H. B., and Hundleby, J. D. (1964). Blood groups and personality traits. *American Journal of Human Genetics*, 16(4):397.

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., Ader, D., Fries, J. F., Bruce, B., and Rose, M. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care*, 45(Suppl 1):S3–S11.

Census, U. S. (2011). American Community Survey 5-Year Estimates. *2011 American Community Survey 5-Year Estimates*, pages 1–1.

Chamorro-Premuzic, T., Stumm, S. V., and Furnham, A., editors (2011). *The Wiley-Blackwell handbook of individual differences*, volume 3. Wiley-Blackwell, London.

Clark, L. A. and Watson, D. (2008). Temperament: An organizing paradigm for trait psychology. In John, O. P., Robins, R. W., and Pervin, L. A., editors, *Handbook of personality: Theory and research*, pages 265–286. Guilford Press, New York.

Clark, S. J. and Houle, B. (2012). Evaluation of Heckman selection model method for correcting estimates of HIV prevalence from sample surveys. *Center for Statistics and the Social Sciences*, Working Paper no. 120:1–18.

Cole, J. S. and Gonyea, R. M. (2009). Accuracy of self-reported SAT and ACT test scores: Implications for research. *Research in Higher Education*, 51(4):305–319.

College Board (2012). 2012 college-bound seniors total group profile report.

Comrey, A. L. (1961). Factored homogeneous item dimensions in personality research. *Educational and Psychological Measurement*, 21:417–431.

Comrey, A. L. (1973). *A first course in factor analysis*. Academic Press Inc, New York.

Comrey, A. L. (1984). Comparison of two methods to identify major personality factors. *Applied Psychological Measurement*, 8(4):397–408.

Condit, C. M. (1999). How the public understands genetics: non-deterministic and non-discriminatory interpretations of the "blueprint" metaphor. *Public Understanding of Science*, 8(3):169–180.

Condit, C. M., Ferguson, A., Kassel, R., Thadhani, C., Gooding, H. C., and Parrott, R. (2001). An exploratory study of the impact of news headlines on genetic determinism. *Science Communication*, 22(4):379–395.

Condon, D. M., Elleman, L., and Revelle, W. (2013). *Correlations between the IPIP100 and ICAR scales in a large online sample*. Retrieved December 3, 2013, from: http://www.personality-project.org/revelle/publications/IPIP100ICAR.pdf.

Condon, D. M. and Revelle, W. (2014). The International Cognitive Ability Resource: Development and initial validation of a public-domain measure. *Intelligence*, 43:52–64.

Confucius (1994). *Analects*. The Internet Classics Archive, http://classics.mit.edu//Confucius/analects.html. (Original work published c. 435 BCE).

Corr, P. J. and McNaughton, N. (2008). The Reinforcement Sensitivity Theory of personality. In Corr, P. J., editor, *The Reinforcement Sensitivity Theory of personality*, pages 155–187. Cambridge Univ Press, Cambridge.

Costa, P. T. and McCrae, R. R. (1976). Age differences in personality structure: A cluster analytic approach. *Journal of Gerontology*, 31(5):564–570.

Costa, P. T. and McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual*. Psychological Assessment Resources, Odessa, FL.

Costa, P. T., McCrae, R. R., and Holland, J. L. (1984). Personality and vocational interests in an adult sample. *Journal of Applied Psychology*, 69(3):390–400.

Courtiol, A., Pettay, J. E., Jokela, M., Rotkirch, A., and Lummaa, V. (2012). Natural and sexual selection in a monogamous historical human population. *Proceedings of the National Academy of Sciences*, 109(21):8044–8049.

Courtiol, A., Rickard, I. J., Lummaa, V., Prentice, A. M., Fulford, A. J. C., and Stearns, S. C. (2013). The demographic transition influences variance in fitness and selection on height and BMI in rural Gambia. *Current Biology*, 23(10):884–889.

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., and Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, 26(4):414–431.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12(11):671–684.

Cronback, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30:116–127.

Cuddeback, G., Wilson, E., Orme, J. G., and Combs-Orme, T. (2004). Detecting and statistically correcting sample selection bias. *Journal of Social Service Research*, 30(3):19–33.

Dahlstrom, W. G., Hopkins, D., Dahlstrom, L., Jackson, E., and Cumella, E. (1996). MMPI findings on astrological and other folklore concepts of personality. *Psychological Reports*, 78(3c):1059–1070.

Dallenbach, K. M. (1955). Phrenology versus psychoanalysis. *The American Journal of Psychology*, 68(4):511–525.

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. John Murray, London, 1st edition.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. D. Appleton and Company, New York.

Darwin, C. (1886). *The expression of the emotions in man and animals*. D. Appleton and Company, New York.

Davies, M. (2008). *T*he Corpus of Contemporary American English: 450 million words, 1990-present. Retrieved February 9, 2013, from http://corpus.byu.edu/coca/.

De Raad, B., Barelds, D. P. H., Levert, E., Ostendorf, F., Mlačić, B., Blas, L. D., Hrebícková, M., Szirmák, Z., Szarota, P., Perugini, M., Church, A. T., and Katigbak, M. S. (2010). Only three factors of personality description are fully replicable across languages: A comparison of 14 trait taxonomies. *Journal of Personality and Social Psychology*, 98(1):160–173.

De Raad, B. and Ceulemans, E. (2001). The trait dimensional scope of the characters of Theophrastus. In Riemann, R., Spinath, F. M., and Ostendorf, F., editors, *Personality and temperament: Genetics, evolution, and structure*, pages 168–184. Pabst Science Publishers, Lengerich.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., and Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, 86(1):130–147.

Dennis, I., Handley, S., Bradon, P., Evans, J., and Newstead, S. (2002). Approaches to modeling item-generative tests. In Irvine, S. H. and Kyllonen, P. C., editors, *Item generation for test development*, pages 53–71. Lawrence Erlbaum Associates, Mahwah, New Jersey.

DeYoung, C. (2010a). Personality neuroscience and the biology of traits. *Social and Personality Psychology Compass*, 4(12):1165–1180.

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, 91(6):1138–1151.

DeYoung, C. G. (2010b). Toward a theory of the Big Five. *Psychological Inquiry*, 21(1):26–33.

DeYoung, C. G. (2012). Intelligence and personality. In Sternberg, R. J. and Kaufman, S. B., editors, *The Cambridge handbook of intelligence*, pages 711–737. Cambridge Univ Press, New York.

DeYoung, C. G., Quilty, L. C., and Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93(5):880–896.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41(1):417–440.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 73(6):1246–1256.

Digman, J. M. and Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, 50(1):116.

Digman, J. M. and Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research*, 16(2):149–170.

Donnellan, M. B., Oswald, F. L., Baird, B. M., and Lucas, R. E. (2006). The Mini-IPIP Scales: Tiny-yet-effective measures of the Big Five Factors of Personality. *Psychological Assessment*, 18(2):192–203.

Duckworth, A. and Seligman, M. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16(12):939–944.

Educational Testing Service (2010). Table of GRE scores by intended graduate major field.

Einstein, A. (1941). Science, philosophy and religion, a symposium. Conference on Science, Philosophy and Religion Their Relation to The Democratic Way of Life. New York.

Ekstrom, R. B., French, J. W., Harman, H. H., and Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Educational Testing Service, Princeton, NJ.

Eliot, J. and Smith, I. M. (1983). *An International Directory of Spatial Tests*. NFER-NELSON Publishing Company Ltd., Great Britain.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment*, 8(4):341–349.

Evans, A. M. and Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, 42(6):1589–1597.

Eysenck, H. J. (1981). *A model for personality*. Springer-Verlag, New York.

Eysenck, H. J. (1983a). Cicero and the state-trait theory of anxiety: Another case of delayed recognition. *American Psychologist*, 38(1):114.

Eysenck, H. J. (1983b). Is there a paradigm in personality research? *Journal of Research in Personality*, 17:369–397.

Eysenck, H. J. (1985). The place of theory in a world of facts. In Madsen, K. B. and Mos, L. P., editors, *Annals of Theoretical Psychology*, pages 17–72. Plenum Press, New York.

Eysenck, H. J. (1991). Dimensions of personality: 16, 5 or 3?—Criteria for a taxonomic paradigm. *Personality and Individual Differences*, 12(8):773–790.

Eysenck, H. J. (1994). The Big Five or Giant Three: Criteria for a paradigm. In Halverson, C. F., Kohnstamm, G. A., and Martin, R. P., editors, *The developing structure of temperament and personality from infancy to adulthood*, pages 37–51. Lawrence Erlbaum, Hillsdale, NJ.

Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology*, 73(6):1224.

Eysenck, H. J. and Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. Plenum Press, New York.

Eysenck, S., Eysenck, H. J., and Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, 6(1):21–29.

Fechner, G. T. (1860). *Elements of psychophysics*, volume 1. Howes, D. H. (Ed.) and Boring, E. G. (Ed.). Holt, Rinehart and Winston, Chicago.

Feist, G. J. (2012). Predicting interest in and attitudes toward science from personality and need for cognition. *Personality and Individual Differences*, 52(7):771–775.

Fernald, G. G. (1920). Character vs. intelligence in personality studies. *The Journal of Abnormal Psychology*, 15(1):1.

Field, T. G. (2012). Standardized tests: Recouping development costs and preserving integrity.

Fraley, C. R. (2004). *How to conduct behavioral research over the Internet: A beginner's guide to HTML and CGI/Perl*. Guilford Press, New York.

Frey, M. C. and Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, 15(6):373–378.

Frucot, V. and Cook, G. (1994). Further research on the accuracy of students' self-reported grade point averages, SAT scores, and course grades. *Perceptual and Motor Skills*, 79(2):743–746.

Fua, K., Horswill, I., Ortony, A., and Revelle, W. (2009). Reinforcement Sensitivity Theory and cognitive architectures. *Association for the Advancement of Artificial Intelligence*.

Fua, K., Revelle, W., and Ortony, A. (2010). Modeling personality and individual differences: The approach-avoid-conflict triad. *Annual Conference of the Cognitive Science Society*, pages 1–6.

Galen (1916). *Galen on the Natural Faculties*. (Brock, A. J., Trans.). G.P. Putnam's Sons, New York. (Original work published c. 460 BCE).

Galton, F. (1869). *Hereditary genius, an inquiry into its law and consequences*. Macmillan and Co., London.

Galton, F. (1884). The measurement of character. *Fortnightly Review*, 42:179–185.

Gambardella, A. and Hall, B. H. (2006). Proprietary versus public domain licensing of software and research products. *Research Policy*, 35(6):875–892.

Gilliland, A. R. (1928). Problems of personality. *The Journal of Abnormal and Social Psychology*, 23(3):369.

Giorgi, A. (1975). Convergence and divergence of qualitative and quantitative methods in psychology. In Giorgi, A., Fischer, C. T., and Murray, E. L., editors, *Duquesne studies in phenomenological psychology*, pages 72–79. Duquesne University Press, Pittsburgh.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In Wheeler, L., editor, *Review of personality and scoial psychology*, pages 141–165. Sage Publications, Beverly Hills.

Goldberg, L. R. (1990). An alternative “description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59(6):1216–1229.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1):26–42.

Goldberg, L. R. (1993a). The structure of personality traits: Vertical and horizontal. In Funder, D. C., Parke, R. D., Tomlinson-Keasey, C., and Widaman, K., editors, *Studying lives through time: Personality and development*, pages 169–188. American Psychological Association, Washington, DC.

Goldberg, L. R. (1993b). The structure of phenotypic personality traits. *American Psychologist*, 48:26–34.

Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several Five-Factor Models. In Mervielde, I., Deary, I., De Fruyt, F., and Ostendorf, F., editors, *Personality Psychology in Europe*, pages 1–7. Tilburg University Press, Tilburg, The Netherlands.

Goldberg, L. R. (2009). How to win a career achievement award in five easy lessons. *Journal of Personality Assessment*, 91(6):506–517.

Goldberg, L. R. (2010). Personality, demographics, and self-reported behavioral acts: The development of avocational interest scales from estimates of the amount of time spent in interest-related activities. In Agnew, C. R., Carlston, D. E., Graziano, W. G., and Kelly, J. R., editors, *Then a miracle occurs: Focusing on behavior in social psychological theory and research*, pages 205–226. Oxford University Press, New York.

Goldberg, L. R. (2014). *International Personality Item Pool: A scientific collaboratory for the development of advanced measures of personality traits and other individual differences.* Retrieved January 18, 2014, from http://ipip.ori.org/.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., and Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1):84–96.

Goldberg, L. R. and Velicer, W. F. (2006). Principles of exploratory factor analysis. In *Differentiating normal and abnormal personality*, pages 209–237. Springer, New York.

Goldstein, G. and Beers, S. R., editors (2004). *Comprehensive Handbook of Psychological Assessment, Volume 1: Intellectual and Neuropsychological Assessment*. John Wiley & Sons, Inc., Hoboken, NJ.

Gorsuch, R. L. (1983). *Factor analysis*. Lawrence Erlbaum Associates, Hillsdale, NJ, 2 edition.

Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin*, 127(1):45–86.

Gosling, S. D. and John, O. P. (1999). Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science*, 8(3):69–75.

Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37:504–528.

Gosling, S. D., Vazire, S., Srivastava, S., and John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, 59(2):93–104.

Gottfredson, G. D., Jones, E. M., and Holland, J. L. (1993). Personality and vocational interests: The relation of Holland's six interest dimensions to five robust dimensions of personality. *Journal of Counseling Psychology*, 40(4):518–524.

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24(1):79–132.

Gould, S. J. (2006). *The mismeasure of man (revised & expanded)*. W. W. Norton & Company.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60:549–576.

Gray, J. A. and McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system (2nd ed.)*. Oxford University Press, Oxford.

Greene, W. H. (2008). *Econometric Analysis*. Pearson Prentice Hall, Upper Saddle River, NJ, 6th edition edition.

Grucza, R. A. and 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(2):167–187.

Haisken-DeNew, J. and Frick, J. R. (2005). *DTC Desktop Companion to the German Socio-Economic Panel (SOEP)*. Deutsches Institut für Wirtschaftsforschung, Berlin.

Harman, H. H. (1976). *Modern factor analysis*. University of Chicago Press, 3 edition.

Harmon, L. W., Hansen, J. C., Borgen, F. H., and Hammer, A. L. (1994). Strong Interest Inventory applications and technical guide. Technical report, Palo Alto, CA.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In Berg, S. V., editor, *Annals of Economic and Social Measurement, Volume 5, number 4*, pages 475–492. NBER, Cambridge, MA.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Hegarty, M., Crookes, R. D., Dara-Abrams, D., and Shipley, T. F. (2010). Do all science disciplines rely on spatial abilities? Preliminary evidence from self-report questionnaires. In Hölscher, C., Shipley, T. F., Olivetti, M., Bateman, J., and Newcombe, N., editors, *Proceedings of the 7th international conference on Spatial cognition*, pages 85–94. Springer-Verlag, Berlin.

Heineman, P. L. (1995). *Temperament and personality*. Retrieved December 25, 2013, from https://personality-project.org/others/heineman/PERSON.HTM.

Hendriks, A. J. (1997). *The construction of the five-factor personality inventory (FFPI)*. PhD thesis, University of Groningen, Groningen.

Hendriks, A. J., Hofstee, W. K., and Raad, B. D. (1999). The five-factor personality inventory (FFPI). *Personality and Individual Differences*, 27(2):307–325.

Hendriks, A. J., Perugini, M., Angleitner, A., Ostendorf, F., Johnson, J. A., De Fruyt, F., Hrebicková, M., Kreitler, S., Murakami, T., Bratko, D., Conner, M., Nagy, J., Rodriguez-Fornells, A., and Ruisel, I. (2003). The five-factor personality inventory: Cross-cultural generalizability across 13 countries. *European Journal of Personality*, 17(5):347–373.

Herrnstein, R. J. and Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. Free Press, New York.

Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., and Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology*, 31(1):3–23.

Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences*, 16(2):107–117.

Hippocrates and Galen (1846). *The writings of Hippocrates and Galen*. Coxe, J. R. (Ed.) Lindsay and Blakiston, Philadelphia.

Hofmann, R. J. (1977). Indices descriptive of factor complexity. *The Journal of General Psychology*, 96(1):103–110.

Hofmann, R. J. (1978). Complexity And simplicity as objective indices descriptive of factor solutions. *Multivariate Behavioral Research*, 13(2):247–250.

Hofstee, W. K. (1991). The concepts of personality and temperament. In Strelau, J. and Angleitner, A., editors, *Explorations in temperament: International perspectives on theory and measurement*, pages 177–188. Plenum Press, New York.

Hofstee, W. K., De Raad, B., and Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63(1):146.

Hofstee, W. K. B. (2007). Unbehagen in individual differences. *Journal of Individual Differences*, 28(4):252–253.

Hogan, R. (1992). *Hogan personality inventory manual*. Hogan Assessment Systems, Tulsa, OK.

Hogan, R., Hogan, J., and Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, 51(5):469.

Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, 6(1):35.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Psychological Assessment Resources, Odessa, Florida.

Horn, J. L. and Engstrom, R. (1979). Cattell's scree test in relation to Bartlett's chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research*, 14(3):283–300.

Hornstein, G. A. (1988). Quantifying psychological phenomena: Debates, dilemmas, and implications. In Morawski, J. G., editor, *The rise of experimentation in American psychology*, pages 1–34. Yale University Press, New Haven.

Hornstein, G. A. (1992). The return of the repressed: Psychology's problematic relations with psychoanalysis, 1909–1960. *American Psychologist*, 47(2):254.

Hoyningen-Huene, P. (1993). *Reconstructing scientific revolutions: Thomas S. Kuhn's philosophy of science*. University of Chicago Press, Chicago.

Hsu, L. M. (1992). Correcting correlations of personality scales for spurious effects of shared items. *Multivariate Behavioral Research*, 27(1):31–41.

Hsu, L. M. (1994). Item overlap correlations: Definitions, interpretations, and implications. *Multivariate Behavioral Research*, 29(2):127–140.

Hu, L. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1):1–55.

ICAR (2014). *International Cognitive Ability Resource*. Retrieved December 19, 2013, from http://www.ICAR-project.org/.

ISSID (2014). *The International Society for the Study of Individual Differences*. Retrieved December 19, 2013, from http://www.issid.org/.

Jastrow, J. (1901). Some currents and undercurrents in psychology. *Psychological Review*, 8(1):1.

Jastrow, J. (1915). *Character and temperament*. D. Appleton and Company, New York.

John, O. P. and Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In Pervin, L. and John, O. P., editors, *Handbook of personality: Theory and research*, pages 102–138. Guilford Press, New York.

Johnson, S. (2010a). *Where good ideas come from: The natural history of innovation*. Riverhead Books, New York.

Johnson, W. (2007). Genetic and environmental influences on behavior: Capturing all the interplay. *Psychological Review*, 114(2):423–440.

Johnson, W. (2010b). Understanding the Genetics of Intelligence: Can Height Help? Can Corn Oil? *Current Directions in Psychological Science*, 19(3):177–182.

Johnson, W. and Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33(4):393–416.

Johnson, W. and Bouchard, T. J. (2009). Linking abilities, interests, and sex via latent class analysis. *Journal of Career Assessment*, 17(1):3–38.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1):31–36.

Kanai, R. and Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, 12:231–242.

Kant, I. (1790). *The critique of judgment*. Henry Holt and Company, Berlin, 1 edition.

Kant, I. (1979). Metaphysical foundations of natural science. In Watson, R. I., editor, *Basic writings in the history of psychology*. Oxford University Press, New York.

Kappeler, P. M., Silk, J. S., Burkart, J. M., and van Schaik, C. P. (2010). Primate behavior and human universals: Exploring the gap. In *Mind the gap: Tracing the origins of human universals*, pages 3–15. Springer, New York.

Kaufmann, P. (2009). Protecting raw data and psychological tests from wrongful disclosure: A primer on the law and other persuasive strategies. *The Clinical Neuropsychologist*, 23(7):1130–1159.

Kenny, D. A. (2012). *Measuring model fit.* Retrieved November 7, 2012, from http://www.davidakenny.net/cm/fit.htm.

Klages, L. (1929). *The science of character.* G. Allen & Unwin, London.

Kluckhohn, C. and Murray, H. A. (1948). Personality formation: The determinants. In Kluckhohn, C. and Murray, H. A., editors, *Personality in nature, society, and culture*, pages 1–9. Alfred A. Knopf, New York.

Koenig, K. A., Frey, M. C., and Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence*, 36(2):153–160.

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*.

Krueger, R. F. and Johnson, W. (2008). Behavioral genetics and personality: A new look at the integration of nature and nurture. In John, O. P., Robins, R. W., and Pervin, L. A., editors, *Handbook of Personality: Theory and Research*, pages 287–310. Guilford Press, New York.

Krueger, R. F. and Markon, K. E. (2014). The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology*, 10(1):131213170432000.

Krueger, R. F., South, S., Johnson, W., and Iacono, W. (2008). The heritability of personality is not always 50%: Gene-environment interactions and correlations between personality and parenting. *Journal of Personality*, 76(6):1485–1521.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* University of Chicago Press, Chicago, 1st edition.

Kuhn, T. S. (1970). *The structure of scientific revolutions.* University of Chicago Press, Chicago, 2nd edition.

Kuncel, N. R., Crede, M., and Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, 75(1):63–82.

Kuncel, N. R., Ones, D. S., and Sackett, P. R. (2010). Individual differences as predictors of work, educational, and broad life outcomes. *Personality and Individual Differences*, 49(4):331–336.

Laland, K. N. and Brown, G. R. (2006). Niche construction, human behavior, and the adaptive-lag hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, 15(3):95–104.

Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist*, 36(3):276.

Lamiell, J. T. (2003). *Beyond individual and group differences: Human individuality, scientific psychology, and William Stern's critical personalism*. Sage, Thousand Oaks, CA.

Lee, K. and Ashton, M. C. (2004). Psychometric Properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, 39(2):329–358.

Lee, K., Ogunfowora, B., and Ashton, M. C. (2005). Personality traits beyond the Big Five: Are they within the HEXACO space? *Journal of Personality*, 73(5):1437–1463.

Legge, J. (1879). *The texts of Confucianism: The Shu King, the religious portions of the Shih King, the Hsiao King*, volume 3 of *The sacred books of China*. Muller, F. M. (Ed.). Macmillan and Co., Oxford.

Liao, H.-Y., Armstrong, P. I., and Rounds, J. (2008). Development and initial validation of public domain Basic Interest Markers. *Journal of Vocational Behavior*, 73(1):159–183.

Loevinger, J. (1987). *Paradigms of personality*. W H Freeman & Company, New York.

Lord, F. M. (1955). Sampling fluctuations resulting from the sampling of test items. *Psychometrika*, 20(1):1–22.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "'General Intelligence,' Objectively Determined and Measured". *Journal of Personality and Social Psychology*, 86(1):96–111.

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences*, 49:344–351.

Lubinski, D. and Benbow, C. P. (2006). Study of Mathematically Precocious Youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, 1(4):316–345.

Lubinski, D. and Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, 24(1):159–201.

May, M. A. and Hartshorne, H. (1927). Personality and character tests. *Psychological Bulletin*, 24(7):395–411.

Mayer, R. E., Stull, A. T., Campbell, J., Almeroth, K., Bimber, B., Chun, D., and Knight, A. (2006). Overestimation bias in self-reported SAT scores. *Educational Psychology Review*, 19(4):443–454.

McAdams, D. (2001). The psychology of life stories. *Review of General Psychology*, 5(2):100.

McAdams, D. and Pals, J. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, 61(3):204–217.

McCaffrey, R. J. and Lynch, J. K. (2009). Test security in the new millennium: Is this really psychology's problem? *Emerging Fields*, 21(2):27.

McCrae, R. R. and Costa, P. T. (1985). Updating Norman's " adequacy taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49(3):710–721.

McCrae, R. R. and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1):81.

McCrae, R. R. and Costa, Jr, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5):509–516.

McDonald, R. P. (1985). *Factor analysis and related methods*. Lawrence Erlbaum, Hillsdale, NJ.

McDougall, W. (1923). *An outline of psychology*. Methuen and Co. Limited, London.

McDougall, W. (1929). The chemical theory of temperament applied to introversion and extroversion. *Journal of Abnormal and Social Psychology*, 24(3):293–309.

McDougall, W. (1932). Of the words character and personality. *Character and Personality*, 1(1):3–16.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37(1):1–10.

McKay, D. A. and Tokar, D. M. (2012). The HEXACO and five-factor models of personality in relation to RIASEC vocational interests. *Journal of Vocational Behavior*, 81(2):138–149.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press, Minneapolis.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46(4):806–834.

Mehl, M. R., Gosling, S. D., and Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, 90(5):862–877.

Mislevy, R. J., Beaton, A. E., Kaplan, B., and Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29:133–161.

Molden, D. and Higgins, E. (2005). Motivated thinking. *The Cambridge Handbook of Thinking and Reasoning*, pages 295–317.

Murphy, L. L., Geisinger, K. F., Carlson, J. F., and Spies, R. A. (2011). *Tests in Print VIII*. An Index to Tests, Test Reviews, and the Literature on Specific Tests. Buros Institute of Mental Measurements, Lincoln, Nebraska.

Navarro, A. and Griffin, R. A. (2002). Matrix sampling designs for the year 2000 census. In *American Statistical Association, proceedings of the section on survey research methods*, pages 480–485.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, 61(6):622–631.

Newcombe, N. S. and Shipley, T. F. (2012). Thinking about spatial thinking: New typology, new assessments. In Gero, J. S., editor, *Studying visual and spatial reasoning for design creativity*, pages 2–13. Springer, New York.

Newcombe, N. S., Uttal, D. H., and Sauter, M. (2013). Spatial development. In Zelazo, P., editor, *Oxford handbook of developmental psychology*, pages 564–590. Oxford University Press, New York.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66(6):574–583.

Norman, W. T. (1967). 2800 Personality Trait Descriptors. Technical report.

Obschonka, M., Schmitt-Rodermund, E., Silbereisen, R. K., Gosling, S. D., and Potter, J. (2013). The regional distribution and correlates of an entrepreneurship-prone personality profile in the United States, Germany, and the United Kingdom: A socioecological perspective. *Journal of Personality and Social Psychology*.

OECD (2012). *PISA 2009 Technical Report*. OECD Publishing.

Paunonen, S. V. (2003). Big Five factors of personality and replicated predictions of behavior. *Journal of Personality and Social Psychology*, 84(2):411–424.

Peabody, D. (1987). Selecting representative trait adjectives. *Journal of Personality and Social Psychology*, 52(1):59.

Pearson, K. (1903). Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 200:1–66.

Peterson, C., Park, N., and Seligman, M. E. P. (2005). Orientations to happiness and life satisfaction: the full life versus the empty life. *Journal of Happiness Studies*, 6(1):25–41.

Peterson, C. and Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press, Oxford.

Pettersson, E. and Turkheimer, E. (2010). Item selection, evaluation, and simple structure in personality data. *Journal of Research in Personality*, 44(4):407–420.

Pinker, S. (2002). *The blank slate*. Penguin Books, New York.

Plomin, R. and Rende, R. (1991). Human behavioral genetics. *Annual Review of Psychology*, 42(1):161–190.

Popper, K. R. (1959). *The logic of scientific discovery*. Basic Books, New York. (Original work published 1935).

Pozzebon, J. A., Visser, B. A., Ashton, M. C., Lee, K., and Goldberg, L. R. (2010). Psychometric characteristics of a public-domain self-report measure of vocational interests: The Oregon Vocational Interest Scales. *Journal of Personality Assessment*, 92(2):168–174.

Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior*, 21(3):259–287.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Revelle, W. (1986). Motivation and efficiency of cognitive performance. In Brown, D. R. and Veroff, J., editors, *Frontiers of Motivational Psychology: Essays in honor of J. W. Atkinson*, pages 105–131. Springer-Verlag, Berlin.

Revelle, W. (2014). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University, Evanston, Illinois. R package version 1.4.12.

Revelle, W. and Brown, A. (2013). Standard errors for SAPA correlations. In *Society for Multivariate Experimental Psychology*, pages 1–10, St. Petersburg, FL.

Revelle, W. and Condon, D. M. (2012). Estimating ability for two samples.

Revelle, W. and Condon, D. M. (in press). Reliability. In Irwing, P., Booth, T., and Hughes, D., editors, *Wiley-Blackwell Handbook of Psychometric Testing*, pages 1–38. Wiley-Blackwell, Hoboken, New Jersey.

Revelle, W., Wilt, J., and Condon, D. M. (2010a). Levels of personality. *European Journal of Personality*, 24:404–422.

Revelle, W., Wilt, J., and Rosenthal, A. (2010b). Individual differences in cognition: New methods for examining the personality-cognition link. In Gruszka, A., Matthews, G., and Szymura, B., editors, *Handbook of Individual Differences in Cognition: Attention, Memory and Executive Control*, chapter 2, pages 27–49. Springer, New York.

Revelle, W. and Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1):145–154.

Rigdon, E. E. (1997). *Not positive definite matrices—causes and cures.* Retrieved January 26, 2014, from http://www2.gsu.edu/ mkteer/npdmatri.html.

Roback, A. A. (1927). *The psychology of character.* Routledge & Kegan Paul LTD., London, 1st edition.

Roback, A. A. (1952). *The psychology of character: With a survey of personality in general.* Routledge & Kegan Paul LTD., London, 3rd edition.

Roberts, B. (2006). Personality development and organizational behavior. *Research in Organizational Behavior: An Annual Series of Analytical Essays and Critical Reviews*, 27:1–40.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., and Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4):313.

Robertson, K. F., Smeets, S., Lubinski, D., and Benbow, C. P. (2010). Beyond the threshold hypothesis. *Current Directions in Psychological Science*, 19(6):346–351.

Rounds, J., Su, R., Lewis, P., and Rivkin, D. (2010). O*NET® Interest Profiler Short Form psychometric characteristics: Summary. *The National Center for O*NET Development*, pages 1–43.

Sackett, P. R. and Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, 85(1):112–118.

Sandy, C. J., Gosling, S. D., and Durant, J. (2013). Predicting Consumer Behavior and Media Preferences: The Comparative Validity of Personality Traits and Demographic Variables. *Psychology & Marketing*.

Saucier, G. (2003). Factor structure of English-language personality type-nouns. *Journal of Personality and Social Psychology*, 85(4):695–708.

Saucier, G. (2009a). Recurrent personality dimensions in inclusive lexical studies: Indications for a Big Six structure. *Journal of Personality*, 77(5):1577–1614.

Saucier, G. (2009b). What are the most important dimensions of personality? Evidence from studies of descriptors in diverse languages. *Social and Personality Psychology Compass*, 3(4):620–637.

Schmidt, F. L. and Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86(1):162.

Schönbrodt, F. D. and Perugini, M. (2013). Journal of Research in Personality. *Journal of Research in Personality*, 47(5):609–612.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Schwekendiek, D. (2009). Height and weight differences between North and South Korea. *Journal of Biosocial Science*, 41(1):51.

Shamdasani, S. (2003). *Jung and the Making of Modern Psychology: The Dream of a Science*. Cambridge University Press, Cambridge, UK.

Shao, L. (2007). *Classified characters and political abilities*. (Yinghuan, L., Trans.). Zhonghua Book Company, Beijing. (Original work published 200 CE).

Shiner, R. L. and DeYoung, C. G. (2013). The structure of temperament and personality traits: A developmental perspective. In Zelazo, P. D., editor, *The Oxford Handbook of Developmental Psychology*. Oxford University Press, New York.

Shipley, W. C., Gruber, C., Martin, T., and Klein, A. M. (2010). *Shipley Institute of Living Scale, 2nd edition*. Western Psychological Services, Los Angeles, CA.

Shipley, W. C., Gruber, C. P., Martin, T. A., and Klein, A. M. (2009). *Shipley-2*. Western Psychological Services, Los Angeles, CA.

Skitka, L. J. and Sargis, E. G. (2006). The Internet as psychological laboratory. *Annual Review of Psychology*, 57(1):529–555.

SMEP (2014). *Society for Multivariate Experimental Psychology*. Retrieved December 19, 2013, from http://www.SMEP.org/.

Smith, E. A., Mulder, M. B., and Hill, K. (2001). Controversies in the evolutionary social sciences: A guide for the perplexed. *Trends in ecology & evolution*, 16(3):128–135.

Smith, T. W., Marsden, P. V., and Hout, M. (2011). *General social surveys, 1972-2010*. National Opinion Research Center; Storrs, CT: The Roper Center for Public Opinion Research, Chicago.

Soto, C. J., John, O. P., Gosling, S. D., and Potter, J. (2010). Age differences in personality traits from 10 to 65: Big five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, 100(2):330–348.

Tellegen, A. and Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In Boyle, G. J., editor, *The SAGE Handbook of Personality Theory and Assessment: Volume 2 Personality Measurement and Testing*, pages 261–292. Sage Publications, Inc, Thousand Oaks, CA.

Thagard, P. R. (1978). Why astrology is a pseudoscience. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1978:223–234.

Thalmayer, A. G., Saucier, G., and Eigenhuis, A. (2011). Comparative validity of brief to medium-length Big Five and Big Six personality questionnaires. *Psychological Assessment*, 23(4):995–1009.

Theophrastus (1927). *The characters of Theophrastus*. Edmonds, J. M. (Ed.). G.P. Putnam's Sons, New York.

Thomas, A. and Chess, S. (1977). *Temperament and development*. Brunner/Mazel, Oxford, England.

Thorndike, R. L. (1949). *Personnel selection: Test and measurement techniques*. John Wiley & Sons Inc, London.

Thurstone, L. L. (1947). *Multiple factor analysis: A development and expansion of vectors of the mind*. University of Chicago Press, Chicago, IL.

Thurstone, L. L. (1954). An analytical method for simple structure. *Psychometrika*, 19(3):173–182.

Tolman, E. C. (1932). *Purposive Behavior in Animals and Men*. Appleton-Century, New York, 1st edition.

Toomet, O. and Henningsen, A. (2008). Sample selection models in R: Package sampleSelection. *Journal of statistical software*, 27(7):1–23.

Tucker, L. R. and Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1):1–10.

Tucker, L. R. and MacCallum, R. C. (1997). Exploratory factor analysis. *Unpublished manuscript, Ohio State University, Columbus*.

Tupes, E. C. and Christal, R. E. (1961). Recurrent personality factors based on trait ratings. Technical report, Personnel Laboratory United States Air Force.

U. S. Census Bureau (2012). Table 6: Resident population by sex, race, and hispanic-origin status: 2000 to 2009. *Statistical Abstract of the United States*, pages 1–1.

Uebersax, J. S. (2000). Estimating a latent trait model by factor analysis of tetrachoric correlations. Statistical Methods for Rater Agreement web site.

Uher, J. (2013). Personality Psychology: Lexical approaches, assessment methods, and trait concepts reveal only half of the story - why it is time for a paradigm shift. *Integrative Psychological and Behavioral Science*, 47(1):1–55.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., and Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2):352.

Vassend, O. and Skrondal, A. (1995). Factor analytic studies of the NEO Personality Inventory and the five-factor model: The problem of high structural complexity and conceptual indeterminacy. *Personality and Individual Differences*, 19(2):135–147.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3):321–327.

Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4):817–835.

Wang, N., Kosinski, M., Stillwell, D. J., and Rust, J. (2012). Can well-being be measured using Facebook status updates? Validation of Facebook's Gross National Happiness Index. *Social Indicators Research*.

Wechsler, D. (2008). Wechsler Adult Intelligence Scale. Technical report, San Antonio, TX.

Weinstein, T. A. R., Capitanio, J. P., and Gosling, S. D. (2008). Personality in animals. In John, O. P., Robins, R. W., and Pervin, L. A., editors, *Handbook of personality: Theory and research*, pages 1–21. Guilford Press, New York.

Weiss, K. M. and Lambert, B. W. (2011). When the time seems ripe: Eugenics, the Annals, and the subtle persistence of typological thinking. *Annals of Human Genetics*, 75(3):334–343.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37(3):395.

Wiggins, J. S. (2003). *Paradigms of personality assessment*. Guilford Press, New York.

Wilt, J., Condon, D. M., and Revelle, W. (2011). Telemetrics and online data collection: Collecting data at a distance. In Laursen, B., Little, T. D., and Card, N. A., editors, *Handbook of Developmental Research Methods*, chapter 10, pages 163–180. Guilford Press, New York.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In Bollen, K. A. and Long, J. S., editors, *Testing structural equation models*, pages 256–293. Sage Publications, Newbury Park, CA.

Yarkoni, T. (2013). *Principled failure to detect a five-factor structure in a canonical Big Five measure*. Retrieved January 18, 2014, from http://figshare.com/.

Zinbarg, R. E., Revelle, W., Yovel, I., and Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ω_h: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70(1):123–133.

Zinbarg, R. E., Yovel, I., Revelle, W., and McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for omega hierarchical. *Applied Psychological Measurement*, 30(2):121–144.