Carroll, J. B. (1983). The difficulty of a test and its factor composition revisited. In H. Wainer & S. Messick (Eds.), *Principals of modern psychological measurement: A Festschrift in honor of Frederic M. Lord* (pp. 257–283). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Carroll, J. B. (1988). Individual differences in cognitive functioning. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., Vol. 2, pp. 813–862). New York: Wiley.

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

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York: Irvington.

Guttman, L. (1954). A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258–348, 430–433). Glencoe, IL: Free Press.

Humphreys, L. G. (1981). The primary mental ability. In M. F. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 87–102). New York: Plenum.

Humphreys, L. G. (1982). The hierarchical factor model and general intelligence. In N. Hirschberg & L. G. Humphreys (Eds.), *Multivariate applications in the social sciences* (pp. 223–240). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

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

Roff, M. (1941). A statistical study of the development of intelligence test performance. *Journal of Psychology, 11,* 371–386.

Education, Achievement, and General Intelligence: What Ever Happened to the *Psycho* in *Psychometrics?*

Stephen J. Ceci
Department of Human Development and Family Studies
Cornell University

Lloyd Humphreys has long been acknowledged as a pioneer in the field of measurement theory and one of the most thoughtful (and thought-provoking) scholars in the study of individual differences in intelligence. Nothing in his target article or in my commentary will alter this assessment. He is once again blazing trails and offering cogent evidence in support of his arguments. I want to make these positive statements about Humphreys at the outset, lest they get lost amid the questions and quarrels that follow.

My quarrels are directed at the research tradition that Humphreys represents, rather than at him specifically. This tradition is, of course, the established psychometric approach. It is one that seems to have forsaken the *psycho* for the *metric.* I find less and less that is psychological about modern psychometrics. I say this while acknowledging the great strides that its proponents have made on thorny measurement issues. Although the *psycho* depends on the *metric* (i.e., reliability and standard errors of estimate are useful in establishing theoretical validities), the former is not the ineluctable result of the latter. A psychologically informed theory requires going beyond correlations; description and explanation can be fundamentally disjunctive enterprises. Just as correlations are useful to test a model, a model is useful in deciding which correlations to test. Increasingly, I find the models tested by psychometricians to be "wooden."

For Humphreys, intelligence is the repertoire of all intellectual skills and knowledge available to a person at a particular point in time. This is a reasonable start on the road to operationalization and explanation. The problem is that, much like the drunk who loses the keys while getting out of the car but searches for them under the street light because the illumination is better there, Humphreys looks in the wrong places for clues about the nature of the repertoire that he equates with general intelligence (*g*). Specifically, he looks at performance on well-fashioned test items from popular IQ tests and psychometric batteries. I argue that these instruments are impoverished measures of the contents of the repertoire and that there is reason for doubting that they truly assess what Humphreys presumes they collectively assess—namely, *g*. I also argue that this repertoire is indistinguishable from achievement, and there is no compelling evidence to prod one to accept the baggage that usually accompanies conceptions of *g*. Let me give some concrete examples of what I mean before arguing that *g* is more illusory than real, notwithstanding its robust psychometric foundations.

Heritability

Some of Humphreys's argument centers on the heritability of intelligence. He correctly comments that "increasing equality of opportunity ... increases heritability" and that, "if heritability ... is lower today than it was a generation ago, ... liberals have no cause for rejoicing." The basis for Humphreys's assertion, which may be counterintuitive to some, is that, by reducing environmental sources of variation in a population, one

is left with disproportionately more genetic variance in the phenotypic outcome, and thus higher h^2 (the variance in a trait that is uniquely the result of genetics). The logic behind his argument is impeccable, but I wish Humphreys had pursued this line of reasoning further. It may be that, although heritability is increased when opportunities are increased, this nevertheless may result in actual reductions in the absolute magnitude of individual and group differences. This seems possible if environmental resources that had been missing from the lives of some are now provided, thus bringing to fruition heretofore absent genetic potential. Hence, although h^2 may go up, differences between people might actually narrow. Elsewhere, my colleague and I (Bronfenbrenner & Ceci, 1993) argued that this can happen with cognitive outcomes. We suggested that certain types of processes are responsible for translating genotypes into phenotypes and simultaneously reducing the size of differences among people, so that any remaining differences, small though they may be, are mostly genetic in origin. A consequence of this line of reasoning is that heritability becomes less important even as it gets larger; h^2 becomes uninteresting if it is unyoked from mean differences.

This relates to another issue—namely, the meaning of the concept of heritability itself. Not only is it highly relativistic, as Humphreys notes, but one can engineer far higher as well as far lower estimates of h^2, depending on the environmental resources available to individuals (Bronfenbrenner & Ceci, in press). Hence, I am unimpressed by figures such as the one Humphreys cites (.8 heritability for intelligence) because they do not imply that 80% of the variance in the capacity for intelligence is uniquely genetic; rather, they merely imply that 80% of the actualized variability includes a genetic component. This distinction strikes me as an important codicil to Humphreys's assertions, because it makes explicit something that is all too often ignored—the h^2 for intelligence tells us nothing about the degree of unactualized genetic potential available to individuals. Heritability analyses are forever locked in time and place, reflecting only that portion of genetic potential that has already been actualized by the conditions existing in the lives of individuals; they tell us nothing about the size of the residual unactualized potential or about how much competence might develop if societal conditions were to change. The question, therefore, ought not be "How heritable is intelligence?" but "How intelligent is heritability?" Humphreys would probably agree with me on this, because it is consistent with statements he has made elsewhere, which show a sensitivity to environmental influences and an intolerance of rigid genetic determinism (Humphreys, 1989).

Hence, although Humphreys is right to pose the important questions of whether the phenotype can be modified by the environment, by how much, how quickly, and at what point in development, these questions are not elucidated by heritability analyses. Humphreys appears to acknowledge this in focusing his argument on the measurement of phenotypes rather than genotypes, but this point gets lost in his later remarks about heritability.

The Problem of Low Cross-Task Correlations

An elaborate construct-validation effort has been underway for decades to define *g* in psychometric terms (e.g., first principle component) and to link it to other constructs (e.g., heritability, speed of processing, central-nerve conductance velocity) as well as to demographic correlates (e.g., racial and ethnic patterns of occupational success, Black–White differences on achievement). In Humphreys's case, the form this validation takes can be inferred from the fact that carefully selected items have positive loadings on the general factor derived from a correlation matrix. This factor has known characteristics and correlates with academic achievement and job success. To Humphreys's credit, he avoids getting sidetracked in the search for the Holy Grail (i.e., "real intelligence") and uses the correlational evidence only to operationalize his repertoire. But operationalization does not illuminate the nature of the repertoire, and tackling this latter task may ensnarl even one as clever and as cautious as Humphreys, should he take the bait and attempt to make inferences about the nature of intelligence. Here is why.

It is hard to reconcile findings from cognitive psychology with the received psychometric wisdom. In the former field, there is ample support for the contextually-dependent nature of intelligence. Study after study shows low cross-task correlations on what are ostensibly the identical cognitive operations in different contexts. Analogical reasoning, syllogistic reasoning, and even microlevel forms of cognitive processing (e.g., encoding speed) all seem to operate more or less efficiently, depending on the nature of the mental representations on which they operate (Ceci, 1990). What is needed is some account of why psychometric researchers are enthralled with the discovery of stable patterns of correlations among the tasks they use, whereas cognitive researchers are routinely impressed by the inconsistency of performance across the tasks and contexts they use (e.g., Johnson-Laird & Byrne, 1991; Reeves & Weisberg, 1993). It is not, to extend Humphreys's metaphor, simply that one is trying to hit a moving target (i.e., trait growth over time) but that different guns are being fired from different angles.

This disjunction of perspectives in the psychometric and cognitive communities inspires reconsideration of the nature of *g*. A contextualized test battery may well

still yield a reliable first principle component, but my hunch is that the general factor loading will be shown to be confined to certain types of tasks and to certain types of individuals. And, it is very likely that the types of tasks saturated with general factor variance are exactly those closest in content to those explicitly or implicitly taught in schools and used by factorists in their batteries (Ceci, 1991). Humphreys's centroid space appears to be primarily composed of verbal/academic content and almost certainly relies on formal schooling to do well. Of course, for Humphreys this is not a problem, because he defines the repertoire in behavioral terms and seems to avoid attaching any definite ascriptions to "real intelligence." I find this acceptable, but it leads me to wonder why he does not just call it "achievement" and avoid the terminological morass that inheres in the concept of *g*.

If *g* is reflected in the integratedness of the cognitive profile, then the test items answered by the lowest functioning 5% of society will possess the highest loadings on the common factor—that is, so-called general intelligence (Detterman, 1991; Detterman & Daniel, 1989; Detterman & Persanyi, 1990).[1] Although this argument may appear counterintuitive on the surface, given the positive connotations associated with *g*, the extent of intercorrelatedness of scores (hence, the size of *g*) is actually greatest among those who are least successful on *g*-loaded tests such as IQ. Humphreys knows this as well as anyone, but he prefers to reverse what for me is the most parsimonious account of the link between schooling and intelligence (or repertoire). Although so-called general intelligence increases with schooling, higher levels of *g* do not dictate more schooling but instead result from it. Humphreys sees the conundrum, and he eschews the "loose thinking" that has been associated with *g*, preferring instead his notion of a repertoire and its behavioral/phenotypic expression. But I would argue that very little of Humphreys's argument would change if the unit of analysis was shifted from scores derived from so-called aptitude tests to scores from off-the-rack achievement tests. Humphreys reports that the relation between a composite achievement score and an IQ score is about as high as one could expect between two scores on the same IQ test, given its level of reliability. Thus, there is nothing mysterious or unponderable about *g;* it is the residue of what is taught in school. So-called general intelligence may be far less responsible for attainments in school, work, and life than its proponents claim. Rather, Humphreys's *g* may be the result of these attainments. To probe the causal pathways between *g* and attainments, one must go beyond Humphreys's conception of the repertoire and make ascriptions that he seems (wisely, I believe) unwilling to make. I think that, if he ventured further from his behavioral/pragmatic stance, he'd soon be ensnarled in this morass.

[1]Detterman and his colleagues have shown that the size of the first principle component is far larger among the lowest aptitude individuals, leading one to suggest that so-called general intelligence should be renamed "general stupidity," because it is most apparent among those with the lowest IQ scores. Among individuals with average to superior IQs, the size of *g* is far smaller; these individuals are more differentiated and exhibit greater test scatter. I think this calls into question the meaning of the central area of the radex, because this is where generality is greatest.

Intelligence as a Predictor

Humphreys asserts that *g* is an important predictor of societal outcomes, such as performance in school, jobs, and the military. He opines that an essential test of his theory is whether the more educated recruits during World War II were more able to acquire the cognitive knowledge and skills that their military occupational specialties called for than were their World War I counterparts. However, the existing data lead me to wonder how much of the relation between Humphreys's predictor and the criterion is the result of "repertoire overlap" and how much is epiphenomenal. As already noted, one such epiphenomenon may be that the predictor and criterion both reflect school-related achievement and are most parsimoniously regarded as just that. After all, why not say that children who are exposed to academic curricula will exhibit more change in their academic skills than others rather than claiming that such exposure fosters greater gains in their intelligence? Humphreys's steadfast insistence on the repertoire does not explain how individuals can behave more complexly in some domains than in others. I believe that we stand to gain greater scientific yield if we call the repertoire what it is—namely, academic achievement and acquired wisdom. This would help explain why individuals with low levels of academic achievement can behave complexly in certain environments, if the test is not "rigged" by requiring academic knowledge. Conversely, it allows for the common observation that individuals with high levels of schooling (and high IQs) often do not behave complexly (Ceci, 1990, in press; Leshowitz, 1989).

As a case in point, often job performance is predicted by *g* scores quite well when they are tested by paper and pencil (Gottfredson, 1986; Hunter, 1983); when tested by supervisor ratings, work products, or rates of completion, however, predictability is reduced. And, when familial social origins are taken into consideration, there is some evidence that *g* does not predict economic success, even among individuals in the data set that Humphreys bases many of his conclusions on—namely, the Project Talent study (Henderson & Ceci, 1992). My colleagues and I have been persuaded that a family's social and economic resources are re-

sponsible for their offsprings' earnings 11 to 15 years after high school. General intelligence, although showing strong direct pathways to earnings, provides no net effect after the family's social origins are taken into consideration. I do not wish to press these findings terribly hard, in view of the fact that most researchers have argued the opposite, but these findings strike me as worthy of consideration, given the extensiveness of both our sample and the mathematical and statistical models we used (Henderson & Ceci, 1992).

Aptitude testing in the military provides another window for examining Humphreys's predictions. There is evidence that the aptitude tests administered to recruits—the four subtests of the Armed Services Vocational and Aptitude Battery (ASVAB) that are combined to form the Armed Forces Qualification Test (viz., Arithmetical Reasoning, Word Knowledge, Numerical Operations, and Paragraph Comprehension)—are positively correlated with a wide range of success on both basic training and later occupational performance. On this, both Humphreys and I would agree. In the late 1970s, however, there were two massive errors in the calibration of the ASVAB. Due to these inadvertent miscalibrations, more than 200,000 recruits with aptitude scores that ordinarily would have rendered them ineligible for military service were admitted. This occurred without their supervisor's knowledge that they actually had very low aptitude scores. Stitch (1991) reported that these individuals performed slightly worse than higher aptitude recruits on paper-and-pencil tests of job knowledge but that they performed comparably to their peers when supervisor ratings, rates of reenlistment, attrition, and job complexity were used as dependent measures, prompting the Defense Department's Director of Accession to conclude:

> Upon looking at their performance, we learned that a surprisingly large number of them became successful members of the military. If the enlistment standards were working properly, those young people should have been marginal performers at best. As it turned out, not only did they not have marginal performance, many of them performed considerably above that level. ... So, the question was not that training grades were somehow flawed, but that a quarter of a million people who did not meet the enlistment standards and should not have been able to do the job did in fact do it pretty well. (Sellman, 1987, p. 420, cited in Stitch, 1991)

In essence, my quarrel with Humphreys is over the meaning that he attaches to his concept of repertoire and how it is sampled. Is it anything more than the accumulation of school-related knowledge and modes of cognizing (e.g., preference for taxonomic sorting) that get reflected on paper-and-pencil testing, or is it reflective of an underlying mental resource pool (e.g., central-nerve conductance oscillation and velocity of processing) that mediates the acquisition of such knowledge (e.g., Jensen, in press)? The answer to this question requires an explicit theoretical model and cannot be inferred from the correlations that Humphreys reports.

As long as an intellectual phenotype is used to select America's elites in education, the military, and industry, it behooves proponents of Humphreys's view to probe more deeply into the nature of the repertoire than is possible with off-the-rack measures like IQ scores. Without such probing, we shall never know when and how low-scoring individuals are able to perform complex jobs, such as those reported by my colleague and me for gamblers (Ceci & Liker, 1986) and by Stitch (1991) for electronics technicians in the air force. Humphreys's claim that complexity resides at the center of the radex, independent of content, does not explain how or why such individuals can perform such complex tasks. They are no more rehearsed on these tasks than are high-aptitude individuals, but they nevertheless demonstrate a fascinating level of complexity that is sometimes absent from those with higher aptitudes and comparable experience (Ceci & Ruiz, 1992). There is no correlation between their IQ and the factors that enter into the measures of complexity (i.e., the number of independent variables they consider interactively). That even a handful of such cases can be found requires psychometricians to rethink their assumptions about *g*.

Notes

Portions of this commentary were supported by National Institute of Child Health and Human Development Grants KO4HD00801 and RO1HD22839A.

Stephen J. Ceci, Department of Human Development and Family Studies, Cornell University, Ithaca, NY 14853.

References

Bronfenbrenner, U., & Ceci, S. J. (1993). Heredity, environment, and the question "how"?: A first approximation. In R. Plomin & G. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 161–181). Washington, DC: American Psychological Association.

Bronfenbrenner, U., & Ceci, S. J. (in press). Nature–nurture in developmental perspective: The question 'how.' *Psychological Review.*

Ceci, S. J. (1990). *On intelligence ... more or less: A bioecological treatise on intellectual development* (Century Psychology Series). Englewood Cliffs, NJ: Prentice-Hall.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components?: A reassessment of the evidence. *Developmental Psychology, 27,* 703–722.

Ceci, S. J. (in press). The now-you-see-it-now-you-don't quality of everyday intelligence. In H. Roselli (Ed.), *The Edythe Bush Symposium on Intelligence.* Orlando, FL: Academic.

Ceci, S. J., & Liker, J. (1986). A day at the races: IQ, expertise, and

cognitive complexity. *Journal of Experimental Psychology: General, 115,* 255–266.

Ceci, S. J., & Ruiz, A. (1992). Transfer, abstractness, and intelligence. In D. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 168–191). Norwood, NJ: Ablex.

Detterman, D. K. (1991). Reply to Dear and Pagliari: Is *g* intelligence or stupidity? *Intelligence, 15,* 251–255.

Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence, 13,* 349–359.

Detterman, D. K., & Persanyi, M. (1990, April). *Mental tests correlate highest in low IQ groups: Evidence from KABC.* Paper presented at the meeting of the American Educational Research Association, Boston.

Gottfredson, L. S. (1986). Societal consequences of the *g* factor in employment. *Journal of Vocational Behavior, 29,* 379–410.

Henderson, C. R., & Ceci, S. J. (1992). Is it better to be born rich or smart?: A bioecological analysis. In K. R. Billingsley, H. U. Brown, & E. Derohanes (Eds.), *Scientific excellence in supercomputing: The 1990 IBM Contest Prize papers* (pp. 705–751). Athens: University of Georgia Press.

Humphreys, L. (1989). Commentary. In R. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pp. 29–73). Urbana: University of Illinois Press.

Hunter, J. (1983). *The dimensionality of the General Aptitude Test Battery and the dominance of general factors over specific factors in the prediction of job performance in the U.S. Employment Service* (Uses Test Research Rep. No. 44). Washington, DC: U.S. Department of Labor, Employment, and Training Administration, Division of Counselling and Test Development.

Jensen, A. R. (in press). Central conductance velocity and oscillation. *Current Directions.*

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Leshowitz, B. (1989). It's time we did something about scientific illiteracy. *American Psychologist, 44,* 1159–1160.

Reeves, L. M., & Weisberg, R. (1993). Abstract versus concrete information as the basis for transfer in problem solving: Comment on Fong & Nisbett (1991). *Journal of Experimental Psychology: General, 122,* 125–128.

Stitch, T. G. (1991). Military testing and public policy: Selected studies of lower aptitude personnel. In B. R. Clifford & L. Wing (Eds.), *Test policy in defense: Lessons from the military for education, training, and employment* (pp. 1–76). Boston: Kluwer.

Toward an Intelligent View of Intelligence

Douglas K. Detterman
Case Western Reserve University

Lloyd Humphreys presents a summary of the clear thinking that has made his career a distinguished one. He outlines an empiricist's theory of intelligence. He also sketches the social implications of his theory of intelligence. I find very little to argue about in his presentation. In fact, I don't think there is much that can be argued about.

I don't think Humphreys's position is behaviorist. I would call it *Dust Bowl empiricism.* Dust Bowl empiricism is a kind of Midwestern functionalism with a strong fondness for data. Although some might feel this is a pejorative characterization of Humphreys's theory, I consider it a compliment. Humphreys relates intelligence to what can be seen (the phenotype) and places a heavy emphasis on measurement. He logically extends this theory to its social implications.

Although I find very little in Humphreys's position to disagree with, there are several points I would like to emphasize or expand upon to indicate how important they are. Humphreys's presentation is extremely concise, almost telegraphic. Some of the implications of the important points he makes may not be obvious.

Importance of Intelligence

The world is rapidly becoming a global community. This observation has been so widely trumpeted in the press, it should come as no shock to anyone. An important factor leading us into a global community is international trade. Large international corporations are establishing a world economy. It could once be said that "what is good for General Motors is good for the country." But that is no longer true. What is more true now is that what is good for General Motors is good for the world. Large corporations are significantly affected by international events. Recession in Europe is a significant problem for large corporations, like General Motors, with significant exposure there.

The important battles of the future will not be fought with armies, and they will not be won by conquering territory. Wars of the future will be fought by international corporations for access to markets. The winners will be decided on the basis of market share and profits. Fortunately, these wars will produce more winners than losers. Consumers will win worldwide with lower prices for products and services.

The pressure on these corporations will be to become ever more productive. To stay in business in the global economy, companies will have to produce more product at lower prices. Two ways to do this are to use lower priced labor and to use automation. Lower priced labor is only a temporary solution to the productivity problem. If the world truly becomes a global community, labor prices will quickly be-