Submitted to *J. Intell.*. Pages 1 - 7.

OPEN ACCESS

Article

Semantic discussions of intelligence and the (un)importance of the study of race and *g*: A comment on Hunt and Jaeggi (2013)

Emil O. W. Kirkegaard [1,*]

[1] Department of Linguistics, Aarhus University, Niels Jensens Vej 8, 3.3, 138, Aarhus, Denmark

Version January 10, 2014 submitted to Joi. *Typeset by LaTeX using class file mdpi.cls*

Abstract: A commentary on parts of Hunt and Jaeggi (2013) dealing with the definition of intelligence, changes in intelligence, and the importance of the issue of race and intelligence.

Keywords: intelligence; definition; g-factor; race and intelligence.

1. Introduction

In their paper, *Challenges for Research on Intelligence*, Hunt and Jaeggi[1] draw attention to what is, in their view, the key interesting research areas on "intelligence" research. This paper is a commentary on that paper.

2. Definitions

2.1. Stipulative definitions

Hunt and Jaeggi claim that "word definitions are changed by usage, rather than by dictate, so we do not think that an elegant linguistic solution is likely". This is however not always the case. It applies to lexical definitions of the type usually found in dictionaries. However, in science (and math and logic), stipulative definitions are quite common[2]. For example, metric units were stipulated by various methods and have also been changed from time to time when problems or better ways of defining the units were found. *meter* has had many definitions throughout time, the most recent based on the distance light travels in a vacuum in $299{,}792{,}458^{-1}$ of a second. An earlier definition from 1799 used a particular prototype stick and *meter* was defined as the length of that stick[3].

As a further example, in astronomy, the definition of *planet* was recently changed. In classical Greek, *planet* meant 'wandering star', but as time went by it become quite foolish to group all the objects that traveled the skies into the category 'planet'. This original category included the Sun, the Moon, Mercury, Venus, Mars, Jupiter, and Saturn. This is an unsuitable definition for term as it excludes the most obvious planet, the Earth, and includes things that are quite dissimilar together (the Sun, the Moon, and all the modern planets). Later, the definition was changed so that whatever orbited the Sun was considered a planet. This removed the Sun and the Moon, and when technology enabled the observation of Neptune, Uranus and Pluto they were added as well. But the inclusion of Pluto broadened the scope of the class "planet" to an unwieldy size and a rather heterogeneous mixture of space objects beyond the familiar nine. For that reason, in 2006 the definition was again amended so as to exclude these objects[4,5].

2.2. The definition of intelligence

Hunt and Jaeggi appear to start another semantic discussion over the word *intelligence*, quoting Boring with his famous quote "what the intelligence tests test" (this is quite true, but not a good definition). This is an unproductive conflation of terms.

Arthur R. Jensen, in *The g Factor*[6], wrote an entire chapter (chapter 3, *The trouble with "Intelligence"*) about the semantic discussions of the word "intelligence", and he concluded that it was best to simply abandon the word as it had become too contaminated with other meanings, or as he wrote in the summary (p. 45):

> The word *intelligence* as an intraspecies concept has proved to be either undefinable or arbitrarily defined without a scientifically acceptable degree of consensus. The suggested remedy for this unsatisfactory condition is to dispense with the term *intelligence* altogether when referring to intraspecies individual differences in the scientific context and focus on specific mental abilities, which can be objectively defined and measured. The number of mental abilities, so defined, is unlimited, but the major sources of variance (i.e., individual differences) among myriad abilities are relatively few, because abilities are not independent but have sources of variance in common.

Hunt and Jaeggi point this out themselves when they cite the OED for giving some eight definitions plus sub-definitions. Other researchers e.g. Linda Gottfredson[7] also think the question is moot (p. 27):

> Theorists have long debated the definition of "intelligence," but that verbal exercise is now moot. *g* has become the working definition of intelligence for most researchers, because it is a stable, replicable phenomenon that–unlike the IQ score–is independent of the "vehicles" (tests) for measuring it. Researchers are far from fully understanding the physiology and genetics of intelligence, but they can be confident that, whatever its nature, they are studying the same phenomenon when they study *g*. That was never the case with IQ scores, which fed the unproductive wrangling to "define intelligence." The task is no longer to define intelligence, but to understand *g*.

Nevertheless, *g* is not a term designed for its ease of use in conversation, though there are alternatives, e.g. general cognitive ability (GCA), general mental ability (GMA), and general intelligence.

Psychometrics researchers, when speaking among themselves, simply call it "intelligence", while knowingly referring only to *g*. It seems rather moot, as Gottfredson put it, to again bring up the verbal definition debates on *intelligence* in any interest beyond endless semantic quibbling as is often found in philosophy (see e.g. [8]).

For those more technically inclined, there are still difficulties, but they are quite small. General factors can be extracted from a dataset in many ways, but they yield mostly the same result when compared by formal methods[9]. Still, it was possible that the different batteries of mental tests and the *g* factors extracted from them yielded different *g* factors. That this was mostly not the case was demonstrated recently[10,11], the authors reporting correlations among *g* factors from different batteries given to the same set of persons around 1 (without correction for measurement error). The lowest correlations were between a battery of tests using only nonverbal matrix type tests (Cattell Culture Fair Test) and even these were .77, .79, .88, and .96 (Table 2 in [11]).

Still, it is possible that these tests, totaling some 46 different subtests, did not include sufficient variation in the tests to capture all the possible variance. This is quite possibly so, because none of these tests were chronometric (or other Elementary Cognitive Tests, see [12]) tests which are known to correlate with *g*[12, Chapter 9], and there is some evidence that a *g* factor extracted from a battery with chronometric tests is a better predictor than a *g* from the same battery without the chronometric tests[13–15]. This finding is in need of further replication. If it holds, it means that the current testing batteries are missing possible variation in *g* which will improve the predictive power of the tests. It would also mean that we have been underestimating the predictive power of *g*, which is often the case anyway since researchers do not correct for measurement error.

For the most conceptual clarity, one must distinguish between IQ, extracted *g* and idealized *g*. Idealized *g* is the human trait which is the active ingredient (as Jensen put it[6, p. 271]) in all IQ tests. It could be defined as the general factor from a battery of tests that capture all the possible variance in mental ability. Using an infinite number of humans of all ages, tested with an infinite number of cognitive tests, which are maximally different as to capture all variance that can be captured in a general factor, and then extracting a general factor from the resulting dataset would yield idealized *g* (see also [6, p. 31]]). In practical terms this is ludicrous. However, one can come quite close to this idealized *g* with extracted *g*'s from already standardized batteries. When the predictive power of the extracted *g*'s doesn't increase with the addition of any type of test, the limit has been reached. Whenever *g* is referred to without any further qualifications, it is assumed to be this idealized *g*.

That is not to say that there is not more to the mind than *g*, and recent studies have attempted to renew interest in some non-*g* abilities (e.g. spatial ability[16]). But for many domains, *g* is quite clearly the most important ability. This was neatly shown in a recent study of about 70,000 children in the United Kingdom, where the *g* factor had predictive validity coefficients .69-.72 while those for a non-*g* verbal ability residue was around .13-.14. As expected, this verbal factor was more important for predicting performance in language class (validities in the .14-.20 range) while lower for science classes (validities in the .00-.14 range with .00 for math)[17].

3. Changes in *g*

Hunt and Jaeggi mention it as an uncontroversial fact that intelligence is rising citing one of James Flynn's works. However, it is not clear that the Flynn-Lynn-Runquest effect (FLR effect, see [18]) is a real increase in what researchers normally call *g*. If one gives people training on how to take an IQ test, or just re-tests them on the same test within a short amount of time, they will increase their raw scores[19], just like the FLR effect. If one then uses the scoring from the test manual, they will have increased their IQ too. However, few, if any, would regard this as an actual increase in *g*, but instead some kind of measurement error.

While test training is itself an interesting topic, it is quite conceptually distinct from changes (increases or decreases) in the construct in question, *g*. Unfortunately, most previous studies of the training of intelligence (see [20]) only looked at IQ scores, and not *g* scores or the *g*-loadedness of the changes in *g*. If it was a real change in *g*, the tests that are the most *g*-loaded should change the most, or in other words, the correlation between a test's *g*-loadedness and the change should be positive. It turns out that training effects and the FLR effect are not *g*-loaded, but effects from inbreeding are. Data needs to be found from the old studies so that these modern analyses can be run, or if the data is lost, new studies need to be done. This is the price to pay for the researchers' lack of data sharing. Hopefully, this journal can make a great contribution to the study of intelligence by both having open access, and data sharing policies (cf. [21]). It would be even better if an open data repository was created. In any case, I look forward to sharing data with other researchers.

4. On the importance of the issue of race and *g*

Hunt and Jaeggi are right that studying race and *g* is probably not a good way to study the nature of *g* although it seems quite likely that some understanding can be gained from that direction, as it indeed can from any direction. We might not know of Spearman's Hypothesis (group differences in *g* are *g*-loaded i.e. highest on the most *g*-loaded subtests, see [22, Section 4]) if we had never studied racial differneces in *g* scores. They are wrong to downplay the issue writing that "In no case, though, do we see research on racial differences in intelligence as being a high-priority scientific topic.". Western countries everywhere have a fertility problem (this section is based on [23]). The number of children per woman (total fertility rate) is too low for sustaining their populations[24]. When these countries also have welfare systems that only work economically with sizeable younger generations (who contribute economically to society) we have the possibility of economic disaster (See Figure 1 for data from Denmark).

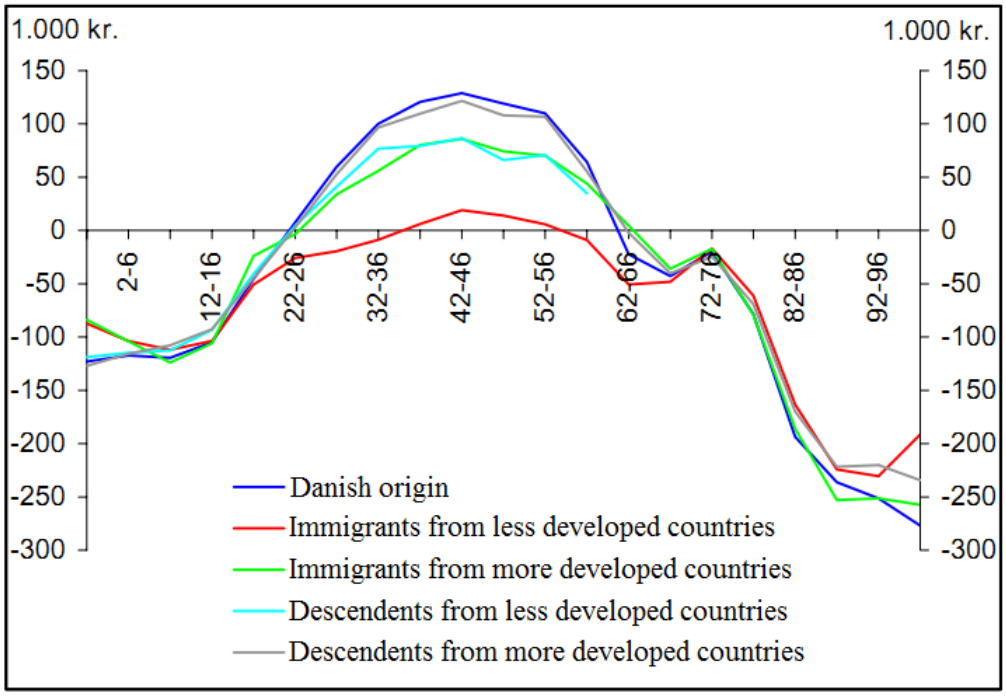

Figure 1. The net contribution to society from different groups over the course of their lives. From [25, p. 386, my translation].

Since it is unlikely that the tendency toward longer educations and low birth rates for Western women will reverse, another solution must be found. The current humanistic, egalitarian tendencies in politics make the choice obvious: Open the borders and let people in. Who is going to say no to refugees fleeing from war, disease, and hunger? The solution can work, but only providing the immigrants contribute to society in the same (or greater) capacity as the current inhabitants. However, if this is not the case, they will instead become an economic burden, and this attempted solution will only make things worse. Since *g* is one of the major determinants of income, social status, crime rates etc., it becomes critically important for predicting the potential economic and societal performance of immigrant groups, and by extension the impact they can be expected to have on the standard of living in the accepting country. If a group's average genetic levels of *g* are lower than that of current inhabitants, the performance gap cannot be expected to close (absent gene therapy or cognitive implants or the like). In that case, immigration will only increase the economic problems in the country in question. In other words, the question of race and *g* has important social policy implications for immigration policy. There are other areas, like affirmative action, where there are also social policy implications. This is not to say that the results imply specific policies themselves; as Rushton and Jensen wrote in their review of 30 years of research on race and *g* "no specific policies necessarily follow from knowing about the causes of group differences"[22]. But research findings can help predict the results of a given policy if adopted. Research into race and *g* is thus vital for evidence-based politics in that area.

Hunt and Jaeggi claim that the issue isn't important because "Due to migration and intermarriage, the identity of different racial groups can change in a very few years.". That depends quite a lot on what is meant by "a very few years"! Due to assortative mating, people tend to marry people from their own racial groups (endogamy, for the case of Ashkenazi Jews, see [26,27]), which seriously slows down the mixing process. Unless racial groups are forced by some means to start interbreeding, race differences

in *g* will not disappear any time soon due to mixing. The authors' downplaying of race differences in *g* is harmful to science as it perpetuates what Gottfredson called "The Egalitarian Fiction"[28]. The goal of science is to find out how the world works. This includes information that is uncomfortable to some. Without correct information, how are societies to make the best decisions?

Acknowledgements

Tegan McCaslin for editorial help.

Conflicts of Interest

The author declares no conflicts of interest.

References

1. Hunt, E.; Jaeggi, S.M. Challenges for Research on Intelligence. *Journal of Intelligence* **2013**, *1*, 36–54.
2. Swartz, N. Definitions, dictionaries, and meanings, 1997.
3. Wikipedia. Metre — Wikipedia, The Free Encyclopedia, 2014. [Online; accessed 10-January-2014].
4. Wikipedia. IAU definition of planet — Wikipedia, The Free Encyclopedia, 2014. [Online; accessed 10-January-2014].
5. CGPGrey. Is Pluto a planet?, 2012.
6. Jensen, A.R. *The g factor: The science of mental ability*; Praeger Westport, CT, 1998.
7. Gottfredson, L.S. Where and why g matters: Not a mystery. *Human Performance* **2002**, *15*, 25–46.
8. Ichikawa, J.J.; Steup, M. The Analysis of Knowledge. In *The Stanford Encyclopedia of Philosophy*, Fall 2013 ed.; Zalta, E.N., Ed.; 2013.
9. Jensen, A.R.; Weng, L.J. What is a good g? *Intelligence* **1994**, *18*, 231–258.
10. Johnson, W.; Bouchard Jr, T.J.; Krueger, R.F.; McGue, M.; Gottesman, I.I. Just one¡ i¿ g¡/i¿: consistent results from three test batteries. *Intelligence* **2004**, *32*, 95–107.
11. Johnson, W.; Nijenhuis, J.t.; Bouchard Jr, T.J. Still just 1¡ i¿ g¡/i¿: Consistent results from five test batteries. *Intelligence* **2008**, *36*, 81–95.
12. Jensen, A.R. *Clocking the mind: Mental chronometry and individual differences*; Access Online via Elsevier, 2006.
13. Luo, D.; Petrill, S.A. Elementary cognitive tasks and their roles in¡ i¿ g¡/i¿ estimates. *Intelligence* **1999**, *27*, 157–174.
14. Luo, D.; Thompson, L.A.; Detterman, D.K. The causal factor underlying the correlation between psychometric¡ i¿ g¡/i¿ and scholastic performance. *Intelligence* **2003**, *31*, 67–83.
15. Luo, D.; Thompson, L.A.; Detterman, D.K. The criterion validity of tasks of basic cognitive processes. *Intelligence* **2006**, *34*, 79–120.

16. Wai, J.; Lubinski, D.; Benbow, C.P. Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology* **2009**, *101*, 817.
17. Deary, I.J.; Strand, S.; Smith, P.; Fernandes, C. Intelligence and educational achievement. *Intelligence* **2007**, *35*, 13–21.
18. Lynn, R. Who discovered the Flynn Effect? A review of early studies of the secular increase of intelligence. *Intelligence* **2013**.
19. te Nijenhuis, J.; van Vianen, A.E.; van der Flier, H. Score gains on¡ i¿ g¡/i¿-loaded tests: No¡ i¿ g¡/i¿. *Intelligence* **2007**, *35*, 283–300.
20. Spitz, H.H. *The raising of intelligence: A selected history of attempts to raise retarded intelligence*; Lawrence Erlbaum Associates, Inc., Publishers, 1986.
21. De Boeck, P. Open Access Intelligence. *Journal of Intelligence* **2013**, *1*, 1–4.
22. Rushton, J.P.; Jensen, A.R. Thirty years of research on race differences in cognitive ability. *Psychology, public policy, and law* **2005**, *11*, 235.
23. Kirkegaard, E.O.W. Befolkningsproblemet, 2013. Danish.
24. Wikipedia. Sub-replacement fertility — Wikipedia, The Free Encyclopedia, 2013. [Online; accessed 10-January-2014].
25. Velfrdskommissionen. Fremtidens velfærd kommer ikke af sig selv. *Analyserapport. Copenhagen* **2004**.
26. Cochran, G.; Hardy, J.; Harpending, H. Natural history of Ashkenazi intelligence. *Journal of Biosocial Science* **2006**, *38*, 659.
27. Lynn, R. *The chosen people: A study of Jewish intelligence and achievement*; Washington Summit Publishers, 2011.
28. Gottfredson, L.S. Egalitarian fiction and collective fraud. *Society* **1994**, *31*, 53–59.