Applied individual differences research is one of the few branches of psychological science that have systematically amassed a coherent body of empirical knowledge withstanding the test of time. Furthermore, it has exemplified how practice can facilitate basic science. While examining relatively stable behavioral attributes (abilities, personality, vocational preferences), which form the bases of their longitudinal forecasts pertaining to broad behavioral patterns and proclivities, differential psychologists appear to have uncovered keys to behavioral development. They also have contributed to a better understanding of human diversity and its many unique ("idiographic") manifestations. This issue of the journal tells a story about this domain of knowledge; its history in facilitating practice in educational, clinical, and counseling psychology; and its role in personnel development and selection in business, industry, and the military. Some ways in which individual differences dimensions converge on other psychological viewpoints (phenomenology, Skinnerian behaviorism) are reflected, as is their bearing on theories of human development more generally. When examining human behavioral phenomena from a psychological point of view, taking a multifaceted (individual differences) approach appears to be critical; this is especially important when designing opportunities and policies for positive development in clinics, in schools, at work, and in society at large.

Applied individual differences research and its quantitative methods has a long and impressive history. Over the years, several names have been offered for this branch of psychology. Binet called it individual psychology, Stern preferred differential psychology, but it was E. L. Thorndike who christened the most frequently used descriptor, the psychology of individual differences. Much of the early work in this area was aimed at assessing behavioral dimensions of human variation—abilities, personality, and vocational preferences—but not just any dimensions. Concentration was restricted to uncovering parameters that carried real-world significance. Dimensions that never traveled outside the laboratory of brass instruments or the factor analyst's Euclidean space were paid little attention. Examining elegant trivia, or what Truman Kelley (1939) called psychological factors-of-no-importance, was shunned. Hugo Munsterberg's well-known substantive appraisal of structuralism ("Yes, it is precise but it is not useful;" cf. Landy, 1992, p. 788) also describes the type of work that was to be avoided in the psychology of individual differences.

Dimensions holding prophecy were what was aimed for—but prophecy for useful things. Thus, validation criteria typically consisted of behaviors observed after large temporal gaps from initial assessments. Because early differential
psychologists working in educational, clinical, industrial, military, and vocational contexts focused on longitudinal forecasts, they were most interested in assessing stable dimensions—relatively life-long behavioral dispositions. Unstable behavioral tendencies would not be helpful. They would not hold prophesy. So attention in the psychology of individual differences became focused on validating measuring instruments (and constructs) capable of manifesting impressive degrees of predictive validity.

We hold a common bond with applied individual differences research and its methods (Hollingworth, 1929; Hull, 1922; Paterson, Schneider, & Williamson, 1938; Viteles, 1932), and we are concerned with: (a) the conservation and optimal development of human resources (Scarr, Dawis, Benbow, & Stanley), (b) quantitative methods for forecasting behavioral tendencies over extended intervals (Hunter & Schmidt, Grove & Meehl), and (c) issues involving the measurement of human psychological attributes—at the individual (Tenopyr) and group (Lubinski & Humphreys) levels. These topics unified early applied differential psychologists; they still do. There is much overlapping between them. Early and present-day investigators of these topics almost always aimed to generate important information for policy research and development. They also hoped to contribute to business, government, and their society’s overall economic well-being. This translation of research findings into informing public policy has been challenging and not always successful. With this in mind, Sechrest and Bootzin have been moved to discuss some difficulties associated with motivating policy makers to attend to solid, empirical findings relevant to policy agendas. They also discuss ways in which psychologists can obtain and present more useful information. Their discussion of mundane realism underscores the need to consider the many important ways in which people differ (i.e., research findings from differential psychology).

Conservation and Development of Human Resources

Assessment tools invented by early differential psychologists were built with the hope of facilitating the optimal development and utilization of human capital (Bingham, 1937; Kelley, 1928; Spearman, 1927; Terman, 1916; Thurstone, 1938). Measures of human behavior were designed to assess current development and also to forecast the projected path of one’s developmental trajectory (provided that certain opportunities are made available). Although, even today, the dimensions and interrelationships within human ability, personality, and vocational preference domains are far from complete adjudication, and tools purporting to measure these constructs undergo continual refinement (Dunnette & Hough, 1990, 1991, 1992), applied psychologists nevertheless routinely find that measures of systematic sources of individual differences are the most predictive of their chief criteria of interest (in learning [Cronbach & Snow, 1977; Snow, 1989], in work [Schmidt & Hunter, 1981; Schmidt, Ones, & Hunter, 1992], and in interpersonal adjustment [Block, 1993; Dahlstrom, 1993; Wiggins, 1988]). In education, counseling, and industrial psychology, applied individual differences researchers consistently have been committed to facilitating human adjustment and development, albeit with focus on somewhat different stages of the developmental process. It is interesting to note that essentially all of the Bingham Lectures (a series devoted to talent development, each published in the American Psychologist) mirrored the aforemen-
tioned developmental sequence (e.g., education [Terman, 1954], counseling [Paterson, 1957] and industrial [Ghiselli, 1963]). Yet what is most impressive about this series (like the contributions of this special issue) is the intersection of their orbits and underlying assumptions about applied psychological research.

A remarkable degree of complementarity is seen in the current issue as well. Dawis’s treatment of vocational psychology builds on earlier formulations stimulated by Parsons (1909)—or perhaps even earlier (Harper, 1905)—and highlights the need to consider constellations of individual differences attributes for vocational counseling. Dawis speaks to the importance of securing a vocationally adjusted work force: the benefits of vocational harmony, the costs of vocational discord. To be sure, Dawis’s treatment is not definitive. Many aspects of vocational behavior need more research. Yet, when reading current treatments of vocational psychology, it is apparent that some are naive of the erudition accumulated over years of systematic inquiry that Dawis brings forth so well. It is ironic that some contemporary writers stress the importance of looking at the “whole person” yet restrict assessment and practice to a limited set of (favorite) relevant factors, sometimes exclusively limited to clients’ self-assessments, rather than taking a more comprehensive, multifaceted-approach, using available tools provided by differential psychology (Dawis, 1992, and herein).

In the same way, for years, applied individual differences psychologists have stressed the importance of assessing conventional ability dimensions (mechanical—spatial, quantitative, verbal) for tailoring optimal educational interventions and opportunities (Paterson, Elliot, Anderson, Toops, & Heidbreder, 1930; Terman, 1956). Benbow and Stanley tell us that seventh-grade students in the top 1% of their age mates in ability routinely assimilate a full year of a high school class (chemistry, Latin, mathematics, physics, etc.) in 3 weeks time. The amount and type of course work for which these talented adolescents can assimilate, can be forecasted from their level and pattern of abilities—assessed with tests initially designed for college-bound high school seniors. This has been documented annually for more than 20 years, and similarly impressive findings about gifted individuals have been reported for decades (Hollingworth, 1927; Seashore, 1922; Witty, 1951). Although they often found their way to premier outlets (through eminent scientists: Hollingworth, 1927; Paterson, 1957; Pressey, 1946a, 1946b, 1967; Seashore, 1922; Terman, 1954), this information is received as news by many contemporary psychologists. Perhaps, given the contemporary standards for what is needed to compete in international markets, these findings will be more widely distributed and assimilated so that what psychological science has to offer may be more fully appreciated. (Revisiting Seashore’s “The Gifted Student and Research” [1922] and his “Gifted Children in the Industries” [1923] might be useful starting points.)

One generalization from Benbow and Stanley’s treatment is that learning potential is unlikely to be realized (among all students) when individual differences in ability is not factored into all facets of educational policy (evaluations and interventions). Indeed, highly gifted individuals, because of their remarkable accomplishments, have been used by educators and differential psychologists—early on (Seashore, 1922) and more recently (Humphreys, 1985)—to underscore the need to take into account individual differences in general. The authors stressed that, in an ideal situation, educational treatments (opportunities) should be tailored toward an individual’s current level of develop-
ment. This perspective is in line with Cronbach’s (1957) formulation of trait-by-treatment interactions, which was stressed as much by early applied differential psychologists as it is by the current contributors.

Since diagnosis is the first step in all effective teaching and training it becomes self-evident that the primary aim of the school can no longer be confined to a “filling station” function in serving the students. Instead, the school must “learn them” in order to teach them. For this reason, each unit in the educational system will require the services of a counselor who is primarily a trained diagnostician. And the counselors effectiveness will depend on the extent to which administrators and teachers adjust their services to the needs of the individual student as disclosed by diagnosis. (Paterson et al., 1938, pp. 300–301)

Years following these remarks, in one of the most scholarly treatments available on the philosophical underpinnings of vocational psychology (after 30 years of utilizing individual differences research to facilitate educational–vocational counseling), Williamson (1965) commented on the lack of scientific professionalism among a number of psychological practitioners—namely, practitioners uniformed by, or disattentive to, the many years of accumulated knowledge about the psychology of individual differences. Educational and vocational counselors were continuing to focus on unreliable sources for advising adolescents and young adults (cf. Brayfield, 1960). It is unfortunate that this trend has continued (Dawis, 1992, & herein). Robyn Dawes’s (1994) recent documentation of practitioners acting without regard to the best of what psychological science has to offer is not new. Dawesian house-of-cards phenomena have colored psychological science and practice for decades.

Those familiar with the writings of E. G. Williamson (1965) and Leona E. Tyler (1953, 1992), two giants in the individual differences approach to facilitating education and vocational development, will find reminiscences within the powerful concatenations across the first three contributions of this special issue: Scarr (on human development), Benbow and Stanley (achieving academic excellence), and Dawis (vocational psychology). Those interested in a topic covered by one will find relevance in all three. The idea of developing a theory of education and counseling to empower students and clients to engage actively in their personal development blends well with these researchers’ emphases on the importance of providing opportunities for development at critical milestones. Williamson (1965) and Tyler (1953, 1992) suggested the same. Educational and vocational counseling is directed toward imparting skills for enabling people to take a more active role in their personal development. “A good theory of counseling would provide for stimulation of the individual to become creative and confidently aggressive in striving to achieve a greater control of his [or her] own development” (Williamson, 1965, p. 164).

This also fits with Scarr’s view of how people select environments (niches)—environments dispositionally congruent with their personal propensities (ability, interest, and personality). Scarr’s treatment helps us understand better why some people feel more comfortable in certain environments rather than others. Learning more about dispositional drives that motivate people to seek out certain environments (while avoiding others), because of the characteristic experience they produce, is in line with what practitioners following the individual differences tradition propounded all along (cf. Paterson et al., 1938; Tyler, 1992; Williamson,
Such differential proclivities are assessed by traditional measures of individual differences (Anastasi, 1958; Tyler, 1965; Willerman, 1979). They also appear to hold scientific significance for achieving a better understanding of human psychological development (Rowe, 1994).

It is intriguing that Scarr’s piece leaves us with a contribution to psychological theory—an interfacing of differential and developmental psychology, a synthetic amalgamation that Scarr has been refining for sometime now (Scarr, 1992, 1993; Scarr & McCartney, 1983). Neither policy makers nor psychologists can give people experiences. Nor can they ensure prescribed developmental paths. At best, they can provide opportunities to have experiences; and, for those whose dispositional propensities have developed to correspond well with such opportunities, characteristic experiences and developmental growth will typically follow.

Optimal development occurs when opportunities are tailored to an individual’s readiness to profit from opportunities. This is something akin to what insightful educators, counselors, and parents frequently do when they surmise that a student, client, or child is not developmentally prepared to profit from a particular setting involving a complex array of novel stimulation. For example, when parents resist, say, sending their children to tour Europe during their mid-teen years, they may do so because, “they wouldn’t appreciate it.” Parents would prefer to send their children after college when the parents believe the kids are more likely to appreciate it. In such scenarios, young adults frequently tell their parents later that they are glad they waited because they knew what to look for and what the things they saw meant. They would not have had these experiences had they not waited to develop more fully. A garden-variety tour of Europe (or any novel situation, opportunity, or learning environment) generates anything but similar experiences among members of a heterogeneous group—that is, when heterogeneity is defined by traditional individual differences dimensions. The historian, the artist, the social scientist all differentially attend to, see, and experience different things. In part, these differential tendencies may be traced to early dispositional antecedents, which served to individuate educational and vocational development down distinctive paths. These differential paths accentuated their differential proclivities to see, hear, and respond to things differentially (through training, a thought-provoking stimulus for one person may be reacted to by another as an aversion).

Several important questions that arise from the previous discussion include, how do elaborate behavioral repertoires actually develop from broad underlying dispositional propensities? How are skills acquired? And at what level of proficiency? Although differential, biological antecedents for developing abilities, personality, and vocational preferences are well documented (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Plomin, 1990), people surely are not born with highly specialized behavioral repertoires. As B. F. Skinner (1969, p. 183) astutely observed in his theoretical writings on behavioral development: “To say that intelligence is inherited is not to say that specific forms of behavior are inherited. . . . What has been selected appears to be . . . greater speed of conditioning and the capacity to maintain a large repertoire without confusion.” As in language-learning, the content of behavior is imparted by our cultural milieu, but broad underlying predispositions make some skills more readily assimilated over various educational, interpersonal, and vocational settings. And beyond this, there are huge individual differences in preferences for many different kinds of
content offered by a culture or subculture. (Scarr discusses this in her contribution, as do Dawis and Benbow and Stanley; for further discussions, see Lubinski & Thompson, 1986; Meehl, 1972, 1986a)

For our purposes, it will suffice to say that much of what is meant by good education, counseling, and parenting is attempting to enhance the sophistication of students’, clients’, and children’s behavioral tendencies so that they may attend and respond effectively to things that would otherwise go unnoticed. Without such formal training, many aspects of life would fail to generate characteristic experiences. They would not be apprehended with profit. They would not make contact with behavior. They would not facilitate positive growth. Good parenting, counseling, and teaching better prepare people for positive growth. Assessing traditional dimensions of individual differences is frequently helpful for ascertaining when someone is ready for a developmental opportunity, to develop more fully, to experience positive growth in order to become a more discerning and sophisticated individual. This also helps in projecting the most profitable avenues for positive growth (following Dawis): for achieving satisfaction (personal fulfillment) and satisfactoriness (competence).

Graduate education involves a similar process: assimilating a domain, manipulating its particulars, expanding its frontiers. There are different ways to do this. Different constellations of personal attributes better prepare some individuals for certain disciplines relative to others. Perhaps the most celebrated bifurcation in intellectual development is C. P. Snow’s (1967) two intellectual cultures: the humanistic and the scientific. Some individuals are well-suited for both intellectual spheres. (Ideally, all psychological practitioners would be.) Snow’s (1967) essay characterizes individual differences within dominant humanistic versus scientific approaches to education, societal problem solving, and orientation to life in general. But, more important, both differential (Humphreys, Lubinski, & Yao, 1993) and experimental (Kimble, 1984) psychologists have provided empirical evidence for the verisimilitude of this distinction. People predominantly humanistic in orientation tend to possess well-developed verbal abilities and preferences for social contact, whereas those predominantly scientific in outlook tend to approach problems with highly developed mathematical and spatial reasoning abilities coupled with preferences for theoretical ideation and working with instrumentation in isolation.

Because contrasting ability–preference patterns motivate these two groups to enter contrasting educational–vocational ecologies, whereby the distinguishing features of their psychological makeup are augmented and further refined, it is not outrageous to suggest that the ability–preference profiles characterizing these two cultures might contribute to their difficulty (at times) in relating to one another. The development of these two cultures has differentially prepared each group—which differed to begin with—to see things somewhat differently, to form different abstractions, to have contrasting experiences when presented with the same problem or situation.

Paul Meehl’s (1954) Clinical Versus Statistical Prediction: A Theoretical Analysis and Review of the Evidence is the stimulus responsible for generating the line of research discussed by Grove and Meehl here. Supporters of clinical prediction routinely generate forecasts of human behavior on the basis of their clinical training and experience, whereas individuals more committed to the
A statistical approach uses optimally weighted regression equations (derived from prior research using similar samples). Once available, the latter approach can be performed by a clerk to derive an estimate of some future behavior (e.g., probably success in a training program, likelihood of attempted suicide, likelihood to commit a violent act).

In his earlier analysis, Meehl (1954, p. 4) listed some honorific adjectives used by supporters of each camp, along with characteristic pejoratives of the other method. It nicely illustrates the point of a humanistic versus a scientific approach to analyzing human behavior. It also possibly serves to highlight why heterogeneous groups composed of humanistic and scientific participants, assembled to address social problems, sometimes have difficulties in generating a unified approach.

For instance, the statistical method is often called [by its supporters]... verifiable, public, objective, reliable... scientific, precise, careful, trustworthy... empirical, mathematical, and sound. Those who dislike the [statistical] method consider it mechanical, atomistic, additive, cut and dried, artificial, unreal, arbitrary, incomplete... static, superficial, rigid, sterile, oversimplified. The clinical method, on the other hand, is labeled by its proponents as dynamic, global, holistic... patterned, organized, rich, deep, genuine, sensitive, sophisticated, real... and understanding. The critics of the clinical method are likely to view it as mystical, transcendent, metaphysical... subjective, unscientific, unreliable, crude, private, unverifiable, primitive... verbalistic, intuitive, and muddleheaded.

It is interesting to note that when the data began to come in indicating the overwhelming superiority for the statistical approach (Meehl, 1986b), many proponents of the clinical orientation did not react with joy. This is probably because their criteria for understanding human behavior were not met. Such criteria were, almost by design, absent from a statistical approach. Yet the statistical approach provided an intellectually satisfactory solution to the problem of predicting human behavior in a variety of applied contexts. But the findings were not unanimously embraced (cf. Grove & Meehl; Meehl, 1986b)—even though some dovetailing with other approaches provided further support. It is what insurance companies have known and used all along: Insurance companies will take expectations based on an actuarial table over any psychologist's clinical impressions of future risk; and the insurance companies will more frequently be correct. Needless to say, supporters of the statistical approach made similar arguments of equivalence of prediction by other disciplines in other settings, and, unlike the clinical prognosticators, they were pleased as the data came in. This example illustrates how similar situations may give rise to different experiences. Why? Perhaps because the supporters of clinical versus statistical prediction approach human behavior with somewhat different intellectual tools and preferences for what constitutes an intellectually satisfying solution to problems involving human behavior.

The psychology of individual differences tells a cogent story about how these two sets of sentiments may have developed in so many people. Empirical, scientifically respectable findings can generate different experiences among sophisticated intellectuals. In the C. P. Snow example (which involves only two groups), criteria involving empirical documentation with quantification are premium for one group; verbal cohesiveness and eloquence dominate the other.
Different constellations of abilities and preferences, refined and sharpened within distinct (esoteric) problem-solving environments, undoubtedly have intensified differences among members of these contrasting orientations. Differential psychology may be helpful in reckoning these contrasting points of view. (It also may be helpful for better understanding the point of view of other groups.) For policy formation and development, it may be important to know about differential orientations of contrasting groups. Often, perhaps typically, groups composed of heterogeneous collections of participants are complementary; but, as in C. P. Snow’s distinction, they may possess a potential for festering into destructive forms of antagonism. Different groups often approach and evaluate solutions to social problems (policies for positive development) in somewhat different ways.

Good counseling, education, and parenting are all processes aimed at facilitating development positively, making people more sophisticated, augmenting their behavioral tendencies, sharpening their capacities to deal with complexities. But different people have different life complexities to deal with. Everyone has differential strengths and liabilities. Some people gravitate toward the humanities, others the sciences, and others embrace athletics, music, politics, the military, or law where somewhat different sets of personal attributes are required to develop behavioral tendencies sufficiently impressive enough to distinguish oneself. Consider your favorite athlete, artist, musician, academician, or military general. The psychology of individual differences holds that their exceptional accomplishments stem from constellations of individual differences continua (to be sure, at the extremes) that have found their way to environments supportive of their expression. We all possess status gradations on these continua. Exceptional performances do not necessarily imply unique qualities. Rather, they likely imply more of certain qualities.

As development unfolds, for better or worse, people seemingly become more active agents in making life choices. Or, put another way, choices become more determined by personal attributes: abilities (strengths and relative weaknesses) and preferences (environments we approach versus those we avoid). People prefer to enter (learn in, work in, socialize in) environments supportive of their abilities and preferences. Once these environments are secured, development does not end. Development is never-ending. Skills are refined for dealing with the demands within these self-selected psychological habitats (educational tracks, occupations, interpersonal relationships). Under such circumstances, people develop more unique behavioral tendencies, idiosyncratic styles, or, following Allport (1937), idiographic behavior patterns. Part of what developing more sophisticated (professional) problem-solving capabilities involves is becoming more multifaceted (more multidimensional, more complex, more refined) and, by necessity, more distinguished from the general population. These ideas are important for policy makers and developers to factor into their work. They appear to be corollaries of a society becoming more specialized (Wolfle, 1960).

Developing this line of reasoning further presents an intriguing conjecture (possibly an inescapable reality) and, if true, another contribution to psychological theory. A bit of phenomenology appears to cut across the human condition (certain aspects of our intraindividual experiences are unique). This phenomenology, however, can be studied scientifically, by quantitatively detailing major dimensions of human variation (because most aspects of experience emanate from a
INTRODUCTION: APPLIED INDIVIDUAL DIFFERENCES RESEARCH

common configural dimensionality generic to humanity). This is in accord with William James (1890), as one can apprehend from his writings. He, too, argued that experiences are not simply given:

Millions of items of the outward order are presented to my senses which never properly enter into my experience. Why? Because they have no interest for me. My experience is what I agree to attend to. Only those items which I notice shape my mind—without selective interest, experience is an utter chaos. (James, 1890, p. 402)

For parents, practitioners, theorists, and policy makers this helps in conceptualizing environments from which different constellations of attributes are most likely to profit (thrive in, seek out), as well as the experiences that dispositionally congruent (adaptive) versus incongruent (maladaptive) environments are likely to engender. This sharpens our capabilities for structuring opportunities for optimal development. But, more than that, the dimensions of human variation build a rich backdrop for better understanding how people build and transform environments for themselves. This analysis sheds light on why people work so hard to transform their physical and social purviews, so as to engender more pleasing experiences, experiences more congenial to their developmental inclinations and personal point of view.

Offered herein are some conceptual tools for better understanding not only our children and others but also ourselves (phenomenologically and otherwise), by appraising ourselves, others, and our interpersonal relationships in the light of what differential psychology has revealed about human diversity and the variegated environments required for its optimal development. As Dawis (1992) has intimated, perhaps the psychology of individual differences will provide our society with an opportunity to develop more cultural empathy. Sometimes, by objectifying the unique nuances of human individuality, the commonalities that we all share become less opaque.

Quantitative Methods for Forecasting Behavioral Tendencies Over Extended Intervals

We now turn to estimating the precision with which we can forecast behaviors and outcomes. Grove and Meehl’s contribution to this issue has been touched on previously. Their piece, like Hunter and Schmidt’s, contains some of the most well-documented empirical generalizations in all of psychology. There are few sounder generalizations in applied psychological science than the superiority of statistical over clinical prediction (Grove & Meehl) and the cross-situational validity generalization of ability–performance predictive validities observed in industrial and vocational settings (Hunter & Schmidt).

Among other things, these works involve predicting behavioral phenomena over extended temporal gaps. Both use meta-analytic strategies for estimating the magnitude of our capability for doing so. Meta-analytic strategies offer a way to integrate independent studies bearing on predictor–criterion, independent–dependent variable covariation. Studies are aggregated onto a common metric to estimate overall effect size (standard deviation units between some variable or set of variables and a criterion). The aim is to gain a purchase on the precise magnitude of functional relationships. Using dimensions of individual differences, applied differential psychologists appear to be able to do much better than most observers initially thought.
There was a time, for example, when industrial and vocational psychologists cut their teeth on the dogma that ability–performance predictive validities were situationally specific, that the utility of a test was highly sensitive to contextual subtleties. Test validities were thought to vary from city to city, state to state, plant to plant, and from group to group. Thus, a leading figure in industrial–organizational psychology wrote in the mid-1960s:

The first and most pervasive generalization that can be made is that jobs within . . . various organizational groupings, as well as the organizational climates in which they may be found, will demonstrate extensive variability. A test or procedure that may be found highly predictive in one situation may, therefore, prove to be no value at all in another apparently similar one. (Guion, 1965, p. 415)

Today, the cross-situational stability of ability–performance forecasts is one of psychology’s most robust empirical findings (and one of the most frequently discussed topics in the Annual Review of Psychology). As meta-analytic strategies developed more fully (Schmidt & Hunter, 1981), most of the random fluctuations in predictive validities were found to be a function of sampling error. For example, most studies in applied psychology consisted of samples of fewer than 100 participants. The standard error of a correlation for \( N = 100 \) is \( 1 \div N^{1/2} \), or \( 1 \div (100)^{1/2} = .10 \). When a study involving 100 participants is conducted and manifests, say, a correlation of \( r_{xy} = .30 \), Fisher’s \( z \)-transformation provides a 95% confidence interval for the true correlation somewhere between .10 and .47! This is not the precision for which one would hope. This explains why validity coefficients tended to fluctuate so wildly in the early literature (Ghiselli, 1966; Guion, 1965).

The inherent instability of correlations (actually, all effect sizes) based on small samples holds for all research findings. In other psychological domains, those involving individual-differences variables as well as other variables, one may wonder how many investigators are engaged in psychological interpretations that essentially reduce to statistical perturbation involving sample fluctuations. Hunter and Schmidt’s recommendation that meta-analytic tools should be a standard component in graduate education for psychologists and consumers of psychological research is sound. Like all powerful analytic tools, however, meta-analytic strategies can be misguided (cf. Snow, 1995). One needs to make sure that apples are being aggregated, when one is interested in generalizing to crates of apples, and that apples, oranges, and pears are being aggregated when one is interested in generalizing to fruit. But meta-analytic strategies are not a methodological fad. They are powerful tools for uncovering the nature and strength of functional relationships. They will be and should be used more in future research. They already have helped researchers to see things that, before their arrival, routinely went unnoticed (the robustness of actuarial prediction and validity generalization).

Measuring Human Psychological Attributes at the Individual and Group Level

Assessing individual differences by building scales through aggregating multiple-choice items has been the preferred medium of differential psychologists.
The impact this instrumentation has had on psychology is akin to the impact of the microscope on biology and the telescope on astronomy (Dawis, 1992).

Since Spearman's (1904) ground-breaking treatise on reliability theory, psychologists have been intrigued by the phenomenon of aggregation—the almost mystical psychometric principle whereby something quite impressive is gleaned by collecting item specimens almost totally saturated with unwanted noise. It was acknowledged that (like an individual's particular response to any given life opportunity) a particular response to any given test item possessed a large noise-to-signal ratio. Items commonly possess 95% construct irrelevant variance. Only a tiny sliver of variance of each item carries genuine information about an attribute under analysis. On an individual basis, such solitary discrete samples of behavior tell us little about people. They tell us even less about what people are likely to do.

Yet, as more items are responded to (like successive opportunities in life), they begin to paint a reliable picture of the individual with respect to the attribute under analysis. When lightly correlated items are added up (much like one's track record in life), the uniqueness associated with each individual instance does not contribute much to the final portrait because, collectively, they share nothing in common. The unique chunks of variance of each item do not coalesce. Each comprises little of the total picture. What does coalesce, however, are the light slivers of communality running through each item (the dominant theme running through each opportunity). These bits of communality pile up. Their influences are successively augmented in a composite (in one's overall track record). The composite's variance consists mostly of the signal that items share. What the items do not share is noise, which aggregation attenuates to a minuscule sliver within the composite. Aggregation turns the large noise-to-signal ratio on its head. The composite is mostly signal even though the individual items are mostly noise. An example might be helpful to make this more concrete.

The Spearman-Brown (Gulliksen, 1950) prophesy formula estimates the overall reliability of a scale (given the number of items and their average intercorrelation). Aggregating, say, 30 items the average intercorrelation of which is .25 (94% noise within each) generates a reliability coefficient having 91% signal, 9% noise. In the words of Bert Green (1978), "Given enough sows' ears, we can indeed make a silk purse" (p. 666). Just as aggregation (of studies sharing a common theme) is the conceptual workhorse for meta-analytic inquiry, aggregation (of items sharing a common theme) is the conceptual underpinning for building scales of systematic sources of individual differences.

The principle of aggregation is given some space because psychometric measures are routinely criticized through semantic dissection at the item level. Particular items are culled from an instrument to illustrate presumed cultural biases, the validity of contrasting distracters, or any number of manifest concerns reflected by the uniqueness of the item under scrutiny. Sometimes these analyses are simply common sense reactions to item content. Criticisms of certain items may be well founded; and indeed, empirical approaches for detecting bias at the item level has generated an important literature (Drazgow & Hulin, 1990). But it is important to keep in mind that from the beginning measurement specialists (including Spearman), agreed that most of the individual differences observed on any given item are completely irrelevant to the psychological attribute under
analysis. A fine-grained analysis would likely uncover all items as systematically biased (with respect to certain socially defined groups). But the question of bias is more cogently addressed at the scale level: Are inferences based on test scores equally valid across variously defined groups? Or are inferences systematically biased in that they over- or underpredict performance?

A given item provides us with no more information about an individual’s normative standing on a psychological attribute than one course reflects 4-year overall GPA. But when aggregation of a number of items (or courses) is systematically sampled, quite a different picture emerges. The uniqueness of test items become successively trivialized as more are added. In the end, we are left with a highly reliable composite reflecting normative standing. We also may have yet another example of why it is important to consult different kinds of expertise when evaluating different mediums for better understanding human behavior. Some differential psychologists have pointed to critics of tests, following a literary criticism of individual items, as a documentary on psychometric ignorance. But literary treatments of psychometric tools (Gardner, 1995) are not the main concern of critics of testing. Most likely, what concerns critics most are group differences on tests used in selection (especially for educational and vocational opportunities).

Forecasting Merit

In the words of a recent APA task force on testing, “The commitment to evaluate people on their own individual merit is central to a democratic society” (Neisser et al., 1996, p. 90). The capacity of psychological tests to find and forecast merit was well-documented by military psychologists in Britain (Vernon, 1947) and the United States (Flanagan, 1947). These findings were extended with profit to academic training programs, about which Pressey (1946a, 1946b) and others (Paterson, 1957; Terman, 1954) argued that training intervals may be trimmed, thereby engendering a cascade of positive psychosocial effects including economic efficiency for personnel and physical resources (Lubinski & Humphreys). Yet, attendant with these findings were group differences on tests, which on analysis reflected group differences in criterion performance.

Counterfactually, if group difference in performance were nonexistent, group differences on the predictors could have been statistically offset by simply adding a constant to predictor tests (McNemar, 1975). This is what many hoped for. But alas, individual differences in test performance mirror individual differences in real-world performance and, in so doing, place on the doorstep of policy makers and developers one of the most formidable conundrums psychological science has ever encountered.

A voluminous psychological and legal literature emerged in response to selection based on tests manifesting group differences—differences that were not biased in terms of their systematic over- or underprediction of groups on relevant performance criteria. These findings were observed in multiple contexts (e.g., Stanley, 1971). They covered behaviors in schools, training opportunities, and the world of work. Yet some writers continue to depict group differences as spurious or as artifacts of a primitive science. This characterization appears to be incorrect. Readers are referred to reports by the National Academy of Sciences (Hartigan & Wigdor, 1989; Wigdor & Garner, 1982) and by earlier (Cleary, Humphreys, ...
INTRODUCTION: APPLIED INDIVIDUAL DIFFERENCES RESEARCH

Kendrick, & Wesman, 1975) and more recent APA task force reports (Neisser et al., 1996).

In a Psychological Review treatment, Gregory Kimble (1994) provided some useful background for approaching the literature on individual differences in human ability (see especially his section on "Anti-Intellectualism Masquerading as Human Sensitivity"). Kimble (1994, p. 257) is correct about the huge jargon-to-substance and feeling-to-thinking ratios that our discipline has generated on politically correct topics and also in his observation that, "How you feel about a finding has no bearing on its truth," something medical practitioners deal with on a daily basis—through the quantification of human phenomena. Group differences on ability and achievement tests are not artifacts, and the psychometric tools that Tenopyr and Lubinski and Humphreys discuss are sophisticated products of a highly developed technology. Asian and Jewish Americans, for example, typically score higher on standardized aptitude and achievement measures, relative to the general population, and these group differences are reflected in outcome criteria such as receiving academic honors, securing professional credentials, and achieving membership in the National Academy of Sciences, just to mention a few. Professionally developed achievement and aptitude tests measure performance-relevant behaviors, which are exceedingly critical to the economic, psychological, and physical well-being of our society (Gottfredson, in press). Yet, it is vitally important to keep in mind that valid inferences based on psychological measurement instruments, no matter how scientifically significant, do not dictate any particular policy agenda.

Policy Is Determined by Values

Empirical findings can, and often should be, consulted before developing and implementing policy, but empirical findings per se do not speak to the goals of society. What psychology can offer is useful information about ways to meet these goals most efficiently and about ways to assess their costs.

Concluding Statement

One of the purposes of Psychology, Public Policy, and Law is to provide an outlet for psychological findings not only for policy makers and legal scholars but also for many kinds of multidisciplinary consumers of psychological research. To be honest, consumers need to be vigilant. There is reason to be critical of much of what passes as psychological science. For useful inoculations in this regard, see Dawes (1994), Lykken (1991), and Meehl (1971, 1978). At the same time, psychologists can be their own worst critics. For methodologists, being critical with alacrity is almost a requirement for their profession.

There are a number of impressive veins of research lining psychology's journals. A communality defined by at least one of them might be ideal for anathematizing Dawesian (1994) house-of-cards phenomena. As the contributors of this issue have documented, within the psychology of individual differences and its quantitative methods exist a number of important concepts and findings, and collectively they have well represented this branch of psychological thought. Practitioners and theoreticians, as well as policy researchers and developers, are all explicitly affected by the material in this set of articles. The authors’ points were anticipated by Walter Dill Scott’s (1920) appraisal of chief concepts in applied
psychology, offered in one of the earliest *Psychological Review* treatments of this topic (and still worth reading): “Possibly the greatest single achievement of members of the American Psychological Association is the establishment of the psychology of individual differences” (p. 85).

It ought to be clear that there are many unprofitable and reoccurring “fads” in psychological research (Dunnette, 1966), that our empirical generalizations frequently manifest a “short half-life” (Cronbach, 1975) and that many of our theories are pathetic in terms of their “scientific respectability” (Meehl, 1978). The contributors of this issue, however, offer some important (and refreshing) exceptions to these problematic trends. For the most part, the products synthesized to illustrate significant psychological findings are accretions from long-established research traditions—lines of inquiry that many modern writers appear to underappreciate or choose to neglect. That much of this material is not standard conceptual background for undergraduate psychology majors points to both a problem and a paradox for our discipline. But when it is not found in graduate training programs in applied psychology, perhaps grounds for scientific malpractice should be entertained? Psychology does have a number of solid empirically established pillars on which a scientific foundation for better understanding human behavior may be built and from which policy makers and developers may enhance their initiatives with more security than had they not consulted this rich body of knowledge: the psychology of individual differences and its quantitative methods.

References


INTRODUCTION: APPLIED INDIVIDUAL DIFFERENCES RESEARCH


Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs (1).*


