Intelligence xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Intelligence

Why teach intelligence?☆

N.J. Mackintosh *

University of Cambridge, UK

ARTICLE INFO

Article history:
Received 9 August 2013
Accepted 9 August 2013
Available online xxxx

Keywords:
Teaching
Intelligence
Advanced course

ABSTRACT

IQ tests are one of psychology's more visible and controversial products. For this reason alone, a student who has graduated with a degree in psychology ought to know enough about the subject to dispute some of the public's misconceptions. Controversy breeds disagreement, and although intelligence researchers are agreed on some of the conclusions suggested by their research, they disagree strongly about others. One reason is that many see desirable or undesirable implications of such research, and their evaluation of the research is influenced by those perceived implications. Another is that the nature of intelligence research, where well-controlled experiment is usually not possible, and conclusions are based on mere correlations or the results of necessarily ill-controlled natural experiments, means that not all conclusions are unequivocally dictated by the evidence. For these reasons an advanced course on human intelligence can teach a student how to evaluate necessarily ambiguous evidence, without being swayed by his or her prior beliefs or wishes.

1. Introduction

It seems reasonable to start on a personal note with my own experience of teaching intelligence. I have taught courses on intelligence at Cambridge for the past 30 years, an advanced final year undergraduate course throughout that time and also, until 10 years ago, a few lectures as part of introductory psychology courses to scientists and medical students. On the whole the lectures have been well received — the main exception being a (thankfully) small minority of medical students who, facing exams in anatomy, physiology, etc., which, if they do not pass well, they will have to resit, have resented the time wasted sitting through any psychology lectures[1]. Of particular relevance to the concerns expressed by the editor when compiling this special issue of the journal, the advanced final year course, although optional, has always been very popular, regularly attracting well over half of the students taking psychology in their final year. This has been in a department of experimental psychology, with a strong emphasis on behavioral and cognitive neuroscience, sensory psychology and animal behavior, which, until very recently, provided essentially no teaching in social psychology or personality. I take little credit for the course's popularity: I am certain that it is the course's subject matter that attracts the students. I think that any such course can be guaranteed to be popular. Perhaps we should listen to what students want.

The structure of the course in recent years (although of course it varies from year to year) is roughly as follows. I start with a bit of history, noting the unfortunate timing that saw IQ tests being first developed when the demographic transition was causing people like Raymond Cattell to worry about the decline of national intelligence. This leads naturally into a discussion of the Flynn effect, which leads into a discussion of environmental and social class effects on intelligence, which leads into a discussion of heritability. This is usually followed by group differences. My intention is to attract students by beginning with the more visible (controversial?) aspects of the subject, before hitting them with factor analysis and the structure of human abilities; processing speed and changes in intelligence with age; the relationship between IQ and executive functions such as working memory; the brain; and

☆ I am very grateful to Scott Barry Kaufman for his helpful comments on a draft of this paper.
* Department of Psychology, Downing Street, Cambridge CB2 3EB, UK.
E-mail address: njsmm10@gmail.com.
[1] My favorite example of a dissatisfied student was the one who wrote in his evaluation of my first lecture: 'Professor Mackintosh has been talking for 40 min, and he still hasn't said anything interesting'.

0160-2896/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.intell.2013.08.001

finally a discussion of the predictive validity of IQ scores and the possibility that IQ tests fail to capture all important aspects of human intelligence: this includes a brief mention of Gardner's multiple intelligences, but more emphasis on social, emotional and practical intelligence and the role of expertise (it is in these last topics that I suspect many psychometricians would find much to disagree with).

I do not hesitate to discuss controversial issues (such as sex differences for example), because I find that most students want to learn about them, and I want them to learn how to evaluate scientific findings without regard to their own personal prejudices (see below). I try to encourage them to reach their own conclusions. Although I do introduce them to factor analysis, I spare them much mathematical detail: even though most of them have a relatively strong scientific background, long experience has taught me that, if I want to keep students in the class, this is a sensible decision. I do not doubt that research in intelligence requires a proper understanding of factor analysis and psychometrics. But this is, after all, an undergraduate, not a graduate, course.

I also give talks to 6th form (high school) students, to various university and other societies, and to the 'University of the 3rd Age' (older people keen to maintain their intellectual curiosity).

2. An obvious reason for teaching human intelligence

An obvious reason why all psychology students should be exposed to an advanced course on human intelligence is that this is an area where significant progress has been made in the past 50 years or so, and many of the important findings of this new research are unlikely to be taught in an introductory course. One version of this argument suggests that there is now general consensus among intelligence researchers about many of the issues that were once bitterly disputed (Neisser, 1996), but this general consensus is still widely misunderstood.

There are other reasons. One is that cognitive psychologists/neuroscientists and intelligence testers actually share some common interests and should sometimes have a common research agenda. Each party therefore needs to know more about the other. For example, in spite of the well established relationship between IQ and working memory, measures of either Gf or Gc appear to be only weakly related to other so-called executive functions (Ardila, Pineda, & Rosselli, 2000; Friedman, Miyake, Corley, et al., 2006). One interpretation of this finding is that 'psychometric intelligence tests do not appropriately appraise intelligence. Or, at least, they are not appraising abilities that, from a neuropsychological perspective (and also, from the point of view of Wechsler's intelligence testing), should be understood as the most important elements in cognition' (Ardila et al., 2000, p. 35). This is clearly a message that intelligence testers should take on board. But the claim that tests of executive function measure the most important elements of cognition needs much stronger justification than Ardila et al. (or Friedman et al. who advance a similar argument) have provided. There is an unfortunate air of confrontation in their claims, which is hardly helpful. Is it reasonable to suggest that it stems from a mistrust of intelligence testing that might be alleviated by more teaching of the subject?

This is not, however, the only reason for believing that many psychologists would benefit from an advanced course on intelligence. I want to advance a rather different argument, one which may be seen by some as an attack on intelligence research. It is certainly not intended as such.

I start with a relatively uncontentious point. For better or worse, IQ tests are one of the more visible products of psychology; the nature of intelligence is a topic of widespread interest; and the possibility that people might differ in 'native intelligence', and the possible causes of such differences, often arouse fierce discussion. Many readers of this journal will argue that the public is seriously misinformed about these topics, citing Snyderman and Rothman's (1988) excellent book in support of their argument. I do not wish to dispute their point, but it can wait. Whether or not the public is misinformed, students who have graduated from university with a degree or major in psychology ought surely to be a great deal better informed about a topic of such widespread interest than members of the general public. They will not be, if all they have learned is from a chapter in an introductory textbook or a couple of lectures in Psychology 100.

At best, such students will have learned that intelligence tests were first developed by Alfred Binet, and will have been shown some examples of such tests; they will have learned that intelligence is affected by both genes and environment, and with luck that heritability is a population statistic, which does not refer to the proportion of any individual's intelligence that is determined by her genes or her environment. They will have heard of *g* and of multiple intelligences and have been told about test reliability and validity, and that IQ predicts educational attainment. This is all fine — but it is not very much.

Would such a student be able to argue with a critic like Murdoch, who claims that IQ tests 'do not test intelligence and have negligible ability to predict academic achievement' (Murdoch, 2007, p. 231), or others who assert that the only reason why they predict any other accomplishments is because they are disguised measures of social class or family background which are the real determinants of such accomplishments? Sackett, Borneman, and Connelly (2008) provide a number of examples of such claims. It is not necessary to insist that the critic is clearly wrong — only that there are at least some counterarguments which can be deployed against his position, and that a well educated psychology graduate should be able to advance some of these counterarguments.

3. The preconceptions (and ignorance) of other psychologists

I suspect that relatively few academic psychologists can or would. It is not only the public that harbors some misconceptions about intelligence tests but also other psychologists. Most experimental psychologists take a decidedly dim view of intelligence testing: being one myself I am familiar with their attitude. When, in 1972, Leo Kamin gave an address to the Eastern Psychological Association (later expanded into a book, Kamin, 1974), in which he denounced Cyril Burt and concluded that 'there exist no data which should lead a prudent man to accept the hypothesis that I.Q. test scores are in any degree heritable' (p. 1), he received a standing ovation at the end of his lecture. When I was writing a review of his book (Mackintosh, 1975), I discussed it with several of my experimental colleagues, and was astonished at their unwillingness to dispute some of his more suspect arguments, or to find anything amiss in Kamin's habit of statistical trawling through data to find the desired result — a practice they would have dismissed out of hand in their own professional work.

Why was this? The obvious answer is because they *wanted* to believe what Kamin was saying, because they found that at least some of the conclusions suggested by research on intelligence conflicted with many of their social values and beliefs. Many people do not like the idea that there might be inherited differences between people that have a substantial impact on their life chances. The difference that has university teachers most worried is a difference in intelligence (in the wider scheme of things, differences in the skills needed to be an outstanding professional footballer or popular musician probably have more impact on fame and fortune, but that is another matter).

4. The importance of seeing through bias

What is the relevance of any of this? Surely that one of the more important lessons that can be learned from an advanced course on human intelligence is that scientific arguments must be judged on their *scientific* merits, not on whether they are consistent with one's social and political beliefs. One of the more pernicious arguments Kamin advanced was that he was writing "about the politics of intelligence testing, as well as the science of intelligence testing. To pretend that the two are separable is either naïve or dissembling" (p. 2).

No one would dispute that, as a matter of fact, much that has been written on the subject of human intelligence has been influenced by the author's social or political views. Here are two examples chosen more or less at random.

The critic who insists that correlations between test scores and later educational and occupational achievements are entirely attributable to the advantages and disadvantages of family background probably *wants* to believe that we live in a class-ridden society. More to the point, he has simply ignored the finding that similar correlations can be observed *within* families: the sibling with the higher test score will be likely to obtain higher educational qualifications and earn a higher income than the sibling with the lower test score (Murray, 1998).

A topic that has generated even more heated argument is the so-called 'Greater Male Variance' hypothesis first advanced by Havelock Ellis and Francis Galton over 100 years ago to explain why there were many more highly distinguished men than women. It is hard to believe that anyone could have seriously thought that this was the only explanation of, say, the observation that more men than women have been awarded Nobel prizes, but that does not mean that the hypothesis can be dismissed. Actually it is more an empirical observation than a hypothesis (and it is not obvious why it deserves capitalizing). There is ample evidence that there are more males than females both at the top and at the bottom of the distribution of scores on many cognitive measures, including both overall IQ and more specialized abilities, such as Math, reading, or spatial ability (several very large studies have shown this for general IQ: Deary, Irwing, Der, & Bates, 2007; Deary, Th orpe, Wilson, Starr, & Whalley, 2003; Strand, Deary, & Smith, 2006; while Hedges & Nowell, 1995, reported that on 34 of 38 measures of specific abilities males were more variable than females). In her book, *Delusions of Gender*, Fine (2010) sets out to challenge the hypothesis, by arguing that although there may be evidence of greater male variability in many countries, it is not universally observed, which proves that there is nothing inevitable or immutable about it. But inevitability and immutability are not a necessary part of the hypothesis (even if some writers have implied that they are): it is better seen as an empirical generalization, whose explanation is open to question. There are, no doubt, plausible genetic explanations for the preponderance of males at the bottom of many distributions, but rather fewer for the preponderance of males at the top end of the distribution. It is entirely possible that social factors are important here. It is worth adding that the large majority of cross-national studies have found evidence of significantly greater male variability: in one study (Machin & Pekkarinen, 2008), for example, in 35 of 41 countries males were significantly more variable than females on tests of reading, and significantly more variable on tests of Mathematics in 36 of 41. In no case were females significantly more variable than males.[2] In most areas of psychology, if 35 or 36 of 41 independent observations recorded a significant difference between two groups always in the same direction, and no significant difference in the opposite direction, I think it would be generally accepted that the difference was 'real'. To imply, as Fine seems to, that it is somehow just an artifact of different cultural attitudes and practices, seems a shade optimistic.

I think it is probably fair to say that in both of these examples the author's conclusions have been influenced by their prior beliefs and wishes. Fine, after all, hardly disguises the central thesis of her book. Some readers will probably think that it is I who am biased in this second case, but although I would be reluctant to accept such a charge, for present purposes it does not matter. My argument is simply that bias is not unknown in writings about some of the issues raised by research on intelligence. Indeed, because of its controversial nature, research on human intelligence has perhaps attracted more than its fair share of suspect arguments. In consequence, an advanced course on human intelligence can provide the opportunity to acquire the necessary attitude and skills to see through them. Students who have taken such a course should be alert to the possibility of bias; they should have learned to check references, be prepared to challenge an author's conclusions, and to take little on trust.

Most readers will no doubt be able to provide other examples of biased arguments, where the author has apparently chosen to present only one side of the case and ignored contrary evidence. But Kamin's argument went further; he insisted that such biases are an *inevitable* part of all arguments about human

[2] Fine cites two other recent studies of international comparisons, Guiso, Monte, Sapienza, and Zingales (2008) and Penner (2008). In the former, there were only two of 40 countries where there were more females than males in the top 10% of the distribution of Math scores (Indonesia and Thailand), and again only two with more females in the top 5% (Thailand and Iceland). Pennington studied 22 European and Western countries; in only three (Germany, Lithuania and the Netherlands — but not Iceland) the standard deviation of females' Math scores was greater than males', and in all 22 there were significantly more males than females in the top 10% of the distribution.

intelligence — in particular that a belief in the heritability of IQ is little more than a reflection of conservative, elitist values. Such a suggestion seems tantamount to condemning all such writings (including, of course, Kamin's own) to the dustbin as probably worthless expressions of political preconception.

5. The inherent ambiguity of some intelligence research

One reason why Kamin's argument has seemed so plausible is that it seems to explain why it is that intelligence researchers themselves often disagree, quite honestly, about the proper interpretation of their data. Sometimes these disagreements are stark. How can it be, for example, that Rushton and Jensen (2005) on the one hand, and Nisbett (2005) on the other, can review exactly the same data set and reach such diametrically opposed conclusions — that the difference in average test scores between African and white Americans is largely caused by genetic differences or entirely by environmental differences? I do not believe for one moment that this difference simply arises from a difference in the authors' political views. A more plausible explanation is that the data themselves are inherently somewhat ambiguous: they do not *unequivocally* force one to accept one conclusion rather than another. Nisbett, for example, can point to the finding (see Dickens & Flynn, 2006) that the test score gap today is smaller than it was 25 or 50 years ago — as an environmentalist would expect given the changes in American society that have occurred over this period. Rushton and Jensen can equally point out that the gap is not *very much* smaller (and certainly has not disappeared), and that some test batteries have shown little or no decrease.

Consider a rather less contentious example. In *A day at the races*, Ceci and Liker (1986) reported that one group of racecourse habitués was substantially more successful than another group at predicting which horses would be the first, second and third favorites in tomorrow's races; that success in such prediction appeared to depend on their use of more complex algorithms that critically took account of the interactions between a number of different variables; but that it was wholly unrelated to their IQ scores. Following an earlier critique of the study by Detterman and Spry (1988), Hunt concluded: “The Ceci and Liker study presents us with good news and bad news. The good news is that when a published study contains major flaws, other scientists point out the errors. The bad news is that almost no one notices the correction” (Hunt, 2011, p.317).

I on the other hand took a notably more favorable view (Mackintosh, 2011, pp. 229–230). Hunt's major criticism was that the prediction of tomorrow's favorites was a very unreliable measure, and this, combined with the relatively small sample sizes, meant that the absence of a significant correlation with the participants' IQ scores was only to be expected. My more favorable take on the study was based on the fact that there was a significant difference between the two groups' ability to predict favorites (on one measure actually no overlap between their scores), and in the complexity of the algorithms they used, but no difference between their IQ scores, which ranged from 81 to 125 in the more successful group, and from 80 to 130 in the less successful. While Hunt's point is surely valid, perhaps unsurprisingly I also believe that the findings I laid stress on do tend to support Ceci and Liker's conclusion. Perhaps we were both guilty of selective reporting? Given what I believe to be the potentially important implications of the study, it is unfortunate that it has not been followed up by others.

Be this as it may, the point I want to emphasize is that the data on human intelligence are often consistent with alternative explanations. They are rarely the product of carefully controlled laboratory experiment; more often they are the result of imperfect natural experiments; sometimes simply the result of naturally observed correlations. I do not want to suggest that research on human intelligence has never produced unequivocal findings. That would be absurd: the meta-analysis of large data sets has surely yielded many securely based conclusions. No one who has studied the evidence could seriously doubt the existence of the positive manifold; that IQ scores are partly heritable; that IQ independent of SES predicts educational attainment; or that children with higher IQ tend to live somewhat longer than those with lower test scores (of course all these points have been disputed by some who have *not* studied the evidence).

It would be equally absurd to claim that the findings of experimental psychology are all so secure that they cannot be disputed, or that even when the data are generally accepted, there is no dispute about their theoretical interpretation. Nevertheless, it remains true that natural experiments are often less than perfectly controlled, and conclusions based on them will sometimes remain open to dispute. Moreover, it is not always easy to find natural experiments that address a particular issue, which is surely one reason why there have not been *very* many studies that have sought to understand the reasons for the black–white difference in average test scores. The relatively small number of such studies, and their almost inevitable imperfections, suggests to me at least that they have not provided a definitive answer to the question. Nor, I venture to predict, are they likely to in the near future.

What the keen scientist may find exasperating, however, the teacher should see as a golden opportunity. An advanced course on human intelligence provides an ideal setting in which to teach students that the proper interpretation of any set of data requires ignoring your preconceptions, looking at the data as a whole and not glossing over some inconvenient aspect that spoils a good story; it may require looking at other data and theories to see how far they support one interpretation rather than another; it is not simply given by the data, but requires careful judgment and the balancing of probabilities. And it almost certainly needs stating with caveats and qualifications. Much science teaching is a matter of presenting the data for the student to memorize, of laying down the law, contrasting today's correct theories with the errors of the past. But even in the experimental sciences, today's theories may well be overturned by tomorrow's, and none should be accorded the status of holy writ. They are at best hypotheses that, for now at least, seem on balance to provide a more convincing explanation of the data at hand than their rivals. The proper education of a scientist requires that she learns this lesson, and precisely because the data on human intelligence are sometimes ambiguous and open to alternative interpretations, an advanced course on intelligence provides one of the best ways to teach this critical lesson. The most satisfying evaluation I have ever received for my own course on intelligence is from the student who said that, unlike most of his other courses, it gave him the opportunity to think for himself.

References

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, *15*, 31–36.

Ceci, S. J., & Liker, J. K. (1986). A day at the races: A study of IQ, expertise and cognitive complexity. *Journal of Experimental Psychology General*, *115*, 255–266.

Deary, I. J., Irwing, P., Der, G., & Bates, T. (2007). Brother–sister differences in the *g* factor in intelligence: Analysis of full, opposite-sex siblings from the NLSY1979. *Intelligence*, *35*, 451–456.

Deary, I. J., Th orpe, G., Wilson, V., Starr, J. M., & Whalley, L. J. (2003). Population sex differences in IQ at age 11: The Scottish Mental Survey 1932. *Intelligence*, *31*, 533–542.

Detterman, D. K., & Spry, K. M. (1988). Is it smart to play the horses? Comment on 'A day at the races: A study of IQ, expertise, and cognitive complexity' (S. J. Ceci and J. K. Liker, 1986). *Journal of Experimental Psychology General*, *117*, 91–95.

Dickens, W. T., & Flynn, J. R. (2006). Black Americans reduce the racial IQ gap: Evidence from standardization samples. *Psychological Science*, *17*, 913–920.

Fine, C. (2010). *Delusions of gender.* London: Icon Books.

Friedman, N. P., Miyake, A., Corley, R. P., et al. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and Math. *Science*, *320*, 1164–1165.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*, 41–45.

Hunt, E. (2011). *Human intelligence.* New York: Cambridge University Press.

Kamin, L. J. (1974). *The science and politics of IQ.* Potomac, MD: Lawrence Erlbaum.

Machin, S., & Pekkarinen, T. (2008). Global sex differences in test score variability. *Science*, *322*, 1331–1332.

Mackintosh, N. J. (1975). Critical notice: Kamin, L. J., The science and politics of IQ. *Quarterly Journal of Experimental Psychology*, *27*, 677–686.

Mackintosh, N. J. (2011). *IQ and human intelligence* (2nd ed.). Oxford: Oxford University Press.

Murdoch, S. (2007). *IQ: The brilliant idea that failed.* London: Duckworth.

Murray, C. (1998). *Income, inequality, and IQ.* Washington, DC: American Enterprise Institute.

Neisser, U. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Nisbett, R. E. (2005). Heredity, environment, and race differences in IQ: A commentary on Rushton and Jensen (2005). *Psychology, Public Policy, and Law*, *11*, 302–310.

Penner, A. (2008). Gender differences in extreme mathematical achievement: An international perspective on biological and social factors. *The American Journal of Sociology*, *114*, S138–S170.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235–294.

Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist*, *63*, 215–227.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy: The media and public policy.* New Brunswick, NJ: Transaction Publishers.

Strand, S., Deary, I. J., & Smith, P. (2006). Sex differences in cognitive abilities test scores: A UK national picture. *British Journal of Educational Psychology*, *76*, 463–480.

Please cite this article as: Mackintosh, N.J., Why teach intelligence?, *Intelligence* (2013), http://dx.doi.org/10.1016/j.intell.2013.08.001