

On the history of “IQ” and aptitude testing – with specific relation to the Kent Test

Dr. Michael Collins, University of Kent

1 Background to establishment of Tripartite System

1.1. Following the conclusions of the Butler report, the subsequent 1944 Education Act established a Tripartite model of education in Britain, with secondary-level schooling divided into grammar, secondary- modern and technical strata. According to the Act these schools were designed in response to the needs of an increasingly diverse British workforce to be separate but equal, teaching curricula specific to the “aptitudes” of each child. Pragmatism dictated that schooling would be divided between primary and secondary at 11, despite no compelling scientific evidence for the choice of this age rather than any other. Unsurprisingly, however, measuring these “aptitudes” was a difficult task. In 1944 neither psychometric testing according to the Stanford- Binet intelligence quotient (IQ), nor curriculum-specific, knowledge- based tests were understood to be wholly acceptable and infallible models of selection. Nonetheless, the 11+ was established as the standard model of educational selection until the late 1970s, when the rise of comprehensive schooling models began to challenge the older “ability” tiers of the Tripartite System in Britain.

1.2 In Kent, grammar schools remain a major element of the educational landscape. In response to Kent County Council Select Committee on Grammar Schools and Social Mobility’s request for responses to their inquiry into getting higher numbers of children on FSM into selective schools I wish to offer the following history of the 11+ plus test, its aims, objectives and outcomes in the hope that the KCC will seriously consider the implications of selective testing at 11. Towards the end I offer some suggestions about how social mobility might be better served. In this paper, I would like to show with specific reference to the Kent Test how the assumption underpinning this model of assessment - that the measurement of “ability” or “intelligence” at 11 is possible and desirable - is a fundamental fallacy that represents an outmoded inheritance from an earlier model of science that has been largely debunked. To do this I will focus on two key areas:

- a) The history of the 11+ and its relation to debates about fixed “intelligence”;
- b) The role of such a measure in the age of a national, as opposed to differentiated, curriculum.

2 The idea of “intelligence”

2.1 To explain why measurement proved so difficult in the early years of the 11+, we must look deeper into the origins of the idea of “intelligence” as a fixed, reified entity. By the late 1910s tests designed by Alfred Binet for the purposes of locating what the influential psychologist and U.S. school reformer Henry Goddard called “feeble- mindedness” in discrete populations

(ostensibly for the purposes of specialised local education policies for those with learning difficulties) had been fused with Francis Galton and James McKeen Cattell's eugenical studies into "hereditary genius" and applied to increasingly large data sets. The rise of demographic studies that allowed for the collection of such data was made possible by the mass- mobilisations of troops during the First World War in Europe and America, but was extended in the post-war era through the consolidation of the welfare state from the 1940s. The Stanford-Binet intelligence test designed by Lewis Terman is the most famous model of this synthesis of disparate areas of inquiry.

2.2. The first IQ tests were measures of neural dexterity (Cattell's tests were of speed of comprehension), but, throughout the 1910s, the tests became synonymous with academic potential. However, these are fundamentally different measures, as was frequently pointed out and can be seen in public debates in the 1914 issues of *Science* between Dr. David Heron and Dr. Charles Davenport. The effect of this synthesis of studies in mental disability with studies of "hereditary genius" was the production of a standard-deviation model of intelligence in a population, which seemed to suggest that individuals had fixed "intelligence" and such abilities conformed to a standard bell-curve distribution through a population. However, as Lelia Zenderland has shown in *Measuring Minds: Henry Herbert Goddard and the Origins of American Intelligence Testing*, the precise nature of this "intelligence" was the subject of serious debate.

2.3. Since the beginning of intelligence testing in the late nineteenth century scientists have continually questioned whether IQ tests actually measured "intelligence" (a quasi-mystical "state of nature" that could express itself free from an individual's social milieu and training) or "scholastic ability" (a measure that would comprise such things and be derived largely from their impacts). Indeed, in the social sciences, hard sciences and humanities, debate continues as to whether something called "intelligence" (free from upbringing, social context and training) is a real, material and measurable entity. Major psychologists and evolutionary biologists working in the last 30 years, such as Stephen Jay Gould, have questioned the validity of the data around innate human intelligence and demonstrated that it has often been misinterpreted for political aims throughout its history (see Gould, *The Mismeasure of Man*). This has led to a state where, according to a 2012 longitudinal study of IQ in Britain after the 1944 Butler Report and Education Act, "Schooling in Adolescence raises IQ scores", which was published in the *National Academy of Sciences of America* and verified by Harvard University, it is possible to state with some confidence that "a growing consensus points to the major role that early childhood environment and interventions play in the development of economically and socially relevant cognitive skills..." (Brinch and Galloway, 425).

2.4 Furthermore, it is also now possible to argue that "education occurring even as late as in the middle teenage years can indeed have a statistically significant effect on IQ scores" (425), questioning the assumption that any model of academic ability or intelligence is fixed at 11. What current academic studies show with reference to a considerable body of research in the years

since the Butler Act is that the “intelligence” is neither fixed at 11 nor universally measurable in a manner that is free from education, parental support, financial background and other factors etc. even through to the middle teenage years and beyond.

2.5 This is not a new story. As early as the 1910s, scientists were unsure that “intelligence” was a historical absolute that was measured simply by a test. By the 1950s heyday of the 11+ it was even less certain. A 1954 study conducted by Alice Heim and published by the National Foundation for Educational Research in England and Wales, *The Appraisal of Intelligence*, concluded that social class and upbringing shaped significantly the potential likelihood of an individual achieving a high IQ score. It was well documented in the middle decades of the twentieth century that IQ and knowledge-based testing at 11 discriminated against those in the lowest socio-economic brackets, while actively benefitting middle-class children. Nonetheless, the 11+ continued to be administered in Kent, maintaining the fallacy that policy was built on a certainty. As a 2012 study by Hart, Moro and Roberts at The University of Stirling notes of the 1950s and 1960s:

a view emerged that the nature of IQ testing under the 11 plus exam was itself not independent of family circumstances. Criticisms of the methods of measuring intelligence began to emerge... with sociologists pointing out ‘the influence of intelligence tests in discriminating against working-class children at eleven-plus’ (Simon and Rubinstein, 1969) (10)

2.6 Indeed, the lack of firm scientific evidence of fixed human intelligence was considered when the 11+ was first implemented in Britain, but governmental will to establish an “intelligence”-based model triumphed for reasons of political and economical expediency. The idea that “intelligence” is innate, fixed and real remains a truism – put down to common sense by many – despite overwhelming contemporary evidence to the contrary. Precisely because of the uncertain facticity of “intelligence”, the initial version of the 11+ test was composed of a mixture of Stanford-Binet IQ reasoning tests with scholastic knowledge that was deemed pertinent to the specific curriculum the child would follow in secondary education. The desire to stratify according to labour need trumped the uncertainty of the science on this issue.

2.7 Early formulators of the test included a knowledge-based component in assessment at 11 (grammatical skills, punctuation, knowledge of famous Shakespearean soliloquies) not out of a wish to specifically select middle-class children for grammar schooling, but precisely because IQ was so uncertain a measure that other elements were deemed necessary to justify the test’s assessment. Despite the best intentions of some of the framers, the effect of this fusion of IQ and knowledge-based testing in the model of the 11+ was that the test served to ultimately reinforce existing class hierarchies. Far more than measuring “intelligence”, or being child-centred in its intentions the 11+ test was designed to hierarchically differentiate by labour demand. Hart, Moro and Roberts note that:

Soon after the reforms became operational, evidence emerged that children with middle-class parents performed especially well in the 11 plus exam (see Simon and Rubenstein), Not only did middle-class children appear to have higher probabilities of attending grammar schools compared to their working-class contemporaries but that they also seemed to be better suited to the aims and objectives of grammar school education” (9)

2.8 Some social mobility did occur through grammar schools, a significant boon for those that benefitted, but this was roughly equal to that which had occurred at a national level in the years prior to the implementation of the Tripartite System. Consequently, there was little overall effect to the tests bar a systematisation of resource allocation at a national level. As Hart, Moro and Roberts expressed this:

The biggest gainers from the free education provision were children from relatively disadvantaged backgrounds who gained competitive entry into the grammar school system. These constituted only about 15 per cent of all children attending tripartite schools. A further 20 per cent were from more advantaged backgrounds and a high proportion of these may well have received a grammar school education in the absence of the new education policy. For the large majority of the remainder who were required to attend secondary modern schools, the policy served generally to stifle educational and post- educational development and this in turn was reflected in relatively poor subsequent labour market outcomes” (25)

2.9 Overall, even in the presumed heyday of the Tripartite System in the late-1950s and early 1960s, the percentage of relatively disadvantaged children who attended grammar schools after passing the 11+ was only 15 per cent nationally. This was in spite of the intention that the 11+ test select approximately 30 per cent of the population for grammar schooling. Given the high numbers of families in the working class than the middle class in the immediate post-war era, this is shocking. For working-class and low income families grammar schooling in the middle decades of the 20th century failed to create significant movement into the middle class. Indeed, the unequal provision of resources and teachers in the secondary-modern schools meant that a significant result of the Tripartite System was “relatively poor subsequently labour market outcomes” – in other words, growth and upward social mobility.

2.10 Following these sociological observations, advocates of the 11+ test attempted to modify it to be more IQ-based in the 1960s, responding to criticism that the test was aimed at cultivating what Pierre Bourdieu would call “bourgeois habitus”, the knowledge, culture and habits of being peculiar to middle-class life. These reforms towards an IQ-based test were similarly unsuccessful. By the 1970s it was quite possible to claim that grammar schools had failed to produce social mobility on any significant scale. The problem lay not in the curriculum per se but the assessment that allowed entry to grammars. The 11+ was neither an IQ test in a clear sense, nor a test of knowledge.

2.11 The existence of the Kent County Council Select Committee on Grammar Schools and Social Mobility in our current moment points to a similar pattern of poor social mobility through grammars. KCC's own statistics show that children on Free School Meals constitute a significantly lower population in grammar schools than in other types of school. The reason for this is the Kent Test itself. Alternatives are not easy for advocates of grammar schools to envisage. A pure IQ test would be discriminatory, since IQ itself is socially conditioned and can (as Brinch and Galloway have shown) be tutored for. A pure knowledge-based test would also benefit only those who can be tutored, have excellent primary schooling, and/or stable, economically secure and supportive family backgrounds. A combination of both IQ and knowledge-based assessment would serve no clear positive purpose in altering rates of social mobility, since both models benefit the middle-classes unfairly.

2.12 Increasingly, selective grammar schools offer free education of the kind given already by private schools to children whose parents could often afford private schooling or tutoring, while cutting out large percentages of the population from access to equal resources and provisions. The Kent Test allows for the further accumulation of capital in the middle and upper-middle classes off the back of the state, since the cost of schooling is retained by the middle classes and not spent, while resources are placed into high quality schooling of the already wealthy. The evidence of extensive tutoring and the use of private primary schools indicates that middle-class parents are more than willing to pay for their children's schooling at junior level in order to qualify for free secondary school. In essence this transforms selective schooling into a benefit offered to those who are least in need of that benefit. A government that wishes to cut social welfare provision should think hard about offering that welfare vicariously to the middle classes through educational capital when doing so is clearly prejudicial to those without sufficient social, economic or cultural capital to successfully qualify for support.

2.13 Much evidence now suggests that IQ itself is not a measure free from social environment and training. However, the pattern of static or diminishing social mobility outlined above is powered through the setting of the Kent Test that attempts to evaluate IQ and knowledge simultaneously. Furthermore, neither one's IQ nor one's knowledge-base is unaffected by tutoring, background and class. This brings me to my second area of inquiry.

3 The new curriculum

3.1 In our current climate of the national curriculum, the existence of the 11+ no longer indicates a test of a person's ability within the terms of a specific curriculum (as it did in the days of the Tripartite System), since this curriculum is now universal. Consequently, children who fail the 11+ are not now taking courses that the test would claim are suited to their interests and abilities, but taking the same curriculum as those in grammars, albeit from the

perspective of having been told that they cannot excel at it. This de-motivation at 11 could have long-term impacts on the educational attainment of those who are not selected for grammar school by the Kent Test. I would suggest that KCC commission research into the achievements of those who take the Kent Test and are not selected for grammars. Little evidence is extant on these individuals, yet they serve as an important test case for the effects of the Kent Test on social mobility. Given the voluntary nature of the test, the decision to take the Kent Test often registers as a marker of one's expectation that they will pass it. This suggests that there are two groups within the cohort who take the test, those whose teachers, parents and other individuals have deemed naturally gifted enough to excel, and those that have been tutored. Since we know that tutoring occurs and has a demonstrable impact on the pass rate of the Kent test, KCC should also consider a study of the psychological impacts upon students who pass by tutoring but subsequently struggle in grammar school environments.

3.2 It is notable that the official language testifying to what the Kent Test actually assesses is vague and this is unsurprising given the distinctly political nature of the test and its history of representing a rejection of science in favour of social engineering by labour need. I have outlined this history above and am happy to expand on any further points the Commission may have. What demonstrates the unclear nature of assessment in the Kent Test is the fact that the official rhetoric of the test has moved away from the IQ-based descriptions of the 1940s/50s/60s, towards a language of "reasoning skills" and "ability in English and maths" (Kent Test Familiarisation Booklet, GL Assessment, p.3) in recent years. These two forms of assessment are fundamentally different, suggesting that the test is uncertain as to what it means to assess beyond being designed to select 30 per cent of its sitters for the benefits of a vaguely-defined model of schooling. Indeed, far more than this, the recent discourse is based in the neoliberal language of "choice". This seems to deliberately imply that success in a tutable, or even an IQ-based assessment, is a matter of preference to the individual when - according to its own logic - it is either a matter of biology or of resources. The familiarisation booklet suggests that the test "assesses whether grammar school is a suitable option for you". (p. 3). This language obscures two facts: that the Kent Test is clearly tutable, and that for significant portions of the population it is not an "option" available to them. The assessment model of the Kent Test is a fundamental paradox. Given the disincentivisation to take the test among working-class parents that KCC has already noted, the language of "choice" ("a suitable option for you") only further serves as a marker for some that the test is not for those in the lower socio-economic brackets. This language of choice and options serves as a rhetorical gatekeeper of middle-class habitus.

3.3. Overall, the Kent Test remains an historical hangover from a period in which the political will to establish Tripartite education meant that policy was forced through that was based on dubious, or pseudo-scientific principles. KCC has a number of options if it wishes to continue to provide selective grammar schools. Firstly, the Kent Test must clarify, rather than obscure, its aims behind the language of choice and options. Is it an IQ test? Is it a knowledge test? What is the purpose for doing both simultaneously? KCC

should also seriously consider the recent studies into the fluidity of intelligence, which challenges the underlying logic of grammar selection policies as they currently stand. As I have said above, “intelligence” has never been understood to be unproblematically defined as innate and fixed. This is even more the case than ever.

4 Suggestions

4.1 Administer a selection test at multiple points of entry. In addition to the test at 11, a test may also be available at 13 and 15 with grammar schools also open to non-tested admissions based on GCSE results into the A-Level programme.

4.2 Adopt a means-tested mixed admissions policy, in which the Kent Test is one means of entry for reserved pupils from families in a higher socioeconomic bracket, with significant affirmative action policies governing entry of children on Free School Meals.

4.3 Abolish the Kent Test and select for grammar-school based on teacher-based assessments or from SATS testing in primary.

5 References

- Bourdieu, Pierre. *Distinction: A Social Critique of the Judgement of Taste*. Trans Richard Nice. Cambridge, Mass.: Harvard University Press, 1984.
- Brinch, Christian N. and Taryn Ann Galloway. “Schooling in adolescence raises IQ scores”. *Proceedings of the National Academy of Sciences of America* Vol 109, No.2 (Jan 10, 2012), pp. 425 – 430.
- Hart, Robert A., Mirko Moro and Elizabeth Roberts, “Date of Birth, Family Background, and the 11 plus exam: short- and long-term consequences of the 1944 secondary education reforms in England and Wales.” Stirling Economics Discussion Paper, May 2012.
- Heim, Alice. *The Appraisal of Intelligence*. London: National Foundation for Educational Research in England and Wales, 1954.
- Gould, Stephen Jay. *The Mismeasure of Man*. New York and London: Norton, 1981.
- Zenderland, Leila. *Measuring Minds: Henry Herbert Goddard and the Origins of American Intelligence Testing*. Cambridge: CUP, 1998