Intelligence 38 (2010) 363–366

Contents lists available at ScienceDirect

Intelligence

The spectacles through which I see the race and IQ debate☆

James R. Flynn

University of Otago, Dunedin, New Zealand

ARTICLE INFO

Article history:
Received 13 May 2010
Accepted 13 May 2010
Available online 9 June 2010

ABSTRACT

The ranking of Wechsler subtests in terms of their *g* loadings is equivalent to ranking them in terms of the cognitive complexity of the tasks measured. Lower performing groups do not always fall behind higher performing groups the more complex the task. But that is the general rule, no matter whether the cause of the lower performance is genetic or environmental. Complex tasks tend to be more affected by genetic differences in inherited traits, have higher heritability, and be more sensitive to inbreeding depression. Therefore, the method of correlated vectors sheds no light on the race and IQ debate. It is irrelevant that black/white score differences on Wechsler subtests rise as their *g* loading, heritability, and inbreeding sensitivity rise.

What is *g* and why is it interesting? People who do better than average on one cognitive task tend to do better than average on a whole range of cognitive tasks, for example, all 10 of the Wechsler subtests. We measure the strength of this tendency by *g*, and can then calculate the correlation between performances on each of the subtests and that factor. This ranks them into a hierarchy according to their *g* loadings, that is, the magnitude of the correlations.

1. The significance of *g* loadings

This hierarchy is interesting because it tallies with the cognitive complexity of the task. Digit span backward has a higher *g* loading than digit span forward. The latter merely involves remembering numbers in the order in which they are read out. The former requires reversing the order, that is, requires an extra mental operation. Presumably high-IQ people would be little better than average at shoe tying, a task with virtually no cognitive complexity. No matter what concept of intelligence you hold, it will focus on solving problems with cognitive complexity; and imply that above average people will soar higher above average the more complex the task.

When comparing groups, it is interesting to see how performance differences vary as *g* loadings rise. Rarely will a "superior" group lose ground, but it can happen, as when females best males on mathematical problems of moderate difficulty yet fall behind on the most difficult problems. Usually it will be an "inferior" group that falls further behind as cognitive tasks become more complex. This allows us to distinguish two kinds of problem-solving gaps between groups: the Wechsler *IQ gap* which treats the 10 subtests as of equal weight; and the Wechsler *GQ gap* which weights the subtest scores in accord with their *g* loadings. For example, the latter gives Vocabulary score differences between two groups double the weight of Coding score differences. Vocabulary has about twice the *g* loading of Coding.

The GQ and IQ gaps between the races differ by no more than one point (SD = 15). This seems surprising until we note that with the exception of Coding, the various Wechsler subtests differ little in terms of their *g* loadings. All of them measure either cognitively complex tasks (fluid *g*) or things like Vocabulary, whose acquisition reflects the cognitive complexity of assimilating the meaning of words (crystallized *g*). The small gap between GQ and IQ is perfectly compatible with

☆ Note: This paper should have followed the publication of: Rushton, J. P., and Jensen, A. R. (2010). The rise and fall of the Flynn Effect as reason to expect a narrowing of the Black–White IQ gap. Intelligence, 38, 213–219. Unfortunately it was not due to the confusion about submission. Douglas K. Detterman, Editor.

E-mail address: jim.flynn@stonebow.otago.ac.nz.

0160-2896/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2010.05.001

a moderate correlation between group difference and the hierarchy of *g* loadings.

You cannot dismiss the score gains of one group on another merely because the reduction of the score gap by subtest has a negative correlation with the *g* loadings of those subtests. In the case of each and every subtest, one group has gained on another on tasks with high cognitive complexity. Imagine we ranked the tasks of basketball from easy to difficult: making lay-ups, foul shots, jump shots from within the circle, jump shots outside the circle, and so on. If a team gains on another in terms of all of these skills, it has closed the shooting gap between them, despite the fact that it may close gaps less the more difficult the skill. Indeed, when a worse performing group begins to gain on a better, their gains on less complex tasks will tend to be greater than their gains on the more complex. That is why black gains on whites have had a (mild) tendency to be greater on subtests with lower *g* loadings.

Reverting to group differences at a given time, does the fact that the performance gap is larger on more complex then easier tasks tell us anything about genes versus environment? Imagine that one group has better genes for height and reflex arc but suffers from a less rich basketball environment (less incentive, worse coaching, less play). The environmental disadvantage will expand the between-group performance gap as complexity rises, just as much as a genetic deficit would. I have not played basketball since high school. I can still make 9 out of 10 lay-ups but have fallen far behind on the more difficult shots. The skill gap between basketball "unchallenged" players and those still active will be more pronounced the more difficult the task. In sum, someone exposed to an inferior environment hits what I call a "complexity ceiling". Clearly, the existence of this ceiling does not differentiate whether the phenotypic gap is due to genes or environment.

Correlations showing that group gaps in basketball skills rise with complexity loading (*g* loading), or rise with the heritability of the skill, or rise with how much the skill is affected by inbreeding depression make sense. Of course height and quickness are more important the more complex the skill, and of course these traits are heritable and adversely affected by inbreeding. But they do not decide the causal question.

2. History of a debate

Originally, Jensen argued: (1) the heritability of IQ within whites and probably within blacks was 0.80 and between-family factors accounted for only 0.12 of IQ variance — with only the latter relevant to group differences; (2) the square root of the percentage of variance explained gives the correlation between between-family environment and IQ, a correlation of about 0.33 (square root of 0.12 = 0.34); (3) if there is no genetic difference, blacks can be treated as a sample of the white population selected out by environmental inferiority; (4) enter regression to the mean — for blacks to be one SD below whites for IQ, they would have to be 3 SDs (3 × .33 = 1) below the white mean for quality of environment; (5) no sane person can believe that — it means the average black cognitive environment is below the bottom 0.2% of white environments; (6) evading this dilemma entails positing a fantastic "factor X", something that blights the environment of every black to the same degree (and thus does not reduce within-black heritability estimates), while being totally absent among whites (thus having no effect on within-white heritability estimates).

I used the Flynn Effect to break this steel chain of ideas: (1) the heritability of IQ both within the present and the last generations may well be 0.80 with factors relevant to group differences at 0.12; (2) the correlation between IQ and relevant environment is 0.33; (3) the present generation is analogous to a sample of the last selected out by a more enriched environment (a proposition I defend by denying a significant role to genetic enhancement); (4) enter regression to the mean — since the Dutch of 1982 scored 1.33 SDs higher than the Dutch of 1952 on Raven's Progressive Matrices, the latter would have had to have a cognitive environment 4 SDs (4 × 0.33 = 1.33) below the average environment of the former; (5) either there was a factor X that separated the generations (which I too dismiss as fantastic) or something was wrong with Jensen's case. When Dickens and Flynn developed their model, I knew what was wrong: it shows how heritability estimates can be as high as you please without robbing environment of its potency to create huge IQ gains over time.

I never claimed that the Flynn Effect had *causal* relevance for the black/white IQ gap. I claimed that it had *analytic* relevance. Jensen had argued that environment (at least between groups both located in a modern Western society) was so feeble that an astronomical environmental difference had to be posited to explain a one SD IQ gap. The Dutch showed that the environmental difference in question was less than whatever environmental enhancement they had enjoyed over 30 years. The gap needed was dragged out of the stars down to earth. If black IQ gains were 0.3 points per year, the environmental lag between blacks and whites would only amount to 50 years (0.30 × 50 = 15 IQ points). In my most recent book (Flynn, 2008, chapter 3), I proved that this was so. Scored against the whites of 1947–48, the blacks of 2002, some 54 years later, had a mean IQ of 104.31 and a GQ of 103.52.

3. The Flynn Effect mantra

Jensen (1998) complains that the Flynn Effect is repeatedly thrown at him as a kind of mantra. My recent book (Flynn, 2008, p. 79) offers an antidote: "Flynn himself . . . does not believe that it shows that blacks can [match whites for IQ] when environments are equal." Misinterpretation by the general public aside, there has been no rise and fall of the Flynn Effect: it never rose. Today, I can say just why casual explanation of IQ gains does not provide the key to the black/white IQ gap because I finally have a hypothesis about the former. The 20th century saw people putting on scientific spectacles that gave them new "habits of mind": rather than differentiating things to capitalize on their differential utility, people find it natural to classify things as a prerequisite to understanding; rather than tying logic to the concrete, people find it natural to take the hypothetical seriously and use logic on the abstract. Thus the huge score gains on Similarities (classification) and Raven's Progressive Matrices (logical sequences of symbols). See Flynn (2009).

This hypothesis erases a preoccupation that affected my 1999 exchange with Ruston. Gains on Raven's were so huge that I believed IQ gains must represent *fluid g* gains. Accordingly, I ranked the WISC subtests in terms of the magnitude of their correlations with Raven's, and found a modest correlation with the magnitude of IQ gains on each subtest. Today I would not be surprised or disturbed if a wider array of evidence negated this result. The significance of IQ gains rests on what they tell us about the evolution of our minds in the 20th century, not on whether we have some kind of *g* advantage on our ancestors. And the new habits of mind are too diverse and complex to be captured by the concept of "enhanced fluid *g*".

4. Status of the race and IQ debate

The fact that the GQ gap between blacks and whites is larger than the IQ gap has causal significance. If blacks did eliminate the IQ gap without eliminating the GQ gap, they would still be less able to solve the most complex cognitive problems, which might be deemed the most significant. Moreover, the fact that blacks have an unusual problem with complexity shows that an explanation of the IQ gap should look for aspects of the black environment that discourage cognitive challenge or at least, downgrade its presence. I took upon myself the burden offering a scenario of a succession of black environments from conception to early adulthood based on the deprivation of complexity (Flynn, 2008). It is significant that when the racial IQ gap was eliminated among post-war German occupation children, the GQ gap was gone. This is not to claim that this study settles the debate; rather it gives us confidence that if the IQ gap proves to be entirely environmental, the GQ gap will prove so as well.

American blacks are not in a time warp so that the environmental causes of their IQ gap with whites are identical to the environmental causes of the IQ gap between the generations. The race and IQ debate should focus on testing the relevant environmental hypotheses. The Flynn Effect is no shortcut; correlations offered by Rushton and Jensen are no shortcut. There are no shortcuts at all.

5. Comments on the text — by numbered section

1. The Flynn sources cited do not give the FE as a reason for expecting the B/W gap to disappear. They note the FE as a reason for entertaining hope of an environmental explanation. Flynn never said real intelligence levels were rising. He always searched for something that would show that gains were neither artifacts nor intelligence gains. The 1999 correlation with fluid *g* has been discussed, as has the significance of the fact that blacks of 2002 exceeded the IQ of whites of 1947–48.
2. The assertion that "if population group differences are greater on the more *g*-loaded and more heritable subtests, it implies they have a genetic origin" is simply false. That "culture-only theories predict a zero relationship between heritability and group differences" is false. I am unsure what "environmentality" means. All of the data from the "method of correlated vectors" is irrelevant. A "Jensen effect" does not evidence a genetic *g* but I suspect that *g* does have some root in brain physiology. That too is irrelevant to the origin of group differences.
3. I do not believe that outbreeding or any other genetic enhancement has caused IQ gains, but that has nothing to do with the insight provide into the cognitive history of the 20th century (Flynn, 2009, cha. 9). The Flynn/Rushton interchange is indeed passé for the reasons I have given. My belief that the GQ gap between black and white "tells us something about causes" has been explained. It does not reverse a past position but is a new insight.
4. Dickens and Flynn (2006a) published data showing that between 1972 and 2002, black Americans had gained 5.5 IQ points on whites and closed the *g* gap by 5.13 points. At both times, blacks lost ground on whites as they aged (from ages 4 to 24) by about 12 points. The critical points Rushton and Jensen raise were answered in our rebuttal to their critique at the time and I urge readers to consult it (Dickens & Flynn, 2006b).

Their survey of black academic achievement (with whites set at 100) shows: (1) Georgia 1954 — age 14 at 86; (2) Coleman report 1966 — ages 12, 15, and 18 at 87, 84, and 82; (3) NAEP 1975 — ages 13 and 17 at 70 and 71; (4) NAEP 2008 — ages 13 and 17 at 85 and 77. Note the tendency to lose ground with age. Note that the NAEP data for 1975 to 2008, a rough match for 1972 to 2002, show academic gains (15 points at age 13, 6 at age 17) larger than our estimate of IQ gains (5.5 points).

However, the data are deeply disturbing. Ruston, Jensen, and Flynn all think black IQ significant largely because it predicts academic achievement. Therefore, black IQ and the achievement values should be roughly the same. Does anyone really believe that blacks in Georgia in 1954 equaled black IQ nationwide (the Coleman report) in 1966? Jensen (1973) gives 80.7 for five Southeastern states including Georgia circa 1960. Does anyone really believe that the IQ of black 13-year olds has been bounding all over the place, from between 84 and 87 in 1966, down to 70 in 1975 (as low as subtropical Africa), up to 85 in 2008? The earlier data are not strictly comparable to NAEP data and perhaps we should trust the NAEP alone. But even it shows that in 1975, blacks performing as if they had a mean IQ of 70. And in 2008, it shows a decline in academic achievement between ages 13 and 17 of 8 points. No one would put the IQ drop between those ages at anything like 8 points.

6. Summary

(1) *g* would be of no interest were it not correlated with cognitive complexity. (2) Given a hierarchy of tasks, a worse performing group (whatever the cause of its deficit) will tend to hit a "complexity ceiling" — fall further behind a better group the more complex the task. (3) Heritability of relevant traits will increase the more complex the task. (4) Thus, the fact that group performance gaps correlate with heritability gives no clue to the origin of group differences. (5) When a lower performing group gains on a higher performing one, their gains will tend to diminish the more complex the task. Thus, blacks have gained 5.50 IQ points on whites since 1972 but only 5.13 GQ points. (6) Recent achievement test data confirm these IQ gains but the data as a whole pose problems

for the external validity of black IQ. (7) The FE is irrelevant to showing that the racial IQ gap is environmental but it was historically valuable in clarifying the debate.

References

Flynn, J. R. (2008). Where have all the liberals gone? Race, class, and ideals in America. Cambridge University Press.
Flynn, J. R. (2009). What is intelligence? Beyond the Flynn Effect [Expanded paperback edition]. Cambridge University Press.
Dickens, W. T., & Flynn, J. R. (2006). Black Americans reduce the racial IQ gap: Evidence from standardization samples. *Psychological Science*, *17*, 913−920.
Dickens, W. T., & Flynn, J. R. (2006). Common ground and differences. *Psychological Science*, *17*, 923−924.
Jensen, A. R. (1973). Educability and group differences. London: Methuen.
Jensen, A. R. (1998). Bias in mental testing. London: Methuen.