Labelled for Life?

Chris N French

We live in a busy world. I am pleasantly surprised you have found time to begin reading this short article. I hope you will be able to finish it! I was reading the newspaper last September and was astounded to find the broadsheet Independent *(and other newspapers)* report *"'Children suffer' if singled out as gifted. Children who are labelled as gifted by their ambitious parents are more likely to grow into unhappy adults than equally bright pupils who are not singled out."* The article was illustrated with pictures of a couple of your typical, average, everyday gifted children - Ruth Lawrence who gained a first class degree at Oxford at age 13 and Sufiah Yusof who ran away from Oxford at age 15 quoting "15 years of physical and emotional abuse".

Commonsense told me that labelling a child as 'thick' or 'dull' was certainly not likely to increase their happiness. Does that then mean that with 'gifted' also ruled out, all children in future were to be labelled as 'average' or 'normal'. Personally I wouldn't have fancied being labelled as 'normal'. Must all labelling therefore be abolished? Anyway, what is this 'labelling' thing!! And how much more likely are these children to grow into unhappy adults? Is it 50 per cent more likely or 10 per cent or is it simply that there is a weak but statistically significant tendency for this to occur? Is this article only going to ask questions?

We all like to be able to depend upon expert opinions, but when they say something strange or apparently outrageous it is unsettling. Is it for the oxygen of publicity? In a society where communications are all too often byte size what do we do? I find it particularly irritating when there is, as there was in this case, insufficient information to enable one to corroborate or evaluate the evidence for oneself. I, therefore, decided to dig deep even if it might take a little time.

The report came from a British Psychological Society (BPS) Developmental/Education Conference, from a paper entitled Following Gifted Children into Adulthood by Professor Joan Freeman (http://www.joanfreeman.co.uk/). This apparently reported on work recently published in a new (2001) book entitled Gifted Children Grown Up[1]

GIFTED CHILDREN GROWN UP

The research study began in 1974 and concerned a longitudinal study of 70 children then aged between five and 14 whose parents had joined NAGC in the north west of England *(61 schools in the Liverpool and Manchester area).* Each pupil was matched with two pupils who were in the same class at school and were the same age and sex. One had the same IQ as the gifted child *(matched according to Ravens' Progressive Matrices)* while the other was selected at random.

Longitudinal studies are rare and valuable and in this case maybe unique. They surely have the potential for providing fascinating insights into the development process. Even those pop studies that appear on TV every seven years have a mesmerising effect. So it was with eager *(albeit belated)* enthusiasm that I bought a copy and caught up on the research. There had been previous publications and this, the third book, marked 27 years of the study with those children now left in the research being adult, potentially aged between 32 and 41 years of age.

The book is an entertaining read and there is much to recommend it. It is particularly interesting to read how individuals's lives have turned out, but it wasn't long before I found myself deeply aggravated.

IQ GRUMBLES AND GRIPES

Throughout the book the author refers to IQs numerically but in a thoroughly confusing manner. I have to admit she is not alone in doing this. The book is written for the general reader but IQ, a concept central to the subject, is poorly explained.

Deviation vs Ratio IQs

Most of the time the IQ references appear to be to Stanford Binet test scores and one might expect, from the initial, 1974, date of the study that this would be to the then most recent

although then 14 years old, 1960, revision of that test, but for some reason she appears to have used the 1937 version, already then 37 years old! Why?

The concept of Intelligence Quotient (IQ) as being Mental Age (MA) divided by Chronological Age (CA) is one that went out of use many years ago, but the reader is not told this. In fact it was a major innovation of the 1960 Stanford Binet when ratio IQs were substituted by deviation IQs[2]. Reference to ratio IQs suggests she used the 1937 test, as do some of the figures on page 15 & 16 of her book. But one shouldn't need to play at detective looking for clues! Which did she use?

IQ Standard Deviations

The average IQ for a population is 100 whatever its age, but any individual IQ score is meaningless unless you know what the spread or standard deviation (SD) is in that population under consideration. You need the SD to tell you how far above or below average an IQ score is. With the 1937 Stanford Binet the SD actually varied substantially with age, from 13 to 21, averaging around 17. Modern IQ tests typically have standard deviations of 16 (eg 1960 Stanford-Binet) or 15 (eg British Ability Scales). Clearly the larger the SD the higher the IQ for the same 'intelligence' performance. *(It remains unclear to me which version of Ravens' Progressive Matrices was used, but as these scores are apparently quoted less often this is probably less important.)*
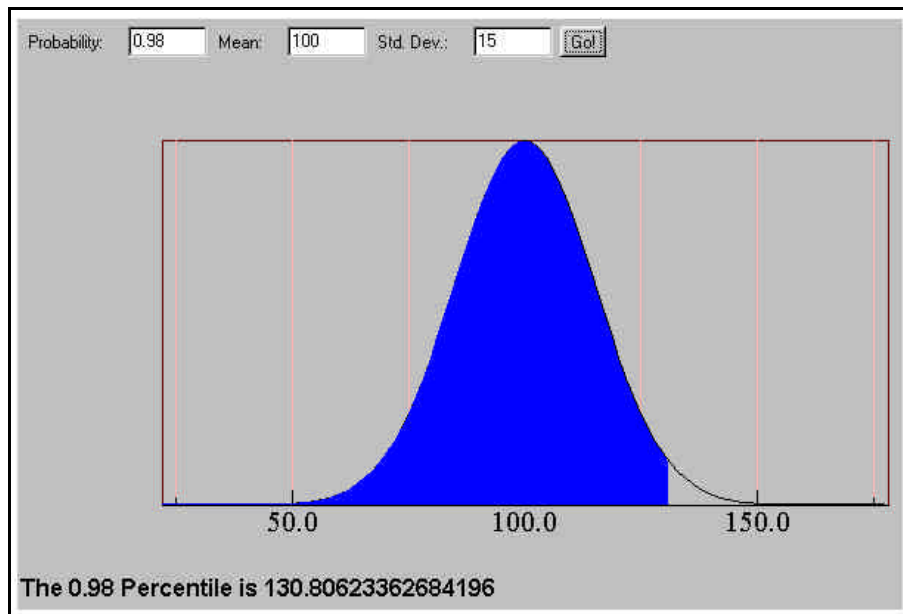
A child taking the Stanford-Binet in 1937 with a calculated IQ of 140 would have been placed in the top 1 per cent of the population. The

equivalent score on a 1960 S-B test would be 136 and on the BAS 133! An old version of Raven's Advanced Progressive Matrices and the old Cattell IIB both had standard deviations of 24. Thus the equivalent score on these tests would have been a whopping 153. Thus, without further information, an IQ score on its own is pretty meaningless!

For example, when you read of TV personalities having stratospheric IQs take them with a pinch of salt! An old-style Mensa test score of 178 would be equivalent to 152 on a modern Stanford Binet and 149 on the BAS. None of this is touched upon in the book.

IQ Inflation, the Flynn Effect
In fact things are more complicated even than this. One wonders what defining norms would have been

Web page output.

IQ Scores and Percentiles as a Function of Population Spread (SD)

SD	examples of tests with given standard deviation (SD)	percentiles for the Normal distribution							
		90	95	98	99	99.9	99.99	99.999	99.9999
15	British Ability Scales I & II	119.2	124.7	**130.8**	134.9	146.4	155.8	164.0	171.3
16	Stanford Binet (1960)	120.5	126.3	132.9	137.2	149.4	159.5	168.2	176.1
17	SB (1937) - aggregate sd	121.8	128.0	134.9	139.5	152.5	163.2	172.5	180.8
24	Cattell IIIB (Mensa)	130.1	139.5	149.3	155.8	174.2	189.3	202.4	214.1

For example, with the BAS tests 98 per cent of scores would be expected to be below **130.8** (in bold above) and two per cent above. For other values consult the web site quoted below. Note that all IQ scores have integer values (none have a decimal point!); that the accuracy (confidence limits) of scores vary so that there is a range (better 'confidence limits') within which the true score is likely to lie. So, for example, a general IQ score of 131 might have confidence limits of (say) 126 to 133. For this reason, a single assessment should be taken with caution. Note also that many tests have low ceilings and don't claim to be able to discriminate at the upper levels. For example BAS (I) covered scores from 51 to 149, while with BAS II the upper limit varies from 159 to 162 according to the subject's age. Leaving aside statistical matters, direct comparisons between tests near their upper limits need to be taken with a pinch of salt! Don't take the right columns very seriously! With any educational assessment the educational psychologist will be able to clarify these matters as part of the consultation process. See: http://www-stat.stanford.edu/~naras/jsm/FindProbability.html

available for the 1937 test in 1974. With the 1937 replaced by the later version in 1960 would there have been any up-to-date norms for the UK? For that matter what norms were available for the 1960 version? This in not an academic question as since the 1980s it has become apparent that IQ scores have been increasing since the beginning of the 20th century - the Flynn effect, named after its discoverer. Children did much better on the old tests than they did on the new ones. Everywhere Flynn looked, he noticed that groups performed much more intelligently on older tests. In 1984 he reported that Americans had gained about 14 IQ points in 46 years[3]. Professor Freeman does touch even more briefly than above upon this IQ 'inflation' but not directly on its implications for her apparent use of an out-of-date 1937 test. Were her norms up to date?

Achievement or Potential IQs
Things get even more complicated because no two IQ tests measure the same thing. Professor Freeman is upfront in pointing out that the Stanford Binet is more of a test of academic achievement than potential. And in fact this and the wealth of data associated with it was an excellent reason for using it. But it does need to be emphasised, as others have pointed out, that it also has not worked well for identifying people who are gifted when its results are compared to those of other tests[4].

Full Information
From this you can see it is very important with IQ scores that, unless you want to treat them simply as primitive ordinal scale numbers, you should quote the precise test being used. If there is any ambiguity then information should also be given on the norms used to interpret that score and, to make things as clear as

possible, the SD should always be given. Reporting scores directly in percentile terms as well makes for even greater intelligibility.

Doubts
It is fair to say that the above comments on IQ probably do not invalidate any of Professor Freeman's conclusions, but for me what they do is unsettle, reduce confidence and plant the seeds of doubt about what else she has to say.

LABELLING HARMFUL?
So where is the evidence that labelling children as having high ability or being gifted is harmful? I couldn't find it. I did write to her as follows *"I was intrigued by the press coverage of your presentation in early September. This appeared to go beyond the book's conclusions (sic?) reporting that children 'labelled' as gifted in 1974 had a tough time because of the labelling? My own hypothesis was*

that there is probably a significant greater tendency for both the parents of 'difficult' bright children and the 'difficult' parents of bright children to join NAGC as part of their process of seeking help. Thus if their children remain 'less happy' than the average bright child then this was to be expected."

I received a polite reply which was very nice *(so many people wouldn't have bothered)* and she concurred with my observation on why some of us might have joined NAGC emphasising that *"obviously, this does not apply to all NAGC members"*. She had in fact researched these reasons for joining in another book.[5] However, she made no comment on my hypothesis. Nor did she repudiate the newspapers' reporting. I followed this up with a further enquiry but the gist of her response this time was that she was far too busy to go into any detail... and I was referred to her PhD thesis.

Despite this I subsequently checked the BPS web site (http://www.bps.org.uk/press/press.cfm?action=details&id=346) in case newspapers had distorted her views. There I found *"Labelling children as gifted can lead to difficulties in their emotional functioning."* This is a weaker statement than the one read in the newspaper. 'Can' means 'potentially capable'. Well, yes, I might go with that. It is not an unreasonable hypothesis. I can imagine certain circumstances where labelling a child as gifted when combined with certain other circumstances or parental behaviour *can* lead to unhappiness with some children. But it is quite different to the bold statement *"Children suffer if singled out as gifted"* and is potentially so weak as to be almost worthless.

I still have problems with even this weak hypothesis. What evidence is there in the book which supports it? There appear to be no numbers.

What evidence could the study have uncovered for the effects of labelling alone? All cases of gifted labelling are confounded. They are accompanied by other circumstances and it is impossible to isolate the two. The principle of parsimony demands that the simplest, most mundane explanation is to be preferred. In 1974 some of the people who had joined NAGC were attracted there because they had problems and thought that the organisation might help solve them. No doubt it did, but 27 years later it shouldn't have come as a surprise if on average these children, now grown-up, should have more

problems than the control group who were reported to have fewer problems back in 1974.

There is no reason to invoke labelling or singling out. Maybe this is where some of the adults looking back on their lives put the blame but that isn't scientific evidence. It is just 'rationalization' - the process of concocting plausible reasons to account for one's practices or beliefs.

Despite this, it might have been interesting to have had even this information quantified. Only 19 per cent of the original sample was lost between 1974 and 1984. However, we are not told how many more were lost by the time this book was written and published in 2001. How big is the sample on which these conclusions are now based? What are the numbers? Which statistical tests were used and what were the results? I am afraid I wasn't able to find out.

Of course it is also questionable whether what happened in 1974 in NAGC circles is comparable to what is happening to day. Yet the reader is not cautioned against extrapolating any conclusions.

Maybe I would hypothesise that being a 'too pushy parent' is not a good thing as far as a child is concerned. In much way that being a 'too unsupportive parent' is unhelpful. But I doubt whether either of these are true hypotheses or true empirical statements!

It is inevitable that we as people will resort to labelling. Labelling is part of what we do to make sense of our world. If the labelling is correct what can be the harm in it? I have seen no evidence to the contrary. It is my belief that self-knowledge is a good thing. It is better to know one's IQ and other abilities than to be ignorant... just as long as one fully appreciates that an IQ is not the be-all or end-all of a person- motivation and character are also very important.

Objective measurement and with it the inevitable labelling has to be far better that subjective, uncertainty and illusions.

CONCLUSIONS

Gifted Soap Stories
What is good about the book are the anonymous accounts of individual case histories. It is an entertaining read and no doubt a fertile source for hypotheses. I have no reasons to deny that and I have considerable sympathy with many of Professor Freeman's observations, particularly the way she questions established beliefs.

But beware of psychologists bearing strange results on gifted children. You have been warned. Keep that sceptical frame of mind... even with the broadsheets!

REFERENCES

[1] Freeman, J, Gifted Children Grown Up. David Fulton Publishers; ISBN: 1853468312

[2] Psychological Testing, A Anastasi, 2nd Edition, 1961 New York: Macmillan, .

[3] You can find much on the Flynn Effect and the strengths and weaknesses of different IQ tests by searching the Web.

[4] Psychology, L T Benjamin, J R Hopkins, J R. Nation, 3rd Edition, 1994, Macmillan, ISBN: 0023082909

[5] Freeman, J (1979) Gifted Children; their Identification and Development in a Social Context. Lancaster: Medical Technical Press.

Author: Chris is an experimental psychologist and retired university lecturer who taught psychology and statistics.