1

INTELLIGENCE: A DIVA AND A WORKHORSE

OLIVER WILHELM

RANDALL W. ENGLE

This introduction will try to set the stage for the subsequent chapters. We first discuss a few issues related to the measurement of intelligence before commenting on the understanding of intelligence. We do so because intelligence testing has always been a very pragmatic job. In an attempt to satisfy the demand for ability testing, the development, use, and evaluation of measurement tools has been a focal activity of researchers working in the area. In the early years of empirical psychological research, the measurement of intelligence was so successful, with approaches related to its understanding so far from being satisfactory, that the latter was overly neglected for a long period. Following these historical lines, we postpone more recent efforts toward the understanding of intelligence to the end of this introduction.

THE MEASUREMENT OF INTELLIGENCE

In English, as in other languages, there is no lack of adjectives to describe intelligent activities, attributes, states, and behaviors. A comprehensive and by no means exhaustive list of adjectives includes the following: abstract, accurate, agile, analytic, artistic, astute, attentive, aware, block-headed, brainy, bright, broadminded, bubble-headed, captivating, cerebral, clever, clod, clumsy, complex, confused, contemplative, crafty, creative, cultured, deep, detail-oriented, distractible, ditsy, doddery, doltish, dreamy, dull, flawless, exacting, fanciful, foolhardy, foolish, foresighted, forgetful, fuddled, gifted, hot-headed, ignorant, imaginative, imperceptible, imprudent, impulsive, inattentive, ingenious, innovative, inquisitive, insightful, instinctive, intellectual, intelligent, introspective, intuitive, inventive, jumbled, knowledgeable, learned, logical, madcap, meditative, mental, meticulous, mindless, musical, oafish, original, perceptive, philosophical, plodding, poetic, pondering, pragmatic, precise, prudent, questioning, quick-witted, reasonable, reckless, refined, responsive, scholarly, shallow, sharp-witted, slow, sly, smart, sophisticated, talented, uncreative, uninhibited, unreflective, vigilant,

AUTHORS' NOTE: Correspondence regarding this chapter should be addressed to Oliver Wilhelm, Department of Psychology, Humboldt-University, Rudower Chaussee 18, 12489 Berlin, Germany. E-mail: oliver.wilhelm@rz.hu-berlin.de.

visionary, well-read, wise, and worldly. These are just a 100 of many more adjectives that can be used for self-reports in everyday language as well as in scientific studies of personality to describe some more or less dispositional features of individuals that somehow have something to do with intelligence.

As with many assessment procedures, we can use factor analysis and other multivariate methods to summarize data on such self-reports, and we would likely find various factors we might label motor abilities, precision in working and thinking, knowledge, reasoning, memory, attention, and the like. Similarly, we can collect self-ratings on abilities by asking respondents to provide us with estimates on how well, compared to other respondents, they would do on certain tests. We can help respondents provide more accurate estimates by explaining ability constructs and giving them examples from actual ability tests. We can ask respondents to provide estimates of how many items of an ability test they solved correctly. We can also rely on the confidence with which individuals respond to individual items of ability tests (Pallier et al., 2002; Stankov, 2000; Stankov & Crawford, 1997). We can also ask individuals to provide us with self-reports about their typical intellectual engagement (Ackerman & Goff, 1994; Goff & Ackerman, 1992), need for cognition (Cacioppo, Petty, Feinstein, & Jarvis, 1996), and openness for new ideas (Costa & McCrae, 1992; Saucier, 1992; Trapnell, 1994).

However, empirical data suggest that none of the above approaches provides an acceptable approximation to the actual intelligence level, as measured by standardized ability tests. The correlations of self-reports of intellect, self-estimates of ability, and confidence ratings for ability tests and standardized ability tests are usually disappointingly low. Despite our environment constantly providing us with information about what we can accomplish and what we cannot, humans do poorly when asked to assess how well they do in a broad variety of ability tests. Apparently, asking for preferences in intellectual engagement, self-ratings of ability, or confidence ratings of responses to problems from intelligence tests does not bridge the gap between putative personality and putative ability measures. From a psychometric perspective, this is intuitively plausible. Cronbach (1949) introduced the distinction between measures of typical and maximal behavior. The distinction between both forms of measurement is strongly associated with the content of a measure. We usually assess typical behavior with self-reports of preferences and valences. Maximal behavior, on the other side, is associated with measuring abilities, aptitude, achievement, and proficiencies. A prototypical measure of maximal behavior can be characterized by several features. The assessed person is aware of the performance appraisal and willing to perform as well as he or she can. The standards for evaluating the performance of a person are explicit, and the assessed person has to exert effort to succeed (see Sackett, Zedeck, & Fogli, 1988, for a similar conceptualization). The pervasive distinction between measures of maximal and measures of typical behavior remains despite attempts to assess constructs from one side of the divide with methods from the other side of the divide.

There are other fundamental differences between measures of maximal and typical behavior. Once it is ensured that participants are motivated to do well and have understood what they are asked to do, precision of measurement in the ability domain can be very high. In addition, there is no need to be afraid of faked answers after taking care of routine precautions because you can hardly pretend to be more able than you actually are by doing better than you actually can. On the other hand, the structure of self-report measures, as well as their validity and vulnerability to artifacts such as response bias, is threatened so strongly that they cannot be applied sensibly for many practical purposes.

Apart from the lack of vulnerability to some critical biases, the use of ability, aptitude, and proficiency measures for practically important decisions is justified by the abundant evidence on the validity and utility of such measures. There can be no serious doubt that ability tests provide incrementally useful information that would improve judgment and decision making. Still, ability tests are not used for all decisions in which they could provide useful information. For example, there are important cultural differences in the frequency with which ability measures are used. These differences are not caused by cultural differences in the validity and utility of intelligence measures: Those are quite stable. It is policy to make use of intelligence measures for high-stakes decisions in one culture but not in another (see Schuler, Frier, & Kauffmann, 1993, for a comparison in the personnel selection context). Leaving the decision uninformed about the abilities of those who are the subject of the decision is likely to be less precise, less dependable, subject to more biases, and less accurate. Although it may be that cultural and social norms and implicit policies place more importance on some aspects of selection than on abilities, the utility of information provided by ability tests can be considered a certainty across cultures.

It is frequently pointed out that predictions based on ability tests do not allow perfect prediction, and it is important to realize that this statement is true and is very likely to stay true for a long time to come. Because human behavior is also affected by such constructs as motivation and personality, abilities will never account for all the variance in performance. In addition, to justify using ability measures, one should demonstrate that they improve the prediction of success, that they cannot be replaced by other predictors or combinations thereof, that the costs are sufficiently smaller than the benefits, and that there is no fundamental problem with fairness in practical applications. Perceived fairness of procedures in practical selection contexts might be the cause for differences in the use of psychometric ability measures across cultures. If this were true, there should be a widespread belief in Germany—a country in which ability measures are used rarely for practical selection—that ability measures are unfair if they are used in high-stakes decisions such as admission to higher education or employment. Similarly, in a country with intense and frequent use of ability measures for crucial decisions, such as the United States, there should be a widespread belief that using ability is a fair way to assist such decision procedures. Available evidence suggests that the cultural differences in the evaluation of various selection procedures either from applicants or jobholders are too small (Marcus, 2003; Phillips & Gully, 2002; Steiner & Gilliland, 1996) to account for the huge differences in the use of ability tests found between cultures. In addition, it is of course possible to raise fundamental fairness problems not only in the case of ability tests (Willingham, 1999) but also for other forms of information used in high-stakes decisions affecting the lives of individuals. It is not clear, then, what causes and maintains the differences across countries with respect to the use of ability measures in high-stakes decisions.

A somehow related problem is the frequency with which various selection methods are used and the validity of such measures. Rationally, it is—ceteris paribus—evident that you will want to use the most valid set of predictors for decisions you are in charge of. Now, the knowledge of which (set of) predictor(s) is actually going to be maximally predictive is no secret science. Psychologists around the world have argued and literally thousands of studies demonstrate the point over and over again: You would not want to neglect intelligence as a predictor in almost any selection decision because the validity of the prediction is going to be worse than it would be with consideration of intelligence. Empirically, there is a surprising gap between what should be done and what is done (Lévy-Leboyer, 1994; Schuler et al., 1993). In fact, in most countries, most selection decisions are not made using the best predictors for success.

The most surprising aspect about test use described above seems to be that no one is too surprised, worried, or shocked about those facts. We have not heard of initiatives of "Test my kid" in countries where ability testing is not a default hurdle in access to higher education. Nor have we heard of fan clubs for college admission tests or similar ability tests in countries where ability testing is a default procedure for many critical decisions.

From all the individual differences constructs that have been investigated and established within the past 100 years, the constructs related to intelligence, ability, aptitude, and achievement have a special status. Ability testing is one of the big success stories of psychology. It is not so easy to assign credit for "inventing" intelligence to patriarchs or leaders in a field of intelligence research. There are two famous and well-acknowledged traditions in intelligence research. First, the methodological innovation of using the then rather new correlation coefficient to express the association between two tests was a brilliant idea. Extending this idea to tetrad analysis and later to factor analysis is certainly one of the most crucial developments for psychology as a scientific discipline, and it can be traced to Charles Spearman most prominently. Second, the initial pragmatic use of intelligence measures (Binet, 1903) and, only a few years later, its widespread use as a selection method in World War I are clearly the other side of a unique success story. Both traditions later joined when it came to standardizing the conditions of administration and the comparisons of individual test results with a group of reference subjects, validating tests by various means, mostly correlating some test value with some other test value, and when the demonstration of utility of measures was warranted.

The successful use of ability tests for many practical purposes and the lack of unequivocal guidelines in deriving indicators from a theory are responsible for the manifold of measures available today. The number of distinct tasks thought to elicit responses that allow an evaluation of the observed behavior as more or less intelligent is hard to estimate, but it is high and rising. Carroll's (1993) review of individual differences research in intelligence lists and classifies a good proportion of tasks proposed so far. One apparent problem with almost all of those tasks is that the gap between what the task is supposed to measure and what it actually measures is not bridged by strict derivation of the task. For example, specification of the measurement intention for a task supposed to measure mental speed usually leaves countless degrees of freedom for how exactly the operationalization takes place. Many of the decisions that need to be made before having a task that can be used to assess individuals are likely to affect the individual differences that can be observed. Thus, it is not clear what exactly individual differences on the task reflect. One can then ask, "Speed of what?" With any individual task, there is thus always a problem of identifying exactly what the task actually measures. The approach of using multiple measures for a construct certainly is major progress in addressing the psychometric problems associated with the use of single indicators. However, analyzing the communality of several things, all of which are pretty fuzzy operationalizations of the constructs they are supposed to reflect, is not a panacea to the problem of indetermination raised above. It is that lack of precision in deriving measures that caused the inflation of available tasks. As a result, frequently the psychological interpretation of the meaning of a specific test score takes not much more space than two or three lines. A related problem is that our interpretation of intelligence constructs, as assessed by several intelligence tests, frequently does not go beyond paraphrasing supposed communality of test content. Although the mathematics of factor analysis has developed rapidly and there is decent software widely available, good factor analysis remains an art much more than a technology because the composition and nature of the included variables are so crucial.

The Understanding of Intelligence

The rapid development of statistical methods and the overwhelming demand for the widespread use of intelligence tests caused an important neglect: trying to understand the basic cognitive mechanisms responsible for the individual differences we label *intelligence*. The initial roots of intelligence research are cognitive in nature. Specifically, Ebbinghaus (1896–1897) and, to a lesser degree, Binet and Henry (1898) were fundamentally interested in cognitive processes and their implications. These roots were almost completely neglected in the first eight decades of the 20th century. The rediscovery of cognition in intelligence research and the related developments caused a number of important changes that might not be apparent to laypersons.

The validity of ability assessment has many faces. The most prominent examples come from personnel selection. There are, however, many more areas in which the assessment of abilities is a crucial aspect. Granting access to higher education, college and university admissions, diagnosis of learning disabilities, diagnosis of all forms of dementia, identifying intellectual talent and giftedness, diagnosis of attentional disorders, and diagnosis of mental retardation and specific disorders associated with intellectual problems are just some of the most prominent applications. Corresponding with this manifold of applications is the prominence of mental abilities throughout our lives. Intelligence is a necessary though not sufficient condition for success in school and college. Intelligence is the best predictor for success on the job, and it is associated with income, health, and instrumental activities of daily living, to name just a few of the abundant correlates. However, all these valuable data do not tell us what we do when we behave intelligently, what is required to solve a specific problem, what can be done to do better on some form of problem, what will help to improve some ability, how our brain operates to solve problems, what processes are accomplished while solving intelligence problems, what makes a problem harder than another one, what we can accomplish and why, what makes one individual smart and another one not, and many more questions that directly relate to the cognition involved in ability testing.

Consider a widely used form of an ability test such as matrices items. This form of test has been used for a long time and is highly recommended by proponents of a general factor theory (Jensen, 1998; Spearman, 1938) but can also be subsumed under the construct of fluid intelligence (Cattell, 1963). Spearman (1938) thought that items of this type required eduction of relations and of correlates as pivotal and that the role of other abilities—specifically, prior knowledge—would be minimal. Cattell thought that items of this type required dealing with novel situations, although he had some sympathy for Spearman's notion of eduction. Multivariate analysis unequivocally demonstrates that items of this type assess something that is at the core of established intelligence tests (Carroll, 1993; Marshalek, Lohman, & Snow, 1983; Snow, Kyllonen, & Marshalek, 1984). However, despite its widespread use, many things were unknown about such matrices items a few years ago. It is in more recent efforts that some determinants of difficulty have been established (Primi, 2001). Other work on matrices items has used specific models of item response theory (Embretson, 1995) and provided us with interpretations of two-person parameters (general control processing and working memory capacity) for such measures. Earlier, more experimental efforts (Carpenter, Just, & Shell, 1990) arrived at similar conclusions.

A larger group of researchers took a more construct-oriented approach in their attempts to explain fluid intelligence. Early work on individual differences in working memory demonstrated a strong link between fluid intelligence and working memory (Kyllonen & Christal, 1990). Subsequently, in a mixture of experimental and correlational work, individual differences in working memory have been established as a crucial construct of human intelligence (Conway, Jarrold, Kane, Miyake, & Towse, in press; Miyake & Shah, 1999). This body of research regresses fluid intelligence and other intelligence constructs on functions of working memory and provides interesting evidence about cognitive processes and cognitive resources that are fundamental for human intelligence.

The developments we just sketched for matrices items can be found for many forms of traditional intelligence assessment. Categorical syllogisms or relational inferences, for example, have a long and scattered history in intelligence research, too (Baddeley, 1968; Carter, Kennedy, & Bittner, 1981; Sternberg, 1980; Störing, 1908; Werdelin, 1958). We refrain from elaborating on this interesting work and want to briefly describe two more very contemporary approaches toward the understanding of intelligence. The first approach is through recent work on interference control and behavioral and cognitive inhibition. The second approach is through evidence from neuropsychology and brain imaging.

Beginning with task classes well known in cognitive psychology, individual differences were investigated in an attempt to find new and promising constructs of human abilities. Tasks that cause conflicts in the cognitive system and require cognitive processes to cope with these conflicts have been used in several investigations. The individual differences that can be observed with such tasks can be considered to reflect interference control, cognitive inhibition, and behavioral inhibition (Nigg, 2000). Among these tasks are the stop-signal paradigm, antisaccadic eye movements, switching a task set, the flancker task, the Stroop test, directed forgetting, and proactive interference, to name just a few. The results of investigations analyzing individual differences on such tasks have been disappointing so far. The observed individual differences are not very stable across task classes, or they are not correlated sufficiently with relevant criteria (Friedman & Miyake, 2004; Kramer, Humphrey, Larish, Logan, & Strayer, 1994; Miyake et al., 2000; Salthouse, Atkinson, & Berish, 2003; Shilling, Chetwynd, & Rabbitt, 2002). It is hoped that future investigations will provide hints for communalities among individual differences in coping with cognitive conflicts, as assessed by various task classes. If intelligence research on this topic is successful, we can add an important ability to our set of usual suspects in predicting important criteria.

The second approach focuses on the physiological level. Fundamental changes for intelligence research are likely to emerge from the rapidly developing fields of neuropsychology and brain imaging. Not only are we likely to learn more about the neural basis of general intelligence (Duncan et al., 2000), but there is also neuropsychological evidence about the dissociation of more specific abilities (e.g., Awh et al., 1996) or on how emotion and cognition interact in decision-making tasks (e.g., Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). These new developments will help us to invent more appropriate and correct models of human intelligence and provide an important addition to our arsenal of methods to investigate cognitive processes. These methods are likely to be of little use for our understanding of intelligence unless informed by psychological theory. Without solid psychological theory, there is no way to understand cognitive processes. Just recording physiological data that accompany some cognitive process will not be instrumental. We do need psychological theory and a psychological understanding before we can make sense of some physiological correlate of intelligence behavior. This objection does not make physiological data any less interesting, but we need to be careful and sensible in their interpretation.

It is possible that, in the long run, the appearance of intelligence tests will change substantially. This could be attributed to new computerized forms of test presentations and new forms of test compilation (adaptive, tailored, and the like). More important, however, the appearance of established intelligence tests might change because ordinary intelligence tests are hard to develop, and what they assess is hard to describe. Forms of assessment that are easier to compile (Kyllonen & Irvine, 2002) because determinants of difficulty and other item statistics are known and subject to experimental manipulations might be established. Similarly, taxonomies of human abilities might change substantially over the years because we succeed in developing a more profound understanding of constructs such as fluid intelligence, crystallized intelligence, or mental speed. These efforts are not likely to provide us with tests that beat existing ones hands down. They just replace traditional forms of assessment that had some basic flaws but many applied victories with forms of assessment that are more thoroughly developed, psychologically profound, but as successful as their ancestors. In the context of this work, it will be desirable to do some fine-tuning on constructs that count, by and large, as established because not all details of measurement and interpretation of these constructs are as refined, solid, and profound as they should be. In our attempts to address these and similar issues, we should have clear expectations. It is not very likely that we will improve the predictions achieved through the application of ordinary and available intelligence tests in many applied settings to a large degree through improving the form of traditional assessment. Many outstanding psychologists have put a lot of effort into establishing what will work in a certain applied context and what will not. This is not to say that modifying the appearance of intelligence tests is not worthwhile. We might slightly improve the validity and utility of traditional intelligence tests, we might improve acceptance in fields where intelligence testing was not accepted as useful information in decision making, and we might have sounder instruments that are easier to develop and maintain.

It is also possible that we succeed in developing tests that assess new abilities. There are several attempts to establish new constructs of intelligence, and they are addressed in various chapters of this book. We should be patient and supportive with such efforts because it is truly difficult to satisfy the necessary requirements before a skeptical crowd of scientists is willing to accept that there is a new player in the field of intelligence. To meet the set of requirements, we will need to have close collaborations between cognitive psychology, differential psychology, psychometrics, and possibly additional disciplines.

The various correlates and possible consequences of individual differences in intelligence have provoked many heated debates and intense political controversies. It is important to note that intelligence for its own sake is important enough for what we do, what awareness we develop, what knowledge we compile, and what we understand, discover, and create throughout our lives to justify almost any level of attention devoted to the construct. It is also important to realize that the individual differences in intelligence and their relevance do not simply disappear if we remove attention from them. The opposite can be, and frequently is, true. Hence, there is a need to deal with causes, consequences, and correlates of individual differences in intelligence so that policies regarding how to use such information can be discussed most coherently. The future is likely to create problems with an even higher propensity to trigger fundamental conflicts. What would our advice be when it comes to problems related to screening for genes associated with different levels of intelligence? Would we recommend the use of devices and drugs that augment intelligence? There are many similar questions ahead, and a profound understanding of intelligence is a prerequisite for sensible advice.

There is a substantial irony in that one of the most successful constructs of psychological research resists a consensual definition. Similarly, there is no broad consensus in the scientific community about the conceptualization of intelligence and measurement methods for intelligence testing. This is why we would label intelligence a workhorse and a diva: The construct is extremely useful, but we do not have a proper definition of what it is and what it is not.

Despite this lack of a consensual definition, our understanding of the principles by which human minds operate, develop, or malfunction is rapidly progressing. Cognitive science is advancing at a very fast rate–much faster than the practical realization of what is or could be possible with these new developments. As a consequence, the basic science of cognition and the technology of intelligence testing are not properly connected—indeed, they have only been connected at their very beginning. The rediscovery of cognition in intelligence research is much more than a fashion. Without an informed cognitive understanding of intelligence as a construct, the technology of intelligence testing is going to make little to no progress. On the other side, testing of cognitive abilities is one of the most important applied fields for cognitive psychology, and the neglect of this aspect is likely to be causal for the lack of implications cognitive psychology has in applied settings.

Hence, we would like to encourage psychologists with a more psychometric background to gain some detailed knowledge about the cognitive processes underlying intelligent behavior. Similarly, we would like psychologists with a more cognitive or experimental background to make more use of applied knowledge from psychometric research, including individual-differences methodology. We hope that this book contributes to serve this purpose.

References

Ackerman, P. L., & Goff, M. (1994). Typical intellectual engagement and personality: Reply to Rocklin (1994). *Journal of Educational Psychology, 86,* 150–153.

Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppe, R. A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science, 7,* 25–31.

Baddeley, A. D. (1968). A three minute reasoning test based on grammatical transformation. *Psychonomic Science, 10,* 341–342.

Binet, A. (1903). *L'Étude expérimentale de l'intelligence* [The experimental study of intelligence]. Paris: Schleicher.

Binet, A., & Henry, V. (1898). *La fatigue intellectuelle* [Intellectual tiredness]. Paris: Schleicher.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119,* 197–253.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence measures: A theoretical account of the processing in the Raven Progressive Matrices test. *Psychological Review, 97,* 404–431.

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

Carter, R. C., Kennedy, R. S., & Bittner, A. C. (1981). Grammatical reasoning: A stable performance yardstick. *Human Factors, 23,* 587–591.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1–22.

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. (in press). *Variation in working memory.* Oxford, UK: Oxford University Press.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO personality and five factor inventory professional manual.* Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J. (1949). *Essentials of psychological testing.* New York: Harper & Row.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A neural basis for general intelligence. *Science, 289,* 457–460.

Ebbinghaus, H. (1896–1897). Über eine neue Methode zur Prüfung geistiger Fähigkeiten und ihre Anwendung bei Schulkindern [On a new method for testing mental abilities and its use with schoolchildren]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 13,* 401–457.

Embretson, S. E. (1995). The role of working memory capacity and general control processes in intelligence. *Intelligence, 20,* 169–189.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent variable analysis. *Journal of Experimental Psychology: General, 133,* 101–135.

Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology, 84,* 537–552.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* London: Praeger.

Kramer, A. F., Humphrey, D. G., Larish, J. F., Logan, G. D., & Strayer, D. L. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging, 9,* 491–512.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14,* 389–433.

Kyllonen, P. C., & Irvine, S. H. (2002). *Item generation for test development.* Mahwah, NJ: Lawrence Erlbaum.

Lévy-Leboyer, C. (1994). Selection and assessment in Europe. In H. C. Triandis, M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organisational psychology* (2nd ed., Vol. 4, pp. 173–190). Palo Alto, CA: Consulting Psychologists Press.

Marcus, B. (2003). Attitudes towards personnel selection methods: A partial replication and extension in a German sample. *Applied Psychology: An International Review, 52,* 515–532.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7,* 107–127.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100.

Miyake, A., & Shah, P. (1999). *Models of working memory.* Cambridge, UK: Cambridge University Press.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126,* 220–246.

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., et al. (2002). The role of individual differences in the accuracy of confidence judgments. *Journal of General Psychology, 129,* 257–299.

Phillips, J. M., & Gully, S. M. (2002). Fairness reactions to personnel selection techniques in Singapore and the United States. *International Journal of Human Resource Management, 13,* 1186–1205.

Primi, R. (2001). Complexity of geometric inductive reasoning tasks: Contribution to the understanding of human intelligence. *Intelligence, 30,* 41–70.

Sackett, P. R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximal job performance. *Journal of Applied Psychology, 73,* 482–486.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132,* 566–594.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300,* 1755–1758.

Saucier, G. (1992). Openness versus intellect: Much ado about nothing. *European Journal of Personality, 6,* 381–386.

Schuler, H., Frier, D., & Kauffmann, M. (1993). *Personalauswahl im europäischen Vergleich* [Personal selection in European comparison]. Göttingen: Hogrefe.

Shilling, V. M., Chetwynd, A., & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. *Neuropsychologia, 40,* 605–619.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 47–103). Hillsdale, NJ: Lawrence Erlbaum.

Spearman, C. (1938). Measurement of intelligence. *Scientia, 64,* 75–82.

Stankov, L. (2000). Complexity, metacognition, and fluid intelligence. *Intelligence, 28,* 121–143.

Stankov, L., & Crawford, J. D. (1997). Self-confidence and performance on tests of cognitive abilities. *Intelligence, 25,* 93–109.

Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology, 81,* 134–141.

Sternberg, R. J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General, 109,* 119–159.

Störing, G. (1908). Experimentelle Untersuchungen über einfache Schlussprozesse [Experimental investigations of simple inference processes]. *Archiv für die gesamte Psychologie, 11,* 1–27.

Trapnell, P. D. (1994). Openness versus intellect: A lexical left turn. *European Journal of Personality, 8,* 273–290.

Werdelin, I. (1958). *The mathematical ability: Experimental and factorial studies.* Lund: Gleerups.

Willingham, W. W. (1999). A systematic view of test fairness. In S. J. Messick (Ed.), *Assessment in higher education: Issues of access, quality, student development, and public policy* (pp. 213–242). Mahwah, NJ: Lawrence Erlbaum.