Giving *g* a Fair Chance: How to Define Intelligence, Survive Falsification, and Resist Behaviorism

James R. Flynn
Department of Political Studies
University of Otago

Humphreys offers an ingenious revision of a *g*-centered theory of intelligence. If it generates measures of cognitive abilities, or combinations of these, that prove to be better predictors of academic and occupational achievement, it deserves to prevail, at least for a while. However, Humphreys fails to confront some of the relevant evidence, prefaces his theory with a circular definition of intelligence, and links his theory to behaviorism—which means to some bad advice about scientific method. It would be a pity if these things, all of which are extrinsic to the theory, were to prejudice its reception. Therefore, I attempt to give the theory a fair hearing by stripping away these extraneous factors.

Humphreys and Behaviorism

Humphreys calls himself a "pragmatic behaviorist" and says that the task of psychology "is to explain and predict behavior, with or without hypothetical mental constructs." He notes that intelligence as a phenotypic trait is "an observable characteristic" and, therefore, that it can be measured and correlated with other traits. He believes that measurement, explanation, and prediction are central to his task—that is, of formulating a theory of intelligence. Other phenomena, such as motivation, learning, and understanding how people solve problems, are legitimate subjects for psychology but are peripheral to his task. He makes common-sense qualifications: For example, motivation has some relevance because, without motivation, the behavior to be measured would not be produced.

Skinner (1972) told us that behaviorism entails the rejection of "mentalism." This does not mean eschewing mental constructs in theory building: Without these, it is hard to see how we could understand anything. Humphreys recognizes this when he defends his core construct *g* "even though it cannot be observed under a microscope." What must be rejected, according to Skinner, is positing mental events as causes. Mental events cannot be observed, and the behaviorist strategy for social science research is sticking to what can be observed. To explain why pigeons hold their heads high, we cite our observations that, whenever they held their heads level, they got no food pellet and whenever they held their heads high, they did. We do not cite the mental event that these pigeons "felt" hungry enough to "want" food. That is better included in the causal analysis by describing a feeding schedule that induced observable weight loss. Skinner did not dogmatically deny that mental events could have causal potency. He merely argued that social scientists would find that inconvenient, precisely because mental events are not directly observable. He suggested that we adopt his research strategy and see how far we can get. Social science might enter a new era of progress so fruitful that no one will miss mentalism or want to resurrect it.

I assume that Humphreys rejects mentalism, despite no explicit statement to that effect. He shows a preference for the observable—his emphasis being on intelligence as an observable characteristic, such as height or length—and it is hard to see why anyone would call himself a behaviorist unless he did. I argue that linking his theory to behaviorism is a mistake. The link is artificial, in the sense that the theory's content owes nothing to behaviorism, and prejudicial, in the sense that anyone who thought the link functional might assume that the defects of behaviorism infect the theory.

How to Resist Behaviorism

Behaviorism attracts social scientists by promising to make social science more scientific; actually, it just makes it more difficult to do. The brute fact is that mental events do operate as causes, and pretending that they do not makes the task of explanation more difficult than it has to be. One example should suffice. We observe a woman laying tiles that make a repetitive, 9-square-foot design on her hall floor. As she nears the end of her task, she slams down her hammer, screams curses, and throws the tile in her hand against the wall. Perhaps running a film of her previous life would lead to an explanation from observable behavior, but asking her what happened might be simpler. Assume she says: "I am such a fool! The floor measured 73 feet wide. I did a bit of mental arithmetic, divided 9 into 73, and got 8. So I thought the pattern would fit neatly. Now I find I have a foot left over, and the whole thing is spoilt." In other words, the mistake in mental arithmetic was an essential part of the causal chain that led to the angry display. Therefore, the best way to get a causal explanation was to do what behaviorism forbids—ask her what was going on in her mind.

An Example From Social Science

It was hypothesized that Black women in America have more children out of wedlock because their ideal of marriage is less conventional than that of Whites. Fortunately, they were asked what they thought: Attitude surveys revealed that they actually hold more conventional views. This led to a better motivational hypothesis. For every 100 Black women, there are only 51 Black men available to play the role of a conventional husband, the rest being dead, in prison, unemployed, itinerant, and so forth. Therefore, Black women often have the choice of either not having children at all or having them outside wedlock (Flynn, 1991, p. 135).

You can obscure the crippling effects of the behaviorist research strategy by selecting your tasks. It does worst when we want to explain or alter behavior, despite its pretensions, and best when we merely want to measure traits. A coach can measure a high-jumper's performance and physical characteristics without reference to mental states. If you want to explain or improve a given performance, it is best to know whether the athlete was worried about being dumped by a partner. Humphreys can ignore mental events because he has given himself the brief of measuring intelligence, correlating it with other traits, and basing predictions on those correlations. He has set aside motivation (which above all involves mental states), learning (which alters behavior), and diagnosis (discovering what goes on in people's heads when they solve problems). If he were trying to improve performance rather than measure it, the attempt to ignore mental events would break down. Perhaps Humphreys is simply saying that a behaviorist approach is usually adequate for trait measurement. If so, we are all behaviorists, and he has no case to argue.

Humphreys's Reasons for Dismissing IQ Gains Over Time

The best way to present a theory is to confront all relevant evidence, particularly evidence thought to count against it. IQ gains over time pose a problem for *g*-centered theory. For example, on tests considered the best measures of *g*, gains from one generation to another are too huge to be plausible as intelligence gains (Flynn, 1987a). Humphreys dismisses these gains, which poses the question as to whether he has good reasons for doing so.

Humphreys prefers Wechsler and Binet IQ tests, usually called measures of *crystallized g*, to tests like Raven Progressive Matrices, a measure of *fluid g*. Indeed, he refers to measures of "so-called" fluid intelligence and says that they prominently involve "visuospatial components." This is preceded by a section in which we are told that individual components of intelligence can change more rapidly than the total. Now it is true that the largest IQ gains over time have been on tests of fluid intelligence, such as Raven's, and, if the largest gains were visuospatial gains only, that would do something to mitigate the problem. There would be a price to be paid—namely, conceding that large differences in visuospatial ability do not cause large differences in socially relevant achievement. However, because I believe that to be the case, I can hardly press the point (Flynn, 1991, pp. 119–122, 140). Nonetheless, the argument is flawed. First, although IQ gains on crystallized tests are less than gains on fluid tests, the former are still massive, ranging from 9 to 20 points over a single generation (Flynn, 1987a, pp. 185–186). Second, as Jensen (1980) pointed out, Raven's simply does not show the male-over-female advantage that all visuospatial tests show (pp. 624, 646–647); Jensen argued that it "measures *g* and little else." Therefore, massive gains on it that do not appear to be intelligence gains pose a grave problem for *g*-centered theory.

The preceding argument is at best implicit in Humphreys, and he might disown it. His main reason for dismissing IQ gains over time is quite explicit. He writes: "Critics use such data to denigrate intelligence tests, claiming that 'real intelligence' did not change. A construct of 'real intelligence' cannot be inferred from any measurement operation". This is not a good reason for ignoring the problems that IQ gains pose.

On one level, those problems can be stated without any reference to a concept of intelligence simply by emphasizing that many of the predictions of *g*-centered theory have been falsified. The theory suggests that massive score gains on *g*-loaded tests would result in the present generation clearly out-achieving the last generation in relevant ways. Yet, despite gains of fully 1 *SD* over 20 or 25 years, teachers do not report speed of learning escalating in the classroom, and vocabularies do not seem much larger. People today who match the scores of normal subjects a generation ago still behave as if they suffer from mental retardation. Conversely, most people a generation ago do not seem to have behaved in that way, despite scoring below the criterion for mental retardation against today's norms (Flynn, in press). The fact that Humphreys does not discuss the falsified predictions of *g*-centered theory seems odd in that he states some of them. He says that *g* differences between groups should predict differences in educational achievement. Yet, despite a *g* advantage over the last generation, the present generation does no better (perhaps worse) on the Scholastic Aptitude Test, something Humphreys acknowledges.

On another level, discussing the problems posed for a *g*-centered theory does assume a concept of intelligence outside the theory, but only because everyone assumes such a concept sooner or later to avoid inco-

herence, including Humphreys himself. Moreover, assuming a concept of intelligence outside theories does not entail making references to "real intelligence." What I call the *primitive* concept of intelligence needs a theory-embedded concept of intelligence to attain scientific status; the theory-embedded concept needs the primitive concept to get off the ground. The relation between the two is symbiotic, and there is no profit in calling one more real than the other.

How to Define Intelligence

These rather cryptic comments necessitate a digression on the problem of how to define intelligence (Flynn, 1987b). This will pay dividends when we get to Humphreys's definition, but it can also free the *g*-men and, indeed, the proponents of all theories of intelligence from certain critics who bedevil them. These critics say that theories of intelligence are not intellectually respectable because they are not prefaced by fully articulated definitions of intelligence. Let us try to dismiss this incubus once and for all by looking at the origins of modern astronomy.

Aristotle thought that the planets were internally programmed to describe circles in the sky. Without a new concept on the pretheory level, modern astronomy was impossible. Fortunately, thinkers like Kepler were grappling with a new concept—celestial influence across distance. They could not escape the notion that the size and proximity of heavenly bodies had something to do with their relative motions. This concept led to theories that have proved far superior in predictive power to Aristotle's. However, it would have been futile to have attempted on the pretheory level to add more precision to the definition of the concept of celestial influence. The thing to do was to embed it in theories that generated predictions and to see which charted the heavens. Making the sun the revolving agitator of a whirlpool that caused the planets to swirl around added definition to the concept but did not work very well. Newton added definition to the concept of celestial influence by assuming that heavenly bodies attracted one another in proportion to their mass and inversely to the distance squared—a theory that proved very fruitful. Einstein added definition by assuming that space curved in the vicinity of mass and that objects in space moved along the shortest distance, which was even more fruitful.

In a brilliant article not sufficiently appreciated, Jensen (1979) elucidated the primitive or pretheory concept of intelligence. He imagined a Crusoe on an island. At first, there is no companion, but, even so, our Crusoe would remember things and acquire knowledge, thus demonstrating the concepts of memory and learning. Then a companion arrives, and they observe that one of them is slower at learning things and is less able to apply knowledge to new problems than the other. This generates the primitive concept of intelligence as a cognitive ability different from memory and learning, having something to do with speed of learning and memory retrieval, and having consequences for achievement dependent on foresight, generalization, invention, and so forth. Following the example of astronomy, I believe that it is futile to take defining intelligence much further than this on the pretheory level. Jensen quite rightly used his theory of intelligence to complete the task of definition: Just as Newton developed the primitive concept of celestial influence into the theory-embedded concept of gravitation, so Jensen, and Humphreys, have added definition to the primitive concept of intelligence by elaborating the theory-embedded concept of *g*.

We now know how we can honorably reject the eternal cry, "Your theory is not respectable because it does not rest on an adequate definition." It is simply wrong-headed to ask for a fully articulated definition on the pretheory level. And, if our critics say, "You don't mean to contend that your theory articulates what most people mean by intelligence, do you? You don't mean to imply that *g* will win universal assent?", we have a ready answer. It is that no theory-embedded concept will mirror the pretheory concept of people in general and that no particular theory, given competing theories, will win universal assent. We are offering something far better than a concept of universal appeal—namely, a concept embedded in what we think to be the best theory going. To ask more is to ask that we usurp the definitional role of theory—that pretheory supply something that only pretheory and theory together can provide.

Humphreys's Attempt to Define Intelligence

Humphreys is quite correct to give a pretheory definition of intelligence despite rejecting the notion of "real intelligence." He writes that "intelligence is the acquired repertoire of all intellectual (cognitive) skills and knowledge available to the person at a particular point in time" and adds that, "to avoid circularity, *intellectual* is defined by the consensus among experts working in the area." Despite its intent, the latter is a paradigm of circularity. We need only ask what "area" are these experts working in? Not physics, presumably, but the area of intellectual skills. So intellectual skills are ones defined as such by people who know what intellectual skills are. This is no better than the classic definition of art as what is defined as such by fellows of the Royal Academy; why do we appeal to them, presumably because they are experts on art? Following Moore (1903/1960), terms like *cognitive* can be clarified, say distinguished from *social* and *athletic,* only

ostensibly—that is, by examples. You can say, "Doing arithmetic is different from being polite or hitting a tennis ball." If that fails, you are dealing either with someone who lacks normal human experience or with someone trying to be clever.

The circular addition to Humphreys's definition does no real harm. However, the core of his definition suffers by not distinguishing primitive concepts of intelligence, memory, and learning from one another. His definition would gain some plausibility if it were offered as a definition of *crystallized intelligence,* which includes the learning an intelligent person normally acquires, although even then it would suffer by including acquired cognitive skills of a trivial nature, like test sophistication. Eventually, of course, Humphreys is forced to give such portions of the "total repertoire" a learning label: The acquisition of the skill of answering IQ-test items correctly is labeled *coaching* if acquired by practice, and such skills are said to lack the significance of those measured by *g* because *g* reflects "intelligence." Like everyone else, Humphreys generates these concepts on the pretheory level, and there seems no merit in their being blurred when he gives pretheory definitions.

By contrast, Jensen (1980) took full advantage of conceptual clarity on the pretheory level: His primitive concepts and his *g*-centered theory continually interact in productive partnership. He used the primitive concept of intelligence to generate a host of expectations about the theory-embedded concept and then demonstrated that *g* admirably fulfills those expectations (1980, pp. 114–115, 248–251). Tests that measure *g* differences show "intelligence" rising from infancy through adolescence, show good students more intelligent than poor students, show elite occupations with intelligence thresholds higher than those in ordinary occupations, show the sibling who rises in socioeconomic status more intelligent than the co-sibling who falls, shows monkeys more intelligent than dogs and dogs more intelligent than chickens, and so forth.

Sometimes Jensen has used *g* to refer to both outside-theory and theory-embedded concepts of intelligence. For example, when commenting on the fact that intervention experiments do not raise intelligence, he has said that there is no real gain in *g* (Jensen, 1989). Or, when referring to IQ gains over time, he has argued that score gains on what are normally *g*-loaded tests do not represent real gains in *g* (Jensen, 1987 pp. 380–381). This usage reveals Jensen's passionate belief in his theory. He believes that *g* capitalizes on the primitive concept far better than any competitor and therefore that there is no harm in assuming that something rather like *g* really does mold "intelligent behavior." This is like a Newtonian who, thanks to the success of the theory, speaks of gravitation defined therein as something at work thereout.

However, such usage is counterproductive. First, no theory has a right to assume that its theory-embedded concept will always best capitalize on a primitive concept. Second, using the same term to refer to both outside-theory and theory-embedded concepts of intelligence blurs the fact that distinct concepts are interacting and blurs the distinctive role played by primitive concepts. To be fair, Jensen's language usually gives unambiguous evidence of his primitive concepts at work. Like Humphreys, he makes the obvious comment that, if intervention experiments coach subjects to answer test items, higher scores merely represent "learning" gains of a rather trivial sort. Unlike Humphreys, he has said that he is troubled by the fact that IQ gains are too massive to be "intelligence" gains (cited in Bower, 1987). Jensen is troubled because IQ gains on *g*-loaded tests do not seem to constitute gains in test sophistication or test technique or school-learned material or skills. After all, if they cannot be explained away as learning gains, they should constitute intelligence gains. The primitive concept of learning, which protects *g*-centered theory from intervention gains so well, seems helpless against gains over time (Brand, Freshwater, & Dockrell, 1989; Flynn, 1987a, 1990, in press).

How to Survive Falsification

That some of its predictions have been falsified does not mean that Humphreys or anyone else must abandon *g*-centered theory. History shows that we never know in advance which falsifying evidence will prove fatal and which will eventually be reconciled. The theory that the earth revolves around the sun generates a prediction falsified by the absence of a stellar parallax: As the earth moves, the position of a fixed star in the heavens should appear to alter. No such phenomenon was observed, which seemed to pose a fatal problem. On the other hand, discrepancies between the predicted and observed orbit of Mercury seemed rather trivial. Yet, the former was eventually reconciled: The fixed stars are at such a great distance from the earth that shifts of position are difficult to discern, something that at the time seemed too incredible to take seriously. However, Mercury's orbit eventually allowed Einstein to transcend the whole of Newtonian astronomy.

The moral is that it is perfectly respectable to soldier on with your theory, despite what appears to be fatal falsifying evidence, just so long as there is no robust competing theory. Einstein himself made this point. No one was ready to give up on Newton, who had explained so much, until a better alternative theory was formulated. If the *g*-men believe their theory has no robust competition, they can turn their backs on IQ gains over time and press on. But they have no right to forget that they are there, and they must regard them as specters whose exorcism cannot be indefinitely postponed.

Summary

Humphreys' *g*-centered theory of intelligence deserves a fair hearing. However, the reception of a theory can be influenced by its mode of presentation. Therefore, Humphreys should vitalize his theory by using better primitive concepts, welcome all attempts at falsification, and eschew praise of behaviorist methodology.

Note

James R. Flynn, Department of Political Studies, University of Otago, Dunedin, New Zealand.

References

Bower, B. (1987). IQ's generation gap. *Science News, 132,* 108–109.

Brand, C. R., Freshwater, S., & Dockrell, N. B. (1989). Has there been a 'massive' rise in IQ levels in the West? Evidence from Scottish children. *Irish Journal of Psychology, 10,* 388–394.

Flynn, J. R. (1987a). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171–191.

Flynn, J. R. (1987b). The ontology of intelligence. In J. Forge (Ed.), *Measurement, realism and objectivity* (pp. 1–40). Dordrecht, The Netherlands: Reidel.

Flynn, J. R. (1990). Massive IQ gains on the Scottish WISC: Evidence against Brand et al.'s hypothesis. *Irish Journal of Psychology, 11,* 41–51.

Flynn, J. R. (1991). *Asian Americans: Achievement beyond IQ.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Flynn, J. R. (in press). IQ gains over time. In R. J. Sternberg, S. J. Ceci, J. Horn, E. Hunt, J. D. Matarazzo, & S. Scarr (Eds.), *Encyclopedia of human intelligence.* New York: Macmillan.

Jensen, A. R. (1979). The nature of intelligence and its relation to learning. *Journal of Research and Development in Education, 12,* 79–95.

Jensen, A. R. (1980). *Bias in mental testing.* London: Methuen.

Jensen, A. R. (1987). Differential psychology: Towards consensus. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy* (pp. 353–399). Lewes, England: Falmer.

Jensen, A. R. (1989). Raising IQ without increasing *g?* A review of the Milwaukee Project: Preventing mental retardation in children at risk. *Developmental Review, 9,* 234–258.

Moore, G. E. (1960). *Principia ethica.* London: Cambridge University Press. (Original work published 1903)

Skinner, B. F. (1972). *Beyond freedom and dignity.* New York: Knopf.

Humphreys's "Behavioral Repertoire" an Epiphenomenon of *g*

Arthur R. Jensen

School of Education

University of California, Berkeley

The clarity of thought and expression in Humphreys's target article is altogether admirable, and it is well nigh impossible to fault any of his empirically based statements or any of his reasoning and conclusions based thereon. I believe that this aspect of Humphreys's article fully accords with the overwhelming consensus of experts in psychometrics and differential psychology.

But I find myself in disagreement with Humphreys on two points of a theoretical nature: (a) Humphreys's insistence on a "pragmatic," "behavioristic" definition of *intelligence* as an acquired "repertoire" of intellectual skills and knowledge and (b) his definition of *intellectual* decided by a consensus of experts working in the area. I am in virtually complete agreement with everything Humphreys says after he makes these two points, which appear early in the article. Because he has reiterated this "behavioristic" definition without modification quite often during the past two decades, he obviously thinks it important. I have taken it seriously enough to be uncomfortable with it, and here I try to explain why.

First, the notion of a behavioral (or phenotypic) repertoire of acquired cognitive skills and knowledge: I argue that this repertoire is an epiphenomenon of a latent trait (or traits); individual differences in the size of the repertoire merely reflect individual differences in the latent trait and therefore can serve as one but not the only means of assessing or measuring individual difference in the latent trait. But I question the idea of using just one epiphenomenon as a definition of the essential phenomenon, which, in this case, is one or more latent traits.

Even pragmatically and operationally, repertoire per se seems to me a troublesome concept. Isn't it rather vacuous if all it means is anything a given person happens to know or can do that is deemed intellectual by a consensus of experts? And where does scientific objectivity come in when we allow a "consensus of experts" to decide what should or shouldn't be included in the repertoire of intellectual behavior? To be able to communicate and get on with their job, scientists must of course agree on certain formal definitions. But, as regards theoretical formulations, they need not agree except as empirical evidence compels them to. The question of which behavior is to be regarded as intellectual is a question science must try to answer and is