

Paranormal IQ Scores?

Andrew M Colman and David D Stretch

A look at assumptions underlying psychometric tests

IN A RECENT ARTICLE in *The Skeptic* [1], an argument was put forward that 'the world's highest IQ' of 228, claimed on behalf of the American writer Marilyn vos Savant, is so improbable as to be unbelievable, and that the claim can therefore be refuted on purely logical (or mathematical) grounds. The argument hinged on the modern statistical conception of IQ, according to which IQ is, by definition, a normally distributed variate with a mean of 100 and a standard deviation of 15. Marilyn vos Savant's alleged IQ of 228 falls 8.53 standard deviations above the mean, and the probability that a sample of a little over 5 billion (the world's population, approximately) drawn from this distribution contains any scores of 228 or above was shown to be less than one in 20 million, which is effectively zero.

Criticisms of this argument have appeared in letters to *The Skeptic*. It is worth reminding readers that the onus clearly falls on those who make or support extraordinary claims to prove these claims, but both letters to *The Skeptic* tried to shift the burden of proof on to the writer who questioned the claim. The specific criticisms could be answered quite briefly in a letter, but we believe that they rest on debilitating though widespread misconceptions that have more general implications in all areas of psychometrics and deserve more thorough examination.

Ratio and Statistical Conceptions of IQ

Hugh Farey [2] criticized the argument against the IQ of 228 on the grounds that the score was derived from an older conception of IQ as the ratio of mental age to chronological age expressed as a percentage (that is, multiplied by 100): 'Miss vos Savant achieved the score expected of a 23-year old (almost) when she was ten, giving a quotient, as it was then defined, of $22.8/10 \times 100 = 228$ '. Farey's formula is mathematically ambiguous; for the sake of clarity it is worth pointing out that what he presumably meant was $(22.8/10) \times 100$, and not $22.8/(10 \times 100)$, which yields an IQ very close to zero and indicates profound mental retardation rather than genius. Farey also reported that Marilyn vos Savant was subsequently tested on a different IQ test and achieved a score equivalent to an IQ of about 175. He concluded that Marilyn vos Savant's claim 'is perfectly acceptable' and the criticism of it 'merely meretricious'.

It is always a pity when writers descend into *ad hominem* attacks on people who present views with which

they disagree. The word *meretricious*, which means 'befitting a harlot' (from the Latin *meretrix*, a harlot) is, in any event, misleading since no money was paid for the article, and it unfortunately tends to distract attention from the ideas themselves. The ideas are worth examining, although we shall show that they do not stand up.

The linchpin of Farey's argument is the suggestion that Marilyn vos Savant's extraordinary score was derived from the old *mental age divided by chronological age* conception of IQ, 'as it was then defined', when she was ten years old. But it is quite wrong to say that IQ 'was then defined' in that way. In fact, the statistical definition was introduced in 1939, before Marilyn vos Savant was born, in an extremely influential monograph by David Wechsler [3]. Farey asserted that after 1939 'for children, the original definition continued, and I believe continues, to be widely used', but he did not disclose on what evidence his belief was based.

To operate with two completely different conceptions of IQ, one for children and another for adults, would be extraordinarily bad science (or, more accurately, pseudoscience) unless of course some evidence came to light that child and adult intelligence were two qualitatively different phenomena. We are not aware of any evidence for such a belief.

Child and Adult IQ

Although it seems a priori improbable (to say the least), let us suppose just for the sake of argument that child and adult IQ are indeed qualitatively different phenomena, and that it is reasonable, as Farey suggested, to define children's IQs using the old ratio formula and adults' IQs according to the modern statistical definition. In that case the claim that Marilyn vos Savant's IQ is 228 collapses, because all that can be claimed is that she *had* an IQ of 228 when she was ten years old, not that she *has* such an IQ now.

On this point, Farey supplied empirical evidence that effectively demolished his own conclusion that 'her claim is perfectly acceptable' when he revealed that she more recently achieved a score equivalent to an IQ of 175 on a different test. This is strongly suggestive of what statisticians call *regression to the mean*, arising from measurement error, and such a large regression to the mean can often arise from the original IQ score having low reliability. Since the scores of 228 and 175 cannot both be valid,

unless Marilyn vos Savant has suffered a catastrophic fall in intelligence, probably unprecedented in the absence of any severe neurological disorder, common sense encourages us to accept the lower IQ, which is at least within the bounds of statistical possibility.

In summary, the assumptions and facts on which Farey based his argument are ultimately self-defeating, because even if they are accepted they show that Marilyn vos Savant's IQ score of 228 is unworthy of belief.

The Distribution of IQ

Martin Hempstead defended Marilyn vos Savant's IQ from a different angle [4]. He criticized the very basis of the statistical definition of IQ: 'you can tailor an IQ test to give a normal distribution, you cannot *define* it thus, and you can only tailor it to an accuracy given by the sample size being used for adjustment' (italics in original). He went on to assert that 'the IQ distribution cannot be normal, anyway, because IQ results are bounded from below, by zero, and from above, by the maximum test score, unlike the normal distribution'. Finally, he proposed a model in which one in a billion people are geniuses, 'perhaps due to genetic mutation', their IQs being drawn from a normal distribution with a mean of 210, not 100; 'such a population component would never show up in mass testing'.

There are several fundamental misconceptions in Hempstead's argument. First, IQ tests do not have to be tailored to give a normal distribution; on the contrary, it is impossible to stop them doing it. It is a remarkable fact of nature that IQ scores, like many other psychological and physical characteristics that are attributable to multiple causes, are approximately normally distributed. The explanation for this is provided by a theorem in mathematical statistics due to Lindeberg [5] called the *central limit theorem* (see [6] for an exposition of the relevant version of the basic theorem and its implications). According to the central limit theorem, provided that certain minimal conditions are satisfied, the sum of a sequence of mutually independent random variables approaches a normal distribution as the number of random variables increases—and in most cases the rate of approach is quite rapid.

If we assume that a person's IQ score is determined by numerous independent random variables, some no doubt determined by genetic inheritance and some by environmental or purely chance factors, then the central limit theorem applies and the cumulative effect of these causal factors will be approximately normally distributed. As one leading authority expressed it, 'the central limit theorem explains why many biometric characters, like height, exhibit an empirical distribution close to the normal distribution' [7].

Hempstead asserted that the IQ distribution cannot be normal because, unlike the normal distribution, IQ scores are bounded from below by zero, and from above by the maximum test score. This is to confuse an underlying variable with the instruments that are used to measure it: it is logically equivalent to saying that temperature is bounded below by the freezing point of mercury and from above by its boiling point. Temperature does not have to be measured with a mercury thermometer, and physicists have

devised different measuring devices to measure lower and higher temperatures. As with IQ, the underlying variable is not bounded by the limitations of any particular measuring instrument. Temperature is, in fact, bounded from below by absolute zero, but this has nothing to do with the limitations of the measuring instruments, and IQ, in its statistical conception, is theoretically unbounded from below and above, despite the limitations of current IQ tests.

Ptolemaic Epicycles

Finally, Hempstead proposed that one in a billion people 'who would never show up in mass testing' are geniuses whose IQs are drawn from a normal distribution with a mean of 210. This is a classical example of an *ad hoc* hypothesis designed to accommodate an errant observation. It is strikingly reminiscent of Ptolemaic epicycles, which were postulated to explain the observed deviations of the orbits of some celestial bodies from perfect circles before Copernicus introduced a heliocentric astronomy in the 16th century. There is not a shred of evidence for any intellectually elite population component that falls outside the normal distribution in which the rest of us languish.

In any event, if these postulated geniuses never show up in mass testing, then the standardization of IQ tests has not taken them into account and the tests may be quite likely to be invalid for measuring their IQs. This means that even if Hempstead's *ad hoc* postulate were to be accepted, far from justifying Marilyn vos Savant's claimed IQ of 228, it would provide yet another argument to undermine it.

Conclusions

1. A claimed IQ of 228 is sufficiently improbable to be described as extraordinary.
2. In an earlier article, Colman [1] argued that there were grounds for disbelieving the extraordinary claim.
3. In reply, two critics have attempted to shift the burden of proof away from those making the extraordinary claim and on to the commentator raising doubts about it.
4. Additionally, *ad hominem* arguments were introduced into the debate.
5. Arguments that were advanced to bolster the extraordinary claim involve a mistaken identity relationship between a measurement method and the underlying concept being measured; an unsupported assertion that two completely different methods of deriving IQ scores are used for children and adults (which leads to a later inconsistency); and a further suggestion, with little supporting evidence, that there exists a population subgroup of genetically determined geniuses that would make the claimed IQ understandable.
6. The arguments advanced showed inconsistencies and often incorporated further, unsupported elaborations designed to shore up the extraordinary claim. Indeed, some of these provided yet further grounds for doubting the extraordinary claim. Consequently, the extraordinary claim should continue to be viewed with suspicion.

References

1. Colman, A M, A supernatural IQ? Investigating a claim to an extraordinary IQ. *The Skeptic*, 7(5), 12—13, Septem-

ber/October 1993.

2. Farey, H, A fair claim [letter]. *The Skeptic*, 7(6), 27, November/December 1993.

3. Wechsler, D, *The Measurement of Adult Intelligence*, Baltimore, MD: Williams & Wilkins, 1939.

4. Hempstead, M, A call for care [letter]. *The Skeptic*, 7(6), 27, November/December 1993.

5. Lindeberg, J W, Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 15, 211—225, 1922.

6. Feller, W, *An Introduction to Probability Theory and Its Applications*, New York: Wiley, 1968. The Central Limit Theorem and its close relatives are discussed in Volume 1, pp. 243—256.

7. Feller, W, *op. cit.*, p. 256.

Andrew M Colman is a Reader specializing in mathematical and social psychology in the Department of Psychology, University of Leicester, and **David D Stretch** is a mathematical psychologist specializing in statistical modelling and research design in the Greenwood Institute of Child Health, University of Leicester.

SKEPTICAL CONTACTS

U.K. Skeptics

Committee: Susan Blackmore, Steve Donnelly, Wendy Grossman, Ian Rowland, Chris Nash, Mike Howgate, Richard Mather, Michael Hutchinson.
10 Crescent View, Loughton, Essex, IG10 4PZ.

London Student Skeptics

Convenor: Bill Harman
21 Manville Road
London, SW17 8JW.

Wessex Skeptics

Secretary: Martin Hempstead
Optoelectronics Research Centre
Southampton University, Highfield
Southampton, SO9 5NH.

Electronic Mail

For information on skeptical information by E-mail or for subscription, back issues, or other magazine enquiries please contact:

Toby Howard: toby@cs.man.ac.uk

The Skeptic now also has World Wide Web page, at: <http://www.cs.man.ac.uk/aig/staff/toby/skeptic.html>