Individual Differences and Their Measurement: A Review of 100 Years of Research

Paul R. Sackett
University of Minnesota

Chad H. Van Iddekinge
Florida State University

Filip Lievens
Ghent University

Nathan R. Kuncel
University of Minnesota

This article reviews 100 years of research on individual differences and their measurement, with a focus on research published in the *Journal of Applied Psychology*. We focus on 3 major individual differences domains: (a) knowledge, skill, and ability, including both the cognitive and physical domains; (b) personality, including integrity, emotional intelligence, stable motivational attributes (e.g., achievement motivation, core self-evaluations), and creativity; and (c) vocational interests. For each domain, we describe the evolution of the domain across the years and highlight major theoretical, empirical, and methodological developments, including relationships between individual differences and variables such as job performance, job satisfaction, and career development. We conclude by discussing future directions for individual differences research. Trends in the literature include a growing focus on substantive issues rather than on the measurement of individual differences, a differentiation between constructs and measurement methods, and the use of innovative ways of assessing individual differences, such as simulations, other-reports, and implicit measures.

**Keywords:** individual differences, ability, personality, interests, motivation

**Supplemental materials:** [http://dx.doi.org/10.1037/ap0000151.supp](http://dx.doi.org/10.1037/ap0000151.supp)

Possibly the greatest achievement of the members of the American Psychological Association is the establishment of the psychology of individual differences. (Scott, 1920, p. 85)

The development of standardized measures of attributes on which individuals differ emerged very early in psychology’s history, and has been a major theme in research published in the *Journal of Applied Psychology* (*JAP*). Numerous questions can be, and have been, asked regarding individual differences, including (a) their origins, including evolutionary, genetic, and situational causes; (b) their dimensionality, with an eye to a parsimonious way of summarizing differences between people; (c) their measurement; (d) their stability over time; and (e) their usefulness for applied purposes, such as forecasting future behavior in contexts such as personnel selection. Issues of the origins of individual differences have rarely been the purview of research in *JAP* (for notable exceptions, see Arvey, Bouchard, Segal, & Abraham [1989] and Shane, Nicolaou, Cherkas, & Spector [2010]). Dimensionality, measurement, stability, and applied use of individual differences have been major themes . . . in *JAP*.

Variables on which individuals differ can be arrayed on a continuum from stable to transitory (Ackerman & Humphreys, 1990). The term “individual differences” is generally reserved for attributes nearer the “stable” end of this continuum. In an organizational context, it is useful to think of individual differences as features that individuals bring with them to the job. Thus, ability, personality, interest patterns, and motivational traits (e.g., achievement motivation, core self-evaluations [CSEs]) fall under the individual differences umbrella, whereas variables that are transient, such as mood, or that are closely linked to the specifics of the work setting (e.g., turnover intentions or perceived organizational climate), do not. We note that “stable” does not necessarily mean “unchangeable”; knowledge and skill are examples of individual difference variables that can be altered through investment of time and effort. Further, it has become clear that variables initially conceptualized as job or organization specific (e.g., job satisfaction) contain some dispositional variance (e.g., Arvey et al., 1989).

Various taxonomies of individual differences have been put forward (e.g., Ackerman & Humphreys, 1990; Murphy, 2012; Peterson et al., 1990). Our review is not driven by any given taxonomic structure, but rather by the individual difference variables that have been the focus of research during the first century of *JAP*. We focus on three major sets of topics. The first is knowledge, skills, and abilities in the cognitive and psychomo-
tor/physical domains. The second is personality, including two
topics of considerable interest in recent years, namely, integrity
testing and emotional intelligence (EI), as well as motivational
traits (e.g., achievement striving and CSE) and creativity. The third
is vocational interests. For each of these domains, we offer a
historical perspective, with a primary focus on articles published in
JAP, though we also cite key works published elsewhere. In
addition, we used citation counts to help identify the most influ-
ential articles in each domain. We conclude by discussing some
general observations and possible future directions for individual
differences research.

Knowledge, Skills, and Abilities

Cognitive Abilities

The study of cognitive abilities in JAP has paralleled the broader
literature, but with a more pragmatic and applied focus. The very
first volume of the journal contained research on topics that are
core topics of research today, including criterion-related validity
(Terman et al., 1917), bias and group differences (Sunne, 1917),
measurement issues (Miner, 1917; Yerkes, 1917), and relation-
ships with learning and the development of knowledge and skill
(Bingham, 1917). The earliest research addressed questions about
the nature and usefulness of ability. Indeed, the majority of em-
pirical papers that appeared in the early years of the journal
focused on measures of cognitive abilities. Unlike other individual
differences, like interests and personality, research on cognitive
abilities has remained prominent and focused around the same
themes. Darley (1968), in his 50-year review of JAP, noted that
“studies of individual mentalities” were heavily published over the
first 50 years of the journal. Although the originally used term
“intelligence” fell out of favor and was largely replaced by “cog-
nitive ability,” research on these themes has continued, but with a
shift that comes with becoming an important and established
determinant of human behavior.

Over the years, studies have reflected the tension between
viewing cognitive abilities as enduring capacities that are largely
innate versus treating them as measures of developed capabilities,
a distinction that has implications for the study of criterion related
validity, group differences, aging effects, and the structure of
human abilities (Kuncel & Beatty, 2013). Terman et al. (1917)
argued for the importance of distinguishing between “the [poor]
intellectual status of an individual . . . [versus] his inability to get
on in the world” (p. 23). Terman et al. were convinced that
intelligence tests were solid measures of the former, writing, “The
diagnosis of intellectual feebleness is absolute, as intelligence is a
definitely measureable thing” (p. 23).

Others researchers have not been so certain. For example, stud-
ies of the intelligence of children, quite common in the early
decades of JAP, debated the causal direction between environment
and abilities. Even preferences for play activities (boxing vs.
drawing vs. reading comics) were discussed theoretically as having
a causal, reverse-causal, or reciprocal influences on developed
verbal abilities (Lehman & Witty, 1928). Other studies wrestled
with the developed versus innate issue in the context of group
differences (Garth, Serafini, & Dutton, 1925; Wang, 1926). Garth
et al. (1925) noted the role of environmental differences in up-
bringing when comparing two groups, and noted that “because of
differences in social status and temperament we cannot conclude
that our results are true and final measures of the intelligence of
Indian [sic] children” (p. 389; italics added). Here “true and final”
means something absolute and innate about the person. In any
case, the concern was with the applied implications of these scores.

This pragmatic focus created a theme of asking, “What is an
intelligence test good for?” But again, the stage was set in the first
volume of JAP, in which it was noted that intelligence was clearly
not the only individual difference that mattered in occupational
performance (Terman et al., 1917). The question was always one
of relative importance, but with gradual changes in how this
question was answered. The greatest change was in the range and
complexity of criteria studied. Early authors appeared to be satis-
ified with using salary as a measure of job performance (Bingham
& Davis, 1924). Over time, notable studies examined decision
making (Taylor & Dunnette, 1974), multidimensional models of
job performance (LePine & Van Dyne, 2001), and leadership
(Judge, Colbert, & Ilies, 2004), among many others.

The types of jobs studied were varied from the beginning, with
some, like police and firefighters, becoming mainstays, and the
study of managers and business leaders a constant and central
focus. Influential papers on managers appeared in early (Bingham
& Davis, 1924), middle (Csoka, 1974; Taylor & Dunnette, 1974),
and recent years (Judge et al., 2004). Later studies have been
heavily influenced by the development of validity generalization
(and meta-analysis), which was initially illustrated with data on
cognitive abilities and training and performance criteria (Schmidt
& Hunter, 1977). This has resulted in a gradual acceptance of the
importance of cognitive abilities. The question is not so much
whether cognitive abilities are important but how they fit among
other factors and to what limits.

Questions about the limits of cognitive ability began early, with
research comparing the accomplishments of gifted students with
others (Lehman & Witty, 1928), as well as with a case study of
Beatty Ford, who had an IQ of 188 and was reading Shakespeare
when 8 years old (Terman & Fenton, 1921). Subsequent studies
grouped participants into score bands to examine whether scores
had relationships at all levels or whether there were points beyond
which scores ceased to matter. Research examining extreme
groups concluded that intelligence scores matter even at 1-in-
10,000 scarcity (Lubinski, Webb, Morelock, & Benbow, 2001).
What has shifted is the advent of stronger methods and the real-
ization that testing at the extremes requires tests designed to
adequately differentiate among test takers.

The study of using ability measures to predict future behavior
has gone hand in hand with a concern for group differences.
Convenience samples dominated most of the early research of
group differences (Pressey & Teter, 1919; Yeung, 1921), although
some studies had large population samples that would make any
author proud and whose analyses must have been very labor
intensive (e.g., Pressey, 1918). A number of authors reacted to
evidence of group differences by raising methodological concerns
about the interpretability of the research. Shared common language
(Wang, 1926), age (Arlitt, 1922), social class (Arlitt, 1921), and
education (Jordan, 1933) were all raised as potential confounds for
drawing conclusions based on observed differences. Creating al-
ternate tests to address differences has also been a mainstay,
starting with early efforts at culture-free tests (Pintner, 1919). It is
important to note that interest in understanding group differences
was not initially motivated by legal concerns, as much of it predates civil rights legislation. Instead, scholars appeared to pursue interesting questions to understand a new technology: the group intelligence test.

Recent decades have frequently yielded research that incorporates ability as one of a number of variables to be considered for topics ranging from team effectiveness (Neuman & Wright, 1999) to ability’s interaction with motivational processes in skill acquisition (Kanfer & Ackerman, 1989). Ability also has played a central role in research on the construct validity of other methods of measurement, including interviews (Huffcutt, Roth, & McDaniel, 1996) and situational judgment tests (SJTs; McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001). Across this work, general cognitive ability is now regularly used as a fundamental individual difference and is positioned as the standard against which other predictors often are compared (Ree, Earles, & Teachout, 1994). In addition, survey data suggest that some of the questions concerning cognitive ability explored at the start of the journal are now topics of broad consensus (Murphy, Cronin, & Tam, 2003). This represents a major change from its start as an interesting but promising characteristic to be explored.

Currently, the field generally agrees on a hierarchical structure, with general ability at its top and more specific abilities below. Although research on the structure of human abilities and issues surrounding their measurement was more prominent in specialty journals, several influential JAP studies addressed fundamental applied concerns. Research covered applied topics like the structure of mechanical abilities (Goodman, 1947), various clerical skills (Bair, 1951), and a series of studies on the structure and nature of psychomotor skills related to performing Morse code (Fleishman & Hempel, 1956). Studies examining the structure of paper-and-pencil cognitive ability tests ultimately waned in parallel with the broader literature on factor analyses of measures of human abilities. Contributions to methods also started early, including one of the earliest comparisons of age-based IQ measures to the now more prevalent point-based methods (Yerkes, 1917). Influential research on measurement equivalence, differential item functioning, and expansions of multivariate-multimethod analysis also have appeared in JAP (e.g., Marsh & Hocevar, 1988; Raju, Laffitte, & Byrne, 2002; Stark, Chernyshenko, & Drasgow, 2006).

Knowledge and Skill

Individual differences in knowledge and skills have long been a part of JAP. In addition to the direct study of knowledge and skill measurement, these individual differences have been a part of research on job analysis, leadership, career development, performance appraisal, training, and skill acquisition, among others. In fact, these characteristics so thoroughly permeate many domains of study that a comprehensive review would be both unmanageable and cross over into too many other domains. Instead, we will limit ourselves to how individual differences in knowledge and skill have been measured and applied.

Ability, aptitude, achievement, knowledge, and skill have all been invoked and measured in reference to individual differences within the cognitive domain. Sometimes a different term is used for effectively the same construct; one calls it a skill, and another, an aptitude. Fundamentally, the measurement of knowledge and skill is intimately connected to the measurement of intelligence or cognitive abilities (Kuncel & Beatty, 2013; Lubinski & Dawis, 1992). Thinking clearly about the underlying distinctions has been an unfolding process. As noted in the cognitive abilities section, even the earliest authors noted the difference between the potential a person has for learning and mental work (something ability-like) versus differences concerning what a person currently can do (something skill-like).

Generally, knowledge has referred to facts a person knows; skills, to what a person currently can do; aptitude, to a person’s potential to learn, achievement, to what a person acquired over a given period of time; and ability has been used to mean all of the above. Even at the start, scholars were concerned with disentangling the sources of differences among these concepts. For example, Gates (1918) discussed differences in expert performance in marksmanship (a skill) as a function of “native ability—ability possessed before practice” and “superiority of the expert . . . due to training in shooting” (p. 12).

Research examining differences in ability, knowledge, and skill measures was ultimately reviewed and organized into a taxonomy in which such measures are said to differ on the type of content (or curricula) measured, the breadth of content covered, how recently that content was learned, and the purpose of the assessment (Cleary, Humphreys, Kendrick, & Wesman, 1975; Lubinski & Dawis, 1992). Therefore, a measure constructed with the purpose of measuring the capability to learn new material quickly might be a broad measure of historical learning from a K-12 education. Similarly, it might be unnecessary to directly measure a skill, as it might be well represented by a broad sampling of curricula (related to the skill that has developed over time). In contrast, assessments of training effectiveness tend to have narrow breadth, very specific content, and reflect very recent learning.

Ideas about breadth and content appeared early in JAP. For example, Cuff (1930) argued that simple vocabulary tests are a good indicator of reading skill and the overall achievement of students. The influence of experience and learning on skill and performance was also studied in detail. Following up on Gates (1918), Humphreys, Buxton, and Taylor (1936) also examined marksmanship performance, noting that (a) body steadiness was strongly associated with performance among experienced shooters, (b) different measures of steadiness exhibited a general factor, (c) practice improved steadiness but, (d) even after considerable practice, most people did not gain enough to put them at a level comparable with competitive college students.

Inherent in these ideas is a theory of how talent, investment, experience, and practice lead to skill and performance. The implicit theory was that there are differences in potential or talent, which, with appropriate learning opportunities, lead to acquiring knowledge and skill better and faster. More talented people with better instruction and considerable practice tend to become the most knowledgeable and skilled. Knowledge and skill, in turn, are likely to be the most proximal determinants of differences in job performance, but only if certain conditions, namely, opportunity to learn, apply. For example, Gates (1918) mocked testing for marksman’s skill potential by a direct skill test, writing, “Some may say, ‘Why not give a man a gun and let him show his ability by shooting a few times!’” (p. 13). He argued that an absence of basic experiences and knowledge (now called declarative knowledge) would interfere with evaluating how good of a marksman a person could become. Measuring more basic abilities (e.g., body stead-
ness), combined with performance after some basic instruction, would be a better indicator of long-term performance.

This research set the stage for two related branches of research: ability and skill acquisition, and models of ability and skill as determinants of job performance. The latter largely incorporates the former by assuming that general ability leads to skill development, and each includes research about the importance of specific measures for prediction compared with broader measures (e.g., general ability). Landmark papers on both topics appeared in JAP.

Longitudinal research measuring individual differences in skill acquisition began in the journal with Chapman (1919), who studied acquiring typewriting skill and noted large individual differences even after practice, as well as strong correlations between initial and later performance. This largely descriptive research was the beginning of a long line of research that examined skill acquisition and discovered the importance of consistent versus inconsistent tasks in understanding the predictive power of ability measures, as well as that ability components were important for understanding complex performance (Ackerman, 1992; Farrell & McDaniel, 2001).

Relationships between ability, skills, and performance were ultimately examined with large-scale studies and supported the model that general ability leads to knowledge and skill, which, in turn, are related to evaluations of job performance (Schmidt, Hunter, & Outerbridge, 1986). Subsequent studies considered motivational determinates as additional mediators between ability and performance (McCloy, Campbell, & Cudeck, 1994), although ability and motivation do not seem to interact (Sackett, Grays, & Ellingon, 1998). Most of this research, however, examined broad test batteries and measures of general cognitive ability instead of very specific measures of ability. Table A1 of the online supplemental materials presents a chronological listing of key JAP articles in the knowledge, skill, and cognitive ability domain.

Sensory, Psychomotor, and Physical Abilities

Sensory abilities (e.g., visual acuity), psychomotor abilities (e.g., finger dexterity, eye-hand coordination), and physical abilities (e.g., explosive strength, cardiovascular endurance) are important contributors to successful performance for select subsets of jobs. Visual acuity was an active research domain in JAP in the 1940s through the 1960s, but has not continued beyond that period. Representative articles include McCormick’s (1950) analysis of the relationship between visual acuity and job performance across 92 jobs for which acuity was judged relevant and Kephart’s (1948) examination of the relationship between visual acuity and turnover. The most cited JAP article on the topic is Burg’s (1966) examination of the relationship between static and dynamic visual acuity (i.e., acuity in perceiving stationary vs. moving objects).

Psychomotor abilities also received considerable research attention during this period. Highly cited articles in JAP include Fleishman and Hempel’s (1956) factor analysis of complex psychomotor tasks, Fleishman’s (1957) comparison of psychomotor performance of skilled and unskilled individuals, and Fleishman and Ellison’s (1962) factor analysis of performance on fine manipulative tests. Further, Locke and Bryan (1966) demonstrated that setting specific, difficult goals has a positive effect on psychomotor test performance, thus demonstrating that there is a cognitive component to performance on psychomotor tests.

Articles on physical ability testing have surfaced sporadically in JAP, commonly describing validation efforts for specific jobs or job families. Representative articles include Reilly, Zedeck, and Tenopyr’s (1979) development of a physical ability battery for craft jobs, Arnold, Rauschenberger, Soubel, and Guion’s (1982) validation of a strength test for steelworkers, and Arvey, Landon, Nutting, and Maxwell’s (1992) development of a physical ability test battery for police officers. Hogan (1991) reported a factor analysis of physical abilities, which revealed three overarching factors: strength, endurance, and movement quality. Handbook chapters by Baker and Gebhardt (2012) and Gebhardt and Baker (2010) have summarized the state of the physical ability testing field. They note that work on this topic comes from multiple disciplines, including exercise physiology, biomechanics, industrial engineering, and medicine, in addition to psychology. There are literatures on key issues in the use of physical ability testing, including job analytic techniques to identify the physical ability requirements of jobs, the development or selection of tests to assess these abilities, and the setting of cut scores.

An important issue is the finding of large mean gender differences on a number of physical abilities, such as muscular strength and endurance (Courtright, McCormick, Postlethwaite, Reeves, & Mount, 2013). This creates tensions between the dual goals of using ability tests for selection into jobs that require certain physical attributes and ensuring a gender-diverse workforce. Courtright et al. (2013) noted that subgroup differences vary across more specific abilities within Hogan’s (1991) broad ability factors, indicating that, in applied work, it will be useful to operate at a more specific level than these three general factors. Another key issue is the choice between using measures of the basic abilities identified by job analysis and the use of measures designed as simulations of specific job tasks (e.g., the use of a standardized measure of muscular endurance vs. measuring performance while engaged in an actual or simulated job task).

In sum, there is a history of research on sensory, psychomotor, and physical abilities within industrial and organizational (I/O) psychology, in general, and in JAP, in particular. This work is quite dated, and there has been very little research on these topics in recent decades, perhaps reflecting the changing nature of work and a reduction of the number of jobs with a substantial physical abilities component. This early work accomplished much in terms of developing measures of these abilities and understanding their factor structure. Today these domains largely constitute an active area of applied practice instead of active research (see Table A2 of the online supplemental materials for key JAP articles in this domain).

Personality

Our discussion of personality includes four sections. We open with a broad historical perspective on the conceptualization and measurement of personality. We follow this with treatments of three topics that can be viewed as part of a broader conceptualization of personality, namely, assessing integrity, EI, motivational traits, and creativity.

Historical Perspective on Personality

As shown by the personality-related articles published in JAP in the last 100 years, personality’s ride through the decades has not
been as smooth as the one taken by cognitive ability, knowledge, and skills. Personality’s research history (see Table A3 of the online supplemental materials) looks more like a rollercoaster and can be broken down in four main eras.

**Emergence (1917–1965).** In the first era, researchers explored the usefulness of personality as a complement to cognitive ability and interests. Further, the objectives for assessing personality were broad. For example, in the first *JAP* article on personality measures, Brandenburg (1925) stated the general purpose as follows: “to determine what factors may be discovered in the young man’s character or personality which would suggest his ultimate vocational selection and at the same time yield valuable information concerning the traits essential to success in the various vocations” (pp. 282–283). Early studies adopted a criterion-referenced strategy in which the personality profile of a successful group was compared with the profiles of less successful groups. This led to studies in *JAP* about personality traits associated with successful teachers, nurses, clerical workers, salesmen, and executives (e.g., Dodge, 1943; J. S. Guilford, 1932).

Inspection of the early personality articles in *JAP* further showed that the personality trait nomenclature and construct space were diverse. “Personality” included not only traditional traits such as introversion/extroversion (Eysenck, 1958) and achievement striving, but also maladaptive traits (e.g., Dorcus, 1944), emotional insight (Tendler, 1930), and sociability (Hunt, 1928). When multiple traits were measured, a term such as “average personality” was often used (e.g., Brandenburg, 1925). In these early days, the measurement of personality was also strikingly varied. Apart from self-reports (e.g., Eysenck, 1958; Gough, 1953; Humm & Wadsworth, 1941), personality was assessed via SJTs (Hunt, 1928), biodata-like questions (Gilliland & Burke, 1926), or projective tests (e.g., sentence completion: Rohde, 1946; picture arrangement: Miner & Culver, 1955; the Rorschach inkblot test: Hertz, 1934).

In a well-cited *JAP* article, Ghiselli and Barthol (1953) summarized this first era by reviewing 113 studies about the validity of personality inventories for predicting job performance. However, the researchers did not report results for individual personality traits but used a composite “personality rating.” Ghiselli and Barthol concluded that under certain circumstances scores on personality inventories correlate better with proficiency on a wider variety of jobs than might be expected. On the other hand there have been enough studies reporting negative results to emphasize caution in their use. These inventories have proved to be effective for some occupations in which personality factors would appear to be of minimal importance (e.g., clerks, and trades and crafts), and ineffective for other occupations in which these factors could reasonably be expected to be of paramount importance (e.g., supervisors and foremen). (p. 19–20)

As the above quote indicates, in this first era, the use of personality measures was not without criticism. Research on response distortion challenged self-reports, and interest grew in response distortion prevention and detection strategies. Meehl and Hathaway (1946) summarized research on the following approaches (of which many are still explored): Forced-choice among equally socially desirable items, insertion of subtle items, and lie and social desirability scales. Meehl and Hathaway also distinguished between unconscious and conscious faking, which are now referred to as self-deception and impression management, respectively. It was also recognized that social desirability “is not necessarily a nuisance factor in personality measurement and therefore something which should be eliminated” (Kriedt & Dawson, 1961, p. 177).

**Downfall (1965–1990).** The second area of personality research began with Guion and Gottier’s (1965) influential review, in which the authors concluded that there is no generalizable evidence that personality measures can be recommended as good or practical tools for employee selection. The number of significance tests resulting in acceptable statements of validity is greater than might be expected by pure chance—but not much. The best that can be said is that in some situations, for some purposes, some personality measures can offer helpful predictions. (p. 159)

This review, along with Mischel’s (1968) book on personality assessment that questioned the role of individual differences, led to a period in which personality was surrounded with substantial skepticism and only few articles on it were published in *JAP*. As an exception, in the late 1980s, there was a thread on Type A personality (e.g., Spence, Helmreich, & Pred, 1987).

**Resurgence (1991–2003).** Several methodological and conceptual developments led to a renaissance of personality research. First, the advent of meta-analysis enabled accumulating research findings across studies, thereby accounting for statistical artifacts (e.g., sampling error, unreliability, range restriction). Second, the five-factor model (FFM) emerged as a uniform conceptual framework for sorting the myriad personality traits. Third, there was recognition that the issue was not whether personality is relevant to work behavior, but rather which personality traits are relevant to which work behaviors.

In addition to the well-known meta-analyses of Barrick and Mount (1991) and Tett, Jackson, and Rothstein (1991) on the criterion-related validity of personality for predicting job performance, *JAP* also played an important role in this comeback of personality. In 1990, Hough and colleagues published a monograph on a construct-oriented examination of the validity of personality using data from the U.S. Army’s Project A (Hough, Eaton, Dunnette, Kamp, & McCloy, 1990). When personality traits were sorted into a taxonomy similar to the FFM and mapped to job performance dimensions (e.g., linking dependability to personal discipline), personality traits predicted targeted criteria. Subsequent studies by Hogan and Holland (2003) and Bartram (2005) corroborated the importance of predictor-criterion matching for personality validities.

In these years, the zeitgeist could best be described as exactly the opposite of the earlier dark period. The FFM was seen as paradigmatic, leading to an explosion of studies with personality traits as independent variables. Meta-analytic evidence solidified the relevance of personality as an individual difference in the work domain. For instance, Salgado (1997) extended earlier results of the validity of the FFM to the European Community. By using only measures designed to assess the FFM, Hurtz and Donovan’s (2000) meta-analysis addressed potential construct-related validity concerns of prior meta-analyses and found evidence for personality as a predictor of task and contextual performance. *JAP* also published meta-analyses that linked personality to variables such as motivation, job satisfaction, leadership, entrepreneurship, cre-
The resurgence of personality also produced a renewed interest in faking. Several meta-analyses, large-scale studies, and simulations were published in JAP that demonstrated the limited impact of socially desirable responding and faking on the construct-related and criterion-related validity of personality measures (Ellingson, Smith, & Sackett, 2001; Ones, Viswesvaran, & Reiss, 1996; and, later on, Hogan, Barrett, & Hogan, 2007; Schmitt & Oswald, 2006). Yet it was also acknowledged that in (top-down selection) applicants who intentionally distort their responses rise to the top of the distribution and have a greater chance of being hired (Mueller-Hanson, Heggestad, & Thornton, 2003).

Refinement (2004 to present). A conference panel discussion on faking in personality instigated yet another pendulum swing in personality research (published afterward in Morgeson et al., 2007). A panel of journal editors critically reviewed the evidence for the validity of personality and explored alternatives. Thus, this fourth era can be described as one in which the field searched for solutions to key contentious issues in the conceptualization and measurement of personality.

A first such longstanding issue deals with the modest size of validity coefficients for self-report measures of personality. Hence, over the last decade, we have witnessed more diversity in personality measurement, which is reminiscent of early personality studies. Research suggests that measuring personality using other-reports (e.g., Connelly & Ones, 2010; Oh, Wang, & Mount, 2011) or with personality inventories that refer to the work context (e.g., Lievens, De Corte, & Scholllaert, 2008) can increase predictive validity relative to more traditional self-report measures. In JAP, research has also appeared measuring personality with conditional reasoning tests (Bing et al., 2007), SJTs (Motowidlo, Hooper, & Jackson, 2006), structured interviews (Van Iddekinge, Raymark, & Roth, 2005), and ideal point models (Stark, Chernyshenko, Drasgow, & Williams, 2006), though not all of these approaches demonstrated incremental prediction over self-reports. Besides alternative measurement, personality validity has also been found to be somewhat larger for compound traits that reflect multiple FFM factors (see Hogan, Hogan, & Busch, 1984, for a pioneering example related to service orientation). There have also been further advancements in predictor-criterion matching. For instance, Chiaburu, Oh, Berry, Li, and Gardner’s (2011) meta-analysis revealed that Openness and Agreeableness were better predictors of contextual performance than of task performance.

Second, researchers have identified personality traits that the FFM may not capture. For instance, Ashton et al.’s (2004) HEXACO model adds a sixth factor (Honesty-Humility) to the FFM. In addition, there has been progress toward a taxonomy of lower level FFM facets (Judge, Rodell, Klinger, Simon, & Crawford, 2013), with the benefit that well-chosen facet measures can enhance the prediction of narrow criteria (e.g., Dudley, Orvis, Lebecki, & Cortina, 2006). Apart from the FFM traits, assessing social skills (e.g., Witt & Ferris, 2003), maladaptive personality traits and integrity (more in section below), and EI (see also below) has (again) become popular.

Third, various theoretical developments have refined conceptualizations of personality. Some of the more notable advancements are that (a) personality can interact with the situation to affect behavior (e.g., Tett & Burnett’s [2003] trait activation theory); (b) motivational forces mediate the effects of personality (e.g., Barrick, Stewart, & Piotrowski, 2002), with both implicit and explicit motives being important (Frost, Ko, & James, 2007; Lang, Zettler, Ewen, & Hulsheger, 2012); (c) personality is stable, yet also prone to change, across life (Woods & Hampson, 2010); (d) personality both affects and is affected by work (Wille & De Fruyt, 2014); and (e) personality traits represent stable distributions of variable personality states (Judge, Simon, Hurst, & Kelley, 2014; Minbashian, Wood, & Beckmann, 2010).

In sum, the history of personality research in JAP is characterized by waves of optimism and skepticism, as well as by waves of construct diversification and homogenization. Another common thread is that JAP has historically favored research on optimizing personality measurement, whereas conceptual developments typically have originated outside of JAP.

Integrity

Our review of the history of personality shows that in the 1920s to 1940s, personality inventories were often used to measure maladaptive personality traits, and therefore to identify people “on a basis of the fact that they showed signs of maladjustment and discontent, or were problem employees” (Dorcus, 1944, p. 302; see also Gilliland & Newman, 1953, and Zickar, 2001). This highlights that a heavy emphasis was put on predicting the counterproductive work behavior (CWB) job performance dimension.

Individual difference measures aimed at the prediction of CWB became a substantial area of research and applied practice in the 1970s with the development of self-report integrity tests. One approach, commonly labeled “overt integrity tests,” assesses attitudes toward and admissions of wrongdoing. The first JAP article on these measures (Ash, 1971) examined how they correlate with polygraph scores; more current work has focused on CWB as the criterion. A second approach, commonly labeled “personality-based” or “disguised purpose” integrity tests, involves standard personality items keyed to predict counterproductive behavior. The earliest article on these measures in the I/O literature was Gough (1971). Despite the “integrity test” label, these measures are designed as predictors of CWB rather than as pure measures of an integrity construct (Becker, 1998).

Sackett and colleagues have published five narrative reviews on integrity testing over three decades, with Sackett and Decker (1979) the earliest and Berry, Sackett, and Wiemann (2007) the most recent. JAP has published several articles that have contributed to the understanding of integrity tests, including Hogan and Hogan (1989) on the measurement of employee reliability, and studies that examined faking (Cunningham, Wong, & Barbee, 1994), subgroup differences (Ones & Viswesvaran, 1998), and difficulties in predicting low base rate behaviors (Murphy, 1987).

Two meta-analyses of integrity test validity have been published in JAP. Ones, Viswesvaran, and Schmidt (1993) is the most cited work in the integrity test field and incorporated a large amount of unpublished work from integrity test publishers. Focusing on predictive studies with non-self-report criteria, Ones et al. reported considerable success in predicting CWB. They also reported relationships with broad job performance measures. This was an attention-getting finding, given that integrity tests had been developed to predict a much narrower portion of the overall criterion space.
Van Iddekinge, Roth, Raymark, and Odle-Dusseau (2012) presented an updated meta-analysis. Their research incorporated some of the primary studies included in Ones et al. (1993), as well as studies published subsequent to their meta-analysis. However, Van Iddekinge et al. used somewhat different inclusion criteria, such as excluding studies that selectively reported statistically significant results and that used extreme-group designs. As such, they did not include many of the primary studies Ones et al. included (and some test publishers did not provide Van Iddekinge et al. access to certain studies that were provided to the previous meta-analysts). Van Iddekinge et al. reported markedly lower mean validity estimates. Sackett and Schmitt (2012) attempted to reconcile the two sets of findings, but found that both meta-analyses contain, to differing degrees, information not available to third parties. Thus, integrity tests are related to CWB, though the strength of the relationship is, at present, under dispute.

It is worth noting that other individual difference measures have also been examined in JAP as predictors of CWB, including biodata (Rosenbaum, 1976), conditional reasoning (Bing et al., 2007; LeBreton, Barksdale, Robin, & James, 2007), and the Big Five personality dimensions (Berry, Ones, & Sackett, 2007). Further, in recent years, there has been a reemergence of scholarly interest in maladaptive traits under the umbrella term of the “dark triad,” namely, narcissism, Machiavellianism, and psychopathy. A meta-analysis by O’Boyle, Forshy, Banks, and McDaniel (2012) found substantial correlations with CWB for narcissism and Machiavellianism. In contrast, correlations between the dark triad traits and job performance were very small, leading O’Boyle et al. to conclude that these traits appear to be related to negative, but not positive, employee behaviors. Thus, the range of constructs under examination as potentially relevant to CWB continues to expand.

In addition, a wider array of approaches has emerged for assessing dark traits, configural scoring of the FFM facets, and omnibus measures of the three traits (Wu & LeBreton, 2011).

Emotional Intelligence

Of all individual differences, in recent years, EI has probably received the most attention among the general public and researchers. Similar to maladaptive traits, however, EI was also covered in early JAP articles (e.g., Pressey & Pressey, 1919; Tendler, 1930). However, this early research lacked a coherent framework to conceptualize EI.

Salovey and Mayer (1990) made important progress toward better understanding the facets of EI. They presented a model that distinguished between various EI-related abilities, such as perceptions of emotions, regulation of emotions, and utilization of emotions. This model is often juxtaposed to a mixed model in which a host of noncognitive factors are thought to reflect EI. Another difference between the two models is that the ability EI model is typically measured via tests, whereas the mixed EI model relies on self- or peer-reports.

The first JAP article on these modern EI conceptualizations was published in 2004 (Law, Wong, & Song, 2004). Since then, JAP has published two EI meta-analyses. Joseph and Newman (2010) examined the validity of EI as conceptualized in the ability model. They found support for a sequential relationship among EI facets (emotion perception, understanding, and regulation) and job performance, with personality and cognitive ability as antecedents of these EI processes. This meta-analysis has become the most cited JAP article in recent years. Then, Joseph, Jin, Newman, and O’Boyle (2015) examined the validity of EI as conceptualized in the mixed model. Although Joseph et al. found a moderate correlation between mixed EI and supervisor-rated job performance, this relationship reduced to zero after controlling for already-established constructs such as ability EI, self-efficacy, the Big Five factors, and cognitive ability. Taken together, these two meta-analyses demonstrate that further progress on EI is to be made via more refined conceptualizations and measurement of the ability EI model.

Motivational Traits

Kanfer, Frese, and Johnson (2017) provide a comprehensive review of JAP’s contributions to work motivation. In this section, we focus more specifically on motivational traits. One challenge with reviewing research in this area is that little consensus exists regarding the construct(s) that best represent individual differences in trait motivation (Diefendorff, 2007; Judge & Ilies, 2002). We focus on traits that reflect, or are thought to influence, choices workers make with respect to goal-directed activities, the amount of effort they expend on those activities, and the duration of time they pursue those activities (Diefendorff, 2007). Further, we focus on stable individual differences and not on more malleable constructs such as goal setting and self-regulatory skills, or on situation-specific constructs such as performance expectancies and specific self-efficacy.

Conscientiousness. Of the Big Five factors, Conscientiousness is thought to be most relevant to work motivation (Chen, Gully, & Eden, 2004; Kanfer, 1992; Schmidt & Hunter, 1992). People who are highly conscientious tend to be achievement oriented, hardworking, and have high expectations of themselves (Barrick, Mount, & Strauss, 1993). However, Conscientiousness is a broad trait that also reflects individual differences in variables, such as dependability and order, which seem less directly relevant to motivation. In addition, correlations between conscientiousness and job performance tend to be modest. Indeed, results of several JAP studies suggest that conscientiousness is a relatively distal antecedent variable that affects performance by influencing more proximal motivational variables. For instance, Gellatly (1996) found that conscientiousness related to performance on a lab study task primarily through goals and expectations concerning performance. Barrick et al. (1993) reported that conscientiousness related to job performance partially through goal setting and goal commitment.

A meta-analysis by Judge and Ilies (2002) examined relations between the Big Five traits and three motivational variables: self-efficacy, expectancies, and goal setting. Results revealed small to modest positive correlations between Conscientiousness and all three variables. However, Neuroticism demonstrated the strongest and most consistent (negative) relations with the other variables. These results appear consistent with models that distinguish between approach-oriented and avoidance-oriented individual differences in motivation. For example, Kanfer and Heggestad (1997) categorized work motivation constructs in terms of achievement—an approach-oriented construct that often is considered a facet of Conscientiousness—versus anxiety—an avoidance-oriented construct related to Neuroticism.
Achievement motivation. Achievement motivation (also referred to as need for achievement or achievement-striving[s]) is often considered the construct that best reflects individual differences in motivation (e.g., Kanfer & Heggestad, 1997). An individual who is achievement oriented “maintains high standards” and “aspires to accomplish difficult tasks” (Jackson, 1974, p. 6). Hermans (1970) developed one of the first measures of achievement motivation that did not use a projective approach such as the Thematic Apperception Test (Murray, 1943). Hermans’s measure is still being used in research today (e.g., Frese, Krauss, et al., 2007). Barrick et al. (2002) developed a measure that assessed accomplishment striving, status striving, and communion striving. The accomplishment and status striving scales (which largely reflect achievement motivation) correlated with job performance and partially mediated the relation between conscientiousness and performance.

Other JAP studies have shown that the achievement strivings dimension of Type A personality relates to academic performance (e.g., Spence et al., 1987; Spence, Pred, & Helmreich, 1989) and job performance (e.g., Bluen, Barling, & Burns, 1990). Phillips and Gully (1997) found that achievement related to academic performance primarily indirectly through self-set goals. Using data from Project A (J. P. Campbell, 1990a, 1990b), Borman, White, Pulakos, and Oppler (1991) found that achievement related to supervisor ratings of job performance via its effect on awards received (see also McCloy et al., 1994). A meta-analysis by Dudley et al. (2006) revealed that achievement related more strongly to task performance than three other facets of conscientiousness (i.e., cautiousness, dependability, and order) and also was a better predictor than global conscientiousness (based on estimates from Hurtz & Donovan [2000]). In contrast, dependability appeared to be somewhat more strongly related to overall job performance, as well as to performance dimensions such as job dedication and CWBs (although many of these results are based on relatively small numbers of primary studies).

Helmreich, Sawin, and Carsrud (1986) showed that relations between work motivation and performance can change over time. Specifically, they found that achievement motivation was not a significant predictor of new hires’ performance during the first few months on the job, but became a significant predictor “after the ‘honeymoon’ period with the work had ended” (p. 187). Similarly, Stewart (1999) found that achievement was more strongly related to performance among long-term workers than was order, whereas the opposite was the case among newly hired employees.

Proactive personality and personal initiative. Two more recent motivational variables are proactive personality and personal initiative. Proactive personality reflects the extent to which individuals “take action to influence their environments” (Crant, 1995, p. 532), and personal initiative reflects the extent to which individuals are self-starters, proactive, and persistent in overcoming barriers (Frese, Garst, & Fay, 2007). Several JAP studies have examined relations between proactive personality and outcomes. For example, Crant found that proactive personality related to job performance after controlling for the Big Five factors, and Seibert, Crant, and Kraimer (1999) found that proactivity correlated with subjective and objective indicators of career success. Major, Turner, and Fletcher (2006) reported that proactive personality predicted motivation to learn (but not developmental activities) beyond the Big Five. Finally, Thompson (2005) found that proactive personality affects job performance, in part, through network-building and initiative-taking behaviors.

Goal orientation. Goal orientation is a construct that originated in the education literature and suggests that individuals hold either a learning or performance orientation toward achievement situations (Bell & Kozlowski, 2002; Phillips & Gully, 1997). Individuals with a mastery or learning goal orientation (LGO) are motivated to increase their competence by acquiring new knowledge or skills, whereas individuals with a performance goal orientation (PGO) are motivated to demonstrate competence compared with others (Dweck, 1986). Further, individuals with a PGO can be motivated by gaining favorable judgments from others (i.e., prove PGO) or by avoiding negative judgments from others (i.e., avoid PGO; Elliot & Harackiewicz, 1996; VandeWalle, 1997).

JAP has published many goal orientation studies in recent years. In what may be the first JAP study in this area, VandeWalle and Cummings (1997) found that LGO was positively related to feedback-seeking behavior, whereas PGO was negatively related to such behavior. Similarly, Colquitt and Simmering (1998) discovered that LGO and PGO were positively and negatively related to motivation to learn, respectively.

Other research suggests that goal orientation affects outcomes through more proximal variables. For example, in the most highly cited JAP article in which goal orientation was a focal variable, Phillips and Gully (1997) found that goal orientation appeared to affect self-set goals and academic performance through specific self-efficacy rather than directly. Similarly, VandeWalle, Cron, and Slocum (2001) found evidence to suggest that specific self-efficacy, goal setting, and effort mediated relations between goal orientation and academic performance. In a lab setting, Ford, Smith, Weissbein, Gully, and Salas (1998) found that LGO related positively to training outcomes through metacognition, whereas PGO related negatively to transfer of training through specific self-efficacy. In an organizational settings, both VandeWalle, Brown, Cron, and Slocum (1999) and Porath and Bateman (2006) found evidence consistent with the idea that LGO relates to job performance through self-regulatory behaviors, such as goal setting and feedback seeking. In contrast, PGO was either unrelated or negatively related to performance. Finally, in a sample of job seekers, Creed, King, Hood, and McKenzie (2009) reported that LGO was positively related to number of job offers, whereas PGO was negatively related to job offers (see also van Hooft & Noordzij, 2009). Further, results were consistent with the idea that self-regulation mediates relations between LGO and job search behaviors.

Finally, JAP has published two major reviews of the goal orientation literature. DeShon and Gillespie (2005) reviewed goal orientation research conducted to date and identified various conceptual and empirical inconsistencies and ambiguities within this literature. In addition, they developed a model—motivation action theory—that attempts to bring clarity to the literature and provide a foundation and organizing framework for future goal orientation research. Payne, Youngcourt, and Beaubien (2007) conducted an

---

1 In contrast to proactive personality, personal initiative appears to be conceptualized and measured more as a set of behaviors than as a disposition. Indeed, several studies have treated personal initiative as a dependent variable rather than as an independent variable (e.g., De Dreu & Nauta, 2009; Frese et al., 2007; Sonnentag, 2003).
empirical review of goal orientation research. Their meta-analysis revealed that LGO was positively and highly correlated with general self-efficacy, and was moderately correlated with need for achievement and Openness to Experience. In contrast, performance goal orientation was negatively and moderately correlated with emotional stability, and avoid performance goal orientation was negatively and strongly correlated with general self-efficacy and moderately correlated with several other traits, including self-esteem and emotional stability. Further, goal orientation was moderately related to more proximal variables (e.g., state goal orientation, specific self-efficacy), but was not a strong predictor of academic or on-the-job performance. These results appear consistent with those described above, which suggests that individual differences in goal orientation exert a more distal influence on performance through more proximal motivational variables.

General self-efficacy and core self-evaluations. General or generalized self-efficacy (GSE) reflects individuals’ perceptions of their ability to perform across various situations (Judge, Erez, & Bono, 1998; Judge, Locke, & Durham, 1997). GSE is considered a relatively stable motivational trait (Chen, Gully, Whiteman, & Kilcullen, 2000), and differs from specific self-efficacy, which is a state-oriented variable that reflects one’s beliefs regarding performance in a particular situation.

Although researchers have tended to focus on specific self-efficacy, several JAP articles have contributed to understanding of GSE. For example, Judge, Locke, Durham, and Kluger (1998) found that GSE related positively to both job satisfaction and life satisfaction. Chen et al. (2000) found evidence consistent with the idea that GSE influences academic performance through specific self-efficacy. Erez and Judge (2001) reported that GSE related positively to job performance, but not beyond the broader CSEs construct of which GSE is thought to be part (more on this below). Foti and Hauenstein (2007) found that GSE related positively to leader emergence and to promotions into such positions, but not to leader performance. DeRue and Morgeson (2007) found that GSE was unrelated to team member performance, but moderated the relationship between person-role fit and performance such that the relation was stronger when GSE was high than when it was low. Finally, a meta-analysis by Judge and Bono (2001) revealed that GSE was positively related to both job satisfaction and job performance. However, the number of GSE studies was small, and the GSE-job performance analyses appeared to include some studies that used self-ratings of performance.

Much of the research on GSE has focused on its role in the broader concept of CSEs, which reflect “assessments that people make about their worthiness, competence, and capabilities” (Judge, Bono, Erez, & Locke, 2005, p. 257). CSEs are thought to represent a latent construct that accounts for the shared variance among GSE and three other constructs: emotional stability, locus of control, and self-esteem. Judge and colleagues have published many highly cited JAP articles on CSE. For example, Judge and Bono’s (2001) meta-analysis revealed that the four CSE traits demonstrate similar relations with job performance. In contrast, relations with job satisfaction were more varied, with GSE being the strongest correlate. Erez and Judge (2001) found empirical support for treating CSE as a higher order factor, which tended to demonstrate stronger relations with effort and job performance than did the average of its lower order traits. The researchers also found evidence to suggest that CSE related to performance directly, as well as indirectly through variables such as goal setting and activity level. Judge, Hurst, and Simon (2009) reported that self-evaluations were positively associated with income and partially mediated the effects of cognitive ability and physical attractiveness on income.

However, other findings have raised questions about the validity of the CSE construct (Chen, 2012). For example, there often is notable variance in the extent to which the lower order traits relate to one another as well as to external variables (e.g., R. E. Johnson, Rosen, & Djurdjevic, 2011). Further, Erez and Judge (2001) reported that in some cases, CSE accounted for only minimal additional variance beyond traits such as locus of control. Relatedly, R. E. Johnson, Rosen, Chang, and Lin (2015) found that the fit of a model that specified CSE as higher order construct was better when locus of control was excluded as an indicator.

Summary

In sum, JAP has published many articles on motivational traits. Table A3 of the online supplemental materials includes some of the key articles in this area. Interestingly, JAP research has tended to focus on broader constructs, such as Conscientiousness and CSE, than on constructs that are assumed to more closely reflect motivation, such as achievement motivation (this general trend is evident in other journals as well). Yet many JAP articles on these broader variables have been highly cited. Finally, with some notable exceptions (e.g., Hermans, 1970), research in JAP has focused more on testing substantive relations between motivation-related traits and other constructs, and less on the measurement of individual differences in motivation.

Creativity

Several themes emerge in work on creativity in work settings. First, although creativity studies in general focus on novelty as the key feature for identifying creative work (e.g., a work of art), in work settings, the focus is jointly on novelty and usefulness as the defining features. Second, there is a distinction between creativity as a set of behaviors or their outcome (typically a dependent variable in research studies), and creativity as an attribute of an individual (typically an independent variable). Third, a distinction is drawn between the generation of ideas and their implementation, with the term “innovation” used to describe success in implementation of creative ideas (e.g., West, 2002). Articles on creativity have appeared sporadically over the decades in JAP. However, Amabile’s (1988) chapter, which presented a broad theory of creativity and innovation, was the main driver of a resurgence of interest in this area. Our interests here are the individual differences component of the theory, which posits three key factors: domain-specific knowledge and skills, creativity-related skills (which include specific abilities, relevant personality factors, and cognitive style), and intrinsic motivation. In terms of research on individual differences as antecedents of creative behavior, our review uncovered five main themes. The first theme focuses on empirical keying of biodata (e.g., Tucker, Cline, & Schmidt, 1967) or personality items (e.g., Gough’s, 1979; “creative personality” scale). Although such studies showed useful prediction of creative outcomes, they shed little psychological insight into attributes of creative behavior.
The second theme involves the use of measures in the cognitive ability domain. For example, Carroll’s (1993) taxonomy identifies idea production as one of the eight second-stratum factors that underlie general cognitive ability (see also the divergent production factor in Guilford’s [1956] structure of intellect model). JAP papers related to this theme include Owens, Schumacher, and Clark (1957).

The third theme is the use of personality constructs to predict creative behaviors (Dilchert, 2008; National Research Council, 2015). At the Big Five level, Openness to Experience emerged as the level of more specific facets, including the dominance facet of Extraversion and the achievement facet of Conscientiousness.

The fourth theme is the use of measures of innovative cognitive style. By far, the most cited article in JAP in the creativity literature is Kirton’s (1976) development of an adaption-innovation scale, which distinguishes between “adaptors,” who prefer to work within existing structures, and “innovators,” who prefer to “break frame.” Factor analyses of the scale produced three factors: originality, efficiency, and conformity.

The fifth theme is the use of measures of intrinsic motivation. Dilchert’s (2008) meta-analysis revealed relationships between intrinsic motivation and supervisor-rated creative behavior. Other motivation-related constructs have also been investigated in relation to creativity, including creative self-efficacy (Tierney & Farmer, 2002), typical intellectual engagement (National Research Council, 2015), and goal setting (Shalley, 1991).

Finally, two highly cited JAP papers examined interactions between individual and situational antecedents of creative behavior. George and Zhou (2001) found that high levels of Openness to Experience resulted in the highest levels of creative behavior in the presence of positive feedback and uncertainty about either job means or ends. High levels of conscientiousness resulted in lower levels of creative behavior in the presence of close supervisory monitoring and low levels of coworker support. Zhou (2003) reported a three-way interaction: individuals low in creative personality were more likely to exhibit creative behavior in the presence of creative coworkers and a low level of supervisory monitoring. In sum, the literature to date would suggest the use of multiple measures from the cognitive ability, personality, cognitive style, and motivation domains to predict creative work behavior.

Vocational Interests

Strong (1951) concluded that “in any attempt to test a man in order to discover what kind of person he is it is evident that an interest test should be included in the battery” (p. 91). Vocational interests reflect individual differences in people’s preferences for certain types of work activities and environments. Interests have been described and measured in various ways, including interests in occupations, job tasks, school subjects, and activities, as well as with respect to personal characteristics and skills. For example, early work focused on people’s interests in particular occupations (e.g., Strong, 1927). However, over the past several decades, most research has focused on the interests attributes from Holland’s (1959, 1973) theory of occupational choice. Holland categorized interests into six main types: realistic, investigative, artistic, social, enterprising, and conventional interests (i.e., RIASEC), which can be used to describe both individuals’ interests and the interests different work environments support.

JAP has a long history of publishing research on interests. We summarize this research according to what appeared to be some major periods and themes of interests research in JAP. Table A4 of the online supplemental materials presents a chronological listing of key JAP articles in this domain.

Early Interests Research (1917–1929)

Some of the very first research on interests was published in JAP during the first decade of the journal. Folsom (1917) appears to have published the first JAP article on interests, finding that interests were one of the main factors that influence occupational choice. Subsequent studies by Freyd (1922) and McNair (1924) reported on some of the first interests measures. Other studies examined the relation between interests and the abilities required to work in occupations aligned with those interests (Fryer, 1927; Uhrbrock, 1926). In summarizing this research, Fryer concluded that individuals often do not possess the abilities the occupations that interest them appear to require. Finally, Strong (1927) reported on the interests of executives using a measure that later would become the Strong Vocational Interest Blank (SVIB). This measure is considered the first standard interests inventory (Fouad, 2007), and many subsequent JAP studies would use or evaluate this measure.

Peak of Interests Research (1930–1979)

Approximately 90% of JAP articles on interests were published in the 50 years from 1930 and 1979. Four main types of studies were published during this period. First, many articles examined issues related to the measurement of interests. For example, more studies developed new interests measures (e.g., Ewens, 1956; Older, 1944; Super & Roper, 1941). The most highly cited article during this period was Holland’s (1958) study on the development and validation of the Vocational Preference Inventory, a measure that used occupational titles for content. Holland suggested that using titles would minimize negative reactions that more personality-oriented measures often create, as well as reduce test takers’ desire to fake. This inventory was the precursor of the Self-Directed Search (Holland, 1994), which has become one of the most well-known and widely used interests inventories.

Other studies examined issues related to the items or scoring of interests inventories. For instance, Perry (1955) compared forced-choice items with the more typical like–dislike items, and Blake (1969) compared items that included pictures with the more typical text-based items. D. P. Campbell, Borgen, Eastes, Johansson, and Peterson (1968) developed a set of basic interest scales based on occupational titles from the SVIB. Also, several studies developed empirical keys to maximize prediction of criteria. For example, Nash (1966) developed and validated a SVIB key for predicting manager job performance. Further, some of the early research on faking was conducted with interests measures. For example, Steinmetz (1932) found that students could fake their responses to the SVIB to be consistent with the interests profile of a particular job. Kirchner (1961) found that job applicants tended to indicate they liked more SVIB items and disliked fewer items compared with current employees.
Second, many JAP articles examined validity or reliability evidence for interests measures. For example, Melville and Frederiksen (1952) found that several SVIB scales correlated significantly with college grade point average. Dyer (1939) reported that interests measured in college were good predictors of individuals’ occupations 10 years later (see also Austin & Hanisch, 1990). Other studies researched relations between interests and on-the-job outcomes. Bills (1938) found that SVIB scores correlated with supervisor ratings of insurance salesmen, and Rosenberg and Izard (1954) reported that scores on the Kuder Preference Record distinguished between Naval cadets who completed or withdrew from training. Kunce (1967) reported that an index based on the SVIB correlated with accidents on the job. Finally, several studies explored the stability of interests over time and generally found them to be quite stable (e.g., Strong, 1951; Thordike, 1949).

Third, studies examined relations between interests and other constructs. For example, several studies found little or no relation between interests and general intelligence (e.g., R. W. Johnson, 1965; Long, 1945; Shultz & Rush, 1942). In contrast, studies found significant relations between interests and constructs such as personality (e.g., Holland, 1960), values (e.g., Super, 1962), and socioeconomic status (Hewer, 1965). Fourth, many studies during this period focused on the vocational interests of particular groups or of individuals in specific jobs or occupations, including tuberculosis patients (Shultz & Rush, 1942), industrial psychology students (Lawshe & Deutsch, 1952), retired YMCA secretaries (Verburg, 1952), and female computer programmers (Perry & Cannon, 1968).

Decline of Interests Research (1980–2008)

The number of interests articles published in JAP began to decline in the 1970s, and from 1980 to 2008, JAP published only seven interests articles. Several factors may have contributed to this decline. First, I/O psychology journals such as JAP tend to focus on outcomes such as job performance and turnover. Although research in JAP and other outlets suggested that interests tend to be good predictors of outcomes such as vocational choice, findings were more mixed for outcomes such as job performance. Second, articles by Guion and Gottier (1965) and others suggested that noncognitive predictors (particularly personality variables) may have limited use for personnel decisions, and this pessimism may have been extended to interests. Third, in their seminal meta-analysis of predictors of job performance, Hunter and Hunter (1984) concluded that the relationship between interests and job performance was small, which may have reinforced pessimism concerning the usefulness of interests and other noncognitive predictors. Finally, in 1971, the Journal of Vocational Behavior was established. This journal (along with Journal of Counseling Psychology, established in 1954) has become the premier outlet for interests-related research. Thus, perhaps interests researchers began to submit their work to journals that focused on vocational and career issues rather than to more traditional I/O journals such as JAP.

Although very few interests articles appeared in JAP from 1980 to 2008, the two most highly cited JAP interests articles were published during this period. Costa, McCrae, and Holland (1984) is by far JAP’s most highly cited interests article. These authors found associations between the RIASEC interests and the NEO (Neuroticism, Extraversion, and Openness to Experience) personality factors. For example, investigative and artistic interests correlated with Openness to Experience, and social interests correlated with Extraversion. Costa et al.’s article provided support for the idea that interests are related to, yet distinct from, individual differences in personality. The second most highly cited JAP article is Lykken, Bouchard, McGue, and Tellegen’s (1993) research on the heritability of interests. The authors measured the interests of twins who were raised apart and twins who were raised together, and found that as much as two thirds of the variance in interests is associated with genetic differences.

Return of Interests Research (2009–present)

More recently, articles on interests have begun to reemerge in JAP. In 2009, Tay and colleagues used item response theory to analyze data from three interests inventories and found that ideal point models better described response processes than did dominance models (Tay, Drasgow, Rounds, & Williams, 2009). Their findings have implications for improving the development and scoring of interests measures, such as obtaining more precise estimates of individuals’ interests and interest profiles.

Two more interests articles were published in 2011. First, Van Iddekinge, Putka, and Campbell (2011) encouraged the field to “reconsider” the use of interests for personnel selection. Using a measure developed for the U.S. military, the authors reported that interests correlated with measures of job knowledge, job performance, and continuance intentions. They also found that interests provided incremental validity beyond cognitive ability and personality variables, and also yielded small to medium subgroup differences (which often favored female and minority individuals). Second, Van Iddekinge, Roth, Putka, and Lanivich (2011) published the first comprehensive interests meta-analysis to include outcomes such as performance and turnover. Results revealed small to moderate corrected correlations between interests and job performance, training performance, and turnover. The meta-analysis also found that validity estimates were larger for interests that were theoretically relevant to the work performed in the target job, as well as for measures designed to predict outcomes in a particular job or vocation than for measures that assess more general attributes such as the RIASEC dimensions.

Finally, two recent studies examined relations between measures of “fit” with respect to interests and outcomes. Le, Robbins, and Westrick (2014) found that interests fit related to college students’ decisions to major in science, technology, engineering, and mathematics versus other majors. Iliescu, Ispas, Sulea, and Ilie (2015) found that interests fit related negatively to CWBs and provided incremental prediction beyond personality and affect.

Discussion

We reviewed 100 years of research on individual differences and their measurement, focusing on JAP’s contribution to this area. In this section, we conclude with some general trends we observed during our review, as well as some thoughts about future directions for individual differences research. A first trend is that the earlier focus in JAP on the measurement of individual differences seems to have shifted toward an emphasis on substantive developments. Compared with other journals in the Organizational
Behavior/Human Resources (OB/HR) domain, this has made JAP a frontrunner in terms of introducing substantive developments in the area of individual differences to the OB/HR field. At the same time, it should also be mentioned that some of these substantive developments were published some years before in journals devoted to individual differences. For example, in the personality and social psychology literature, Fleeson (2001) published the first article on personality states. Minbashian et al. (2010) then relied on this notion in the work context. As another example, Winter, John, Stewart, Klohnen, and Duncan (1998) published a seminal piece on implicit motives, and later on these ideas started to appear in JAP (e.g., Bing et al., 2007; Frost et al., 2007; Lang et al., 2012). These examples show that JAP has remained a journal of applied psychology, although the theory-driven focus and grounding of the papers has substantially increased, and we expect this to continue in the future.

Another trend we observed was an increased diversity in the measurement of individual differences. Consistent with the construct-method distinction (Arthur & Villado, 2008), we are now able to measure a given individual differences construct via a variety of methods. In other words, individual differences are no longer tied to a specific measurement method. For example, personality can be measured via self-reports or other-reports, and with structured interviews, conditional reasoning tests, SJTs, or biodata inventories. An implication is that the choice of a specific measurement method might affect construct saturation, which refers to the degree to which total score variance in a measure reflects intended construct variance (Lubinski & Dawis, 1992). For example, if the choice of particular measurement method adds unwanted cognitive load to a measure designed as noncognitive, construct saturation is reduced.

A crucial question for future research is the extent to which different measurement methods make individual differences scores differentially saturated with constructs (e.g., cognitive ability for SJTs or Extraversion for interviews), thereby adding (intended or unintended) variance. To shed light into this, future research will need to go beyond the measurement method as a holistic entity by considering the effects of separate method factors such as stimulus format, response format, or response instructions on construct saturation (Lievens, De Cote, & Westerveld, 2015, Lievens & Sackett, 2016). Kuncel and Sackett (2014) illustrated this notion by decomposing assessment center variance into various intended and unintended sources. They manipulated the degree to which intended sources dominated other unintended sources, from “nominal dominance” (in which the intended construct is the largest source of variance in a measure, yet accounts for only a modest proportion of the total variance) to “full dominance” (in which the intended construct accounts for more variance than all other sources combined). These ideas could be applied to a range of individual difference domains.

Sackett and Lievens (2008) outlined various ways that the quality of employee selection systems can be improved, which include, among others, measuring new constructs and measuring old constructs better. Progress in the use of individual differences measures in organizational settings is possible via both paths. Given the length of time that individual differences have been studied, one perspective is that most constructs of value have been identified. Yet new insights do emerge on occasion, such as ability-based measures of EI. Although it is hard to anticipate where an insight into a new construct will come from or what that construct might be, it is easier to see a way forward in terms of new ways of measuring existing constructs. One avenue is the broadening of measurement in the personality arena to approaches beyond self-report. As noted earlier, one promising approach is the use of other-reports (Connelly & Ones, 2010). Although research supports their value, work is needed on developing and evaluating systems for incorporating other-reports into operational use in organizations.

Second, there is a growing research base on the use of implicit measures to assess personality. One notion here is that individuals may have limited insight into their thought processes and patterns of behavior, thus limiting the value of self-reports. Another is concern for favorable self-presentation when using self-reports. Implicit approaches to personality measurement include the use of conditional reasoning measures and implicit association tests. In conditional reasoning measures, individuals respond to what they perceive to be a reasoning test, choosing the “correct” response to items; in fact, the responses reflect personality attributes (Bing et al., 2007). Implicit association measures examine differences in response latencies in pairing personality attributes with categories such as “me” versus “others” (Schnabel, Asendorpf, & Greenwald, 2008). This is a growing area of research, and other approaches to implicit assessment are also possible.

A third alternative to self-reports is the measurement of personality via simulations and games. Although work is in the very early stages, we are aware of attempts to design games that present individuals with different courses of action that can shed light on typical patterns of action (National Research Council, 2015). A fourth alternative involves drawing inferences about personality and other individual differences from online information about an individual, such as social media postings. In the era of so-called “big data,” access to enormous amounts of information is possible, and tools for harvesting such data are developing at a rapid rate (National Research Council, 2015). Recent research has begun to evaluate the validity of inferences based on online information (Back et al., 2010), as well as the prospects for use of such information as a predictor of workplace outcomes (Back et al., 2010; McFarland & Ployhart, 2015; Van Iddekinge, Lanivich, Roth, & Junco, 2016). Although it is too early to assess which, if any, of these approaches to individual differences measurement that will supplant or supplement current approaches, we anticipate interesting and productive work will be forthcoming. However, we caution against moving forward quickly with any of these new approaches without the thorough research foundation underlying more established predictors (e.g., reliability, factor analysis, convergent and discriminant validity research).

Finally, we were often surprised to discover early JAP research that addressed issues of interest to researchers today. For example, studies that examined faking detection and prevention began to appear in JAP as early as the 1930s (e.g., Steinmetz, 1932). So, sometimes it seems that we have “reinvented the wheel.” Many of these early studies used simple, yet elegant, methods to investigate the phenomena of interest. We believe reviewing early work in an area often can be quite beneficial, and we encourage researchers to broaden their literature reviews to include early studies. Further, it is important to recognize that even if good ideas do not become part of the mainstream initially, they are likely to resurface and perhaps be better received in the zeitgeist of a different era.
References


Received August 12, 2015
Revision received June 2, 2016
Accepted June 27, 2016

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.