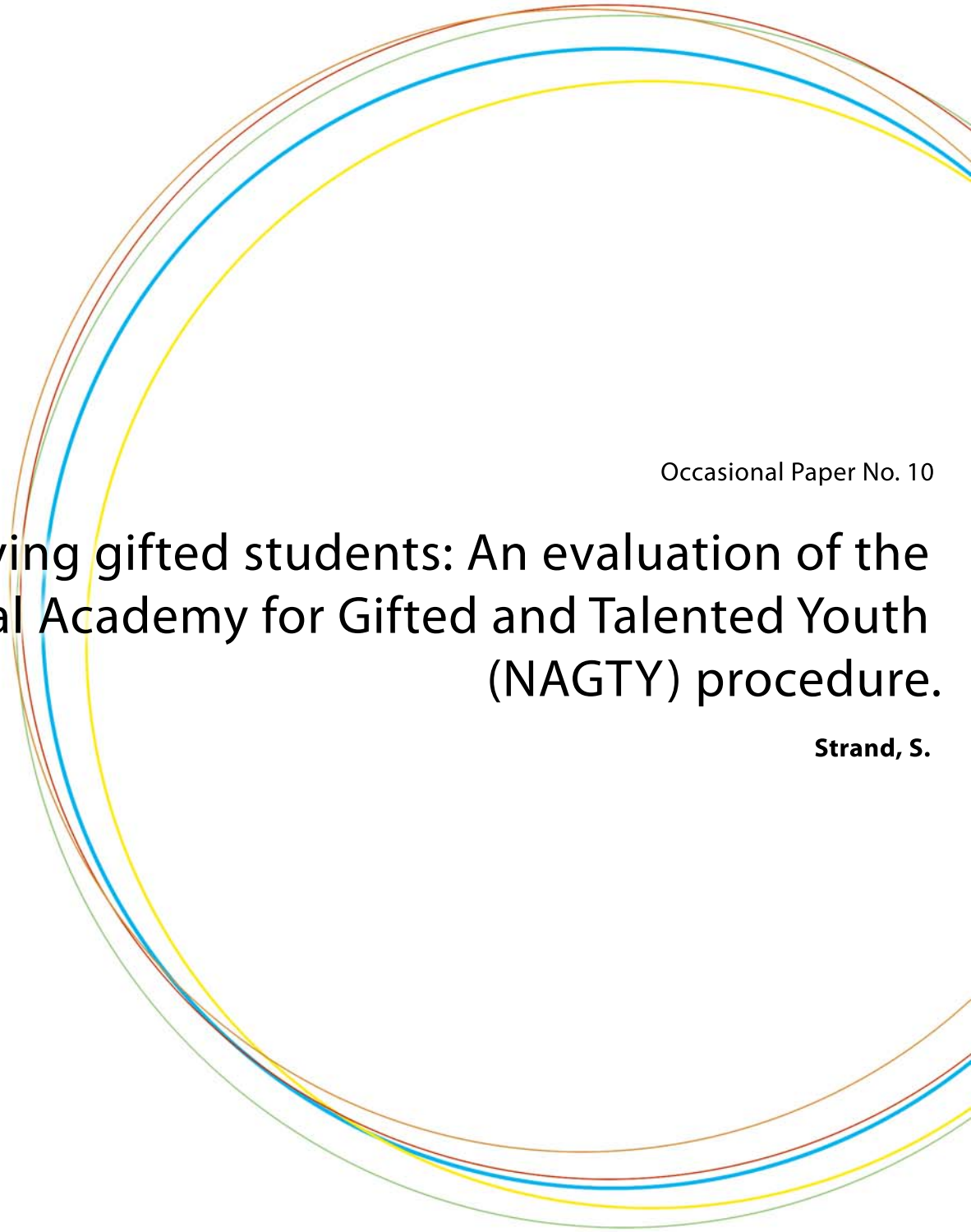


Occasional Paper No. 10

Identifying gifted students: An evaluation of the National Academy for Gifted and Talented Youth (NAGTY) procedure.

Strand, S.

gifted talented youth
THE NATIONAL ACADEMY

1-905877-10-2
978-1-905877-10-2
OCCASIONAL PAPER 10
February 2006
(online version)

Dr Steve Strand
Reader in Education
CEDAR, University of Warwick
Coventry CV4 7AL

A report commissioned by the National Academy for Gifted & Talented Youth
Nagty-research@warwick.ac.uk

CONTENTS

EXECUTIVE SUMMARY

Purpose of the report

The National Academy for Gifted and Talented Youth (NAGTY) was created by the government in February 2002 as a key partner in the Government's approach to meeting the needs of 'gifted and talented' pupils within England. The academy is based at the University of Warwick. The Academy has developed an admissions process for identifying students who are working in, or have the potential to work in, the top 5% of the ability range in one or more subjects. Student membership is open to secondary-level students aged 11-19 who are both resident and being educated in England.

The aim of this paper is to review the admission process and the eligibility criteria for identifying gifted pupils for membership of the student academy. The review is concerned only with gifted pupils, who are generally considered to be those having high levels of ability usually across a range of academic subjects, rather than those described as 'talented' in arts subjects (art & design, dance, drama, music and/or sport).

Identifying gifted pupils

Identifying the top 5% of the ability range might, at first glance, seem a relatively straightforward issue, but in reality it is difficult. The fundamental problem is the range of views on what constitutes giftedness or high ability. Hany (1993) observes "there are perhaps 100 definitions of 'giftedness' around, almost all of which refer to children's precocity, either in psychological constructs such as intelligence and creativity, but more usually in terms of high marks in school subjects". While there is no universally accepted definition or measure, it is clear that "high intelligence, as measured by IQ, is by far the most popular criterion for defining children as very able or gifted" (Freeman, 1998). The Excellence in Cities evaluation has also reported a general consensus across secondary schools that gifted pupils are those academically able across a range of subjects (Pocklington, Fetcher-Campbell & Kendall, 2002).

Since the definition of 'giftedness' is subject to debate it follows that there will be no such thing as a perfect identification system. All approaches to identification will be a balance between competing pressures. The aim of the review is to evaluate the NAGTY admissions criteria with reference to the dimensions of:

- Fitness for purpose
- Reliability and validity
- Practicality and manageability
- Equity & access
- Cost (including time and resources).

The outcomes of the review in relation to each of these dimensions are described in detail in the body of the report. A brief summary of each is given below.

Fitness for purpose

The 'fitness for purpose' of the admissions criteria has to be judged in the context of government policy. This is evolving and a new context is signalled in the recent White Paper (2005). The White Paper confirms the government's commitment to create a National Register of the top 5% of pupils. The Specialist Schools and Academic Trust (SSAT) has proposed using the national KS2 tests to identify each year the 30,000 pupils with the highest total test marks. Over five years this will build to a national register of the country's most gifted 150,000 children (Times, 05/12/05). Inherent in the concept of a national register must be that it is inclusive and will identify all gifted students, otherwise what is the rationale for the register? By the same argument, pupils who are not on the register may not be considered gifted. This 'raises the stakes' of the identification procedure and has implications for the assessment evidence employed in the admissions process.

This most obvious implication is an increased reliance on reliable and objective measures. Without a standardised approach across all schools, so that pupils are given the same opportunity to show their ability and are measured consistently against national standards, equity is impossible. This raises questions about the possible need for a NAGTY admissions test, and clarification of the role of teacher references. It also requires a clearer specification of the assessment domain: should giftedness be considered multi-dimensional or general 'all-round' ability? The manageability and cost of various options, and the equity implications, also need to be addressed.

Reliability

The reliability of an assessment is the extent to which it can be said to be accurate and not influenced by, for example, the particular occasion on which the assessment is completed or the person who is administering the assessment. Evidence of reliability is a key component in

the development of tests, and most of the tests listed have published evidence of their reliability. However there are questions over the reliability of teacher references in the identification process. Teacher-completed structured checklists do not appear to significantly improve reliability. However pupils identified through teacher reference constitute only a minority of the NAGTY membership (less than 5%). The academy should monitor the situation so that no more than 5% of members are admitted through this route. Additionally the academy should monitor closely the subsequent engagement and performance of students admitted through teacher reference compared to those admitted through the test based routes.

Validity

The validity of an assessment refers to what is assessed and how well this corresponds with the construct that it is intended to measure. The problem in assessing the validity of the identification process is the lack of agreement about the construct of giftedness. The consensus across the literature is that secure identification should be based on triangulation across multiple criteria (Montgomery, 1996; Freeman, 1998). A weakness of the current process is that, while multiple criteria for eligibility are documented, the applicant only has to select one of the criteria. Candidates can also be selective about what evidence they choose to disclose. There is no simple resolution to this dilemma. The initial Loc8or procedure, that did require both evidence from a formal test result and a teacher reference, was widely viewed by schools as complex and time consuming and was also resource intensive for NAGTY. Consequently test evidence provides the most practical basis for considering students as potentially among the gifted cohort.

The validity of the eligibility criteria are questioned since they are only ever applied *at one point in time*. If a student successfully meets the eligibility criteria then s/he gains membership of the academy in perpetuity. There is no system for reappraising performance or determining whether a student remains eligible after any period of time. Assessing whether pupils are highly able at one particular point in time is not necessarily a secure way of deciding what their potential is, or what educational experiences would best suit them. Criteria based on ability tests are probably least vulnerable, since empirical evidence shows that ability test scores tend to be more stable over time than scores from attainment tests, school grades or other measures of performance.

To periodically reassess members would require a more detailed and comprehensive assessment that is beyond the scope of the current procedures. It is therefore important to explicitly recognise that the procedure i*dentifies a student as eligible to join the academy using*

a particular piece of evidence at a particular point in time, it does not define a pupil as gifted. The basis for eligibility is that the student has demonstrated some degree of exceptional performance that suggests they *may* benefit from the distinctive range of provision offered by the academy.

Giftedness as multi-dimensional or general 'all-round' ability

NAGTY has argued that 'giftedness' is multi-dimensional and that using a single criterion to define eligibility would identify too narrow a range of skills. However there are problems with identifying pupils in each and all of the National Curriculum subjects. The QCA website contains checklists for identifying the G&T student for every subject of the National Curriculum (English, maths, science, design & technology, ICT, history, geography, modern foreign languages, religious education)[1]. This is appropriate to help teachers in different subjects consider differentiation and the needs of gifted pupils in their classes and departments. However it is inconsistent with the reported consensus that gifted pupils are those academically able across a range of subjects (Pocklington, Fetcher-Campbell & Kendall, 2002). It is also incompatible with a procedure to identify the national top 5% of ability. The more routes that there are, the greater the number of students who will qualify for entry. If our definition of gifted pupils is this inclusive, identifying the top 5% in each of the separate subjects, this makes problematic the conception of a simple "top 5%".

A rationale for this plurality is often rooted in the ideas of Gardner (1983, 1993) and his conception of multiple intelligences. However the measurement basis of Gardner's theory is weak (Klein, 1997, 2003). It remains the case that a 'g' based factor hierarchy, with a general ability factor at the apex and various specialised abilities arrayed below it, is the most widely accepted current view of the structure of intellectual abilities, and general cognitive ability remains the best single predictor of academic success and job performance (Neisser et al, 1996). However, while we may reject the notion of 10 intelligences, many tests of reasoning abilities do make broad distinctions below the level of 'g'. For example the most commonly used test for nomination for NAGTY (the Cognitive Abilities Test or CAT) recognises that within general cognitive ability there are strengths and weaknesses in reasoning with words, with numbers or with shape and space, but still remains grounded in the assessment of students' reasoning abilities rather than their attainment in individual school subjects. This addresses equity issues which show that pupils from some social and ethnic groups or speakers of English as an Additional Language (EAL) often demonstrate particularly low scores on verbal reasoning

1. *Art & design, music and physical education are also included but are not listed here since they cover the area of talent.*

tests depressing their overall mean score, even though they demonstrate high ability on quantitative reasoning or non-verbal reasoning tests. On balance using these three broad abilities is recommended, particularly because of the equity issues.

The use of the national end of KS2 and KS3 tests

Jesson (2005) has recently proposed identifying the top 5% of pupils in the KS2 tests on the basis of total test marks. KS2 tests are not included among the tests specified in the current eligibility criteria and this is a significant omission. KS2 tests should be included, using total test marks to identify pupils with exceptional performance placing them in the top 5% on the basis of total test score[2]. In relation to KS3 tests, the current eligibility criteria include a level 8 assessment in any one of the national curriculum subjects. This is problematic for the reasons of multi-dimensionality identified above. Therefore a similar method to that used for KS2, identifying the top 5% of the cohort based on total test marks, is proposed in relation to the KS3 tests.

The above criteria will identify pupils who have exceptionally high-level performance, however they may be relatively weak in identifying pupils with potential. School performance is much more susceptible to factors external to the child such as the extent of curriculum covered in class or the quality of teaching received. Performance in an attainment test, particularly a poor performance, may therefore have little to do with a pupil's ability. Reasoning tests offer an opportunity to minimise the effect of specific curricular experience and to identifying general, transferable learning abilities applicable across a wide range of tasks and situations. Therefore it is important that where attainment tests are used, reasoning tests are also used as well. Identification of gifted students would not be served well if identification was solely based on national end of key stage tests.

It is important to stress that the use of reasoning tests should in no way be seen to imply assessment of innate ability or fixed capacity. This was summarised succinctly by Mackintosh & Mascie-Taylor (1985): *"If a child has been deprived of intellectual stimulation or educational opportunity, it is small wonder that his intellectual performance will reflect this fact. An IQ test is no more able to gauge a child's true innate potential regardless of the circumstances of his upbringing than is a pair of scales to measure his true potential weight regardless of what he has been fed" (Mackintosh & Mascie-Taylor, 1985, Swann Report, Annex D).*

---

2. *See the main text for a discussion of some important technical issues to be resolved before such a procedure could be implemented.*

As with any test we should be careful of the apparent precision of KS2 total test marks (ranging from 0 to 280 marks) since any line drawn in a continuous distribution is subject to threshold effects. It would be nonsensical to consider a pupil with a mark above a certain threshold 'gifted', but a pupil with one mark less as 'not gifted'. This serves to further reiterate the point made in the section on validity that the eligibility criteria do not define giftedness: they indicate the student may have the potential to benefit from some of the distinctive provision offered by the academy, but *are no more than indicators.* It is important that this is highlighted to avoid any public perception that NAGTY membership defines or assures giftedness.

Equity issues in identification

Concerns are often expressed that test results may be biased or unfair to certain groups. However it is important to remember that objective testing has historically been seen as an instrument of equity. A standardised test offers a degree of distance and impartiality, though if equality of educational opportunity has not preceded the test then the fairness of the test may be called into question. What is less often considered is the evidence that teacher assessment can be systematically biased in relation to student characteristics, such as gender, ethnicity, social class or behaviour, and that test scores may be less likely to suffer distortion from such extraneous pupil characteristics. The research concerning special educational needs (SEN) and giftedness is complex since there is debate about the definition of both high ability and of certain types of SEN. However by including both attainment and ability tests within the criteria, the NAGTY framework allows for some flexibility. The issues are complex and beyond the scope of this present paper (see Winstanley, 2005).

A recent paper by Campbell et al. (2005) provides a detailed geo-demographic analysis of the students admitted to NAGTY in 2002/03 and 2003/04. The analysis shows that, whilst having a student membership skewed towards groups with high levels of cultural and economic capital, the academy has nonetheless reached significant numbers of students in the poorest areas, around 9% of the student members. This is comparable with the proportions of such students achieving the highest 'A' level grades and the proportion entering university. It is important that this monitoring and evaluation continues. Of particular importance is to monitor the subsequent achievement of students identified through different types of evidence. For example, those identified through the various different test routes and those identified through teacher reference. Specific recommendations for further research are given below.

Cost and resources implications

The academy has in the past considered the option of creating its own entrance test for academy membership. However a single national test would be costly to develop. It is arguable whether such costs could be justified, since funds would necessarily have to be diverted from delivering actual provision and activities to members. A framework approach, whereby variously widely used tests are endorsed for the purpose of identification is recommended. It is important to note however that the framework approach might be undermined if the tests became widely perceived as 'high stakes' tests, for example if significant social or educational advantage were perceived to accrue from academy membership *per se*. The role of NAGTY membership within the University Admissions process must be clarified. If NAGTY membership is a box to be ticked on the UCAS form, and this is perceived to improve the chances of admission to university or to an elite institution, this will undermine the validity of the assessment process.

Recommendations for amendments to the identification criteria

1. The academy should continue with a framework approach towards identification. An effective framework will utilize the wide range of assessment data currently available while at the same setting clear quality standards (see 3 below). The costs and resource implications do not currently warrant the development of a separate NAGTY admissions test.

2. Any framework must be regularly revised and refined. Some amendments to the current framework are recommended:

    - Increase the cut-score on each of the three CAT batteries to 129 or above, and drop the criterion of 120 or above on any two batteries, so that the overall identification is limited to 5%;

    - Discuss further with CEM the appropriateness of continuing with a Yellis mean score of 126, but to allow cut-scores of 129 or above on the separate vocabulary, maths and non-verbal sub-tests of MidYIS, subject to restricting overall identification to the top 5%;

    - Include the KS2 tests which are not currently part of the eligibility criteria. This should be based on total test marks to identify the top 5% in terms of overall performance;

- Replace the current KS3 subject based criteria with a measure of total test marks to again identify the top 5%;

- Remove ‘UK Maths Challenge’ and ‘World Class Tests in Mathematics’ from the list of eligible tests (but retain World Class Tests in Problem Solving).

3. To be included in the framework, test publishers must provide documentary evidence of:

   - reliability coefficients of at least 0.90 or above for any test or sub-test used as an eligibility criteria;

   - the nature of the sample employed in the standardisation process, including evidence that it is nationally representative;

   - (a) a national standardisation of the test in the last 10 years, (b) equating of the test to more recent norms, or (c) clear and defensible data as to why certain level of performance on the test should be equated to top 5% or ‘gifted‘ range.

4. Pupils identified through teacher reference constitute only a minority of the NAGTY membership (less than 5%) but this is in many ways the least secure of the criteria. The academy should monitor the situation so that no more than 5% of members are admitted through this route. Additionally the academy should monitor closely the subsequent engagement and performance of students admitted through teacher reference compared to those admitted through the test based routes.

Recommendations for further research

5. There is a need to determine the equity impact of the different criteria in terms of the social composition of the populations they identify. Where the tests used cover a large proportion of the national population, these questions can be addressed by linking the national results to the Pupil Level Annual School Census (PLASC) data. For example, Jesson (2005) reports that only 3% of the 30,000 pupils in state primary schools identified in his analysis of KS2 test scores were entitled to FSM, against an average for secondary schools of 14%. We need further research on the extent to which differences associated with ethnicity, sex, entitlement to FSM, EAL or age exist for the cohort identified through KS2 test scores.

6. The academy also need to determine whether any group differences as identified above are similar across other criteria such as reasoning ability tests and teacher reference. NAGTY should commission research to investigate these issues. The availability of national PLASC makes this practical in ways which were previously impossible and represents a unique opportunity to better understand these complex issues.

7. It is also important for the academy to monitor the subsequent attainment and achievement of students identified through different types of evidence. NAGTY should collect and collate evidence of the type of activities that different students undertake with the academy and relate this to the route through which they were identified.

8. NAGTY should seek to link its database with the national pupil test and assessment databases so that the progress of pupils who are NAGTY members can be tracked. This is important to establishing the extent to which students initially identified as members continue to be identified among the top 5% over time, or the extent to which performance may be variable over time.

9, Finally, the above link to the national test and assessment databases would give the academy the important opportunity to evaluate whether NAGTY members make any more, or less, progress than similar students who do not become members of the academy. This would allow the academy to establish the 'value-added' of NAGTY membership. It would allow the academy to prove its overall worth, and whether there were any differential effects, for example whether its impact was particularly great for pupils from disadvantaged areas or from ethnic minority groups.

Main Report

Introduction

Identifying the top 5% of the ability range might, at first glance, seem a relatively straightforward issue, but in reality it is difficult. The House of Commons inquiry into Highly Able Children (1999) stated:
"The identification of the target group proved the most complex aspect of our inquiry… it would be fair to say that a definitive answer to the question "Who are the highly able?" has not emerged" (par 7).

The fundamental problem is the range of views on what constitutes giftedness or high ability. Hany (1993) observes *"there are perhaps 100 definitions of 'giftedness' around, almost all of which refer to children's precocity, either in psychological constructs such as intelligence and creativity, but more usually in terms of high marks in school subjects".* While there is no universally accepted definition or measure, it is clear that "high intelligence, as measured by IQ, is by far the most popular criterion for defining children as very able or gifted" (Freeman, 1998). The Excellence in Cities evaluation has also reported a general consensus across secondary schools that gifted pupils are those academically able across a range of subjects (Pocklington, Fetcher-Campbell & Kendall, 2002).

Despite the lack of clarity on the domain to be assessed, NAGTY was charged with developing criteria to identify the top 5% of the ability range. We described below the origin and development of the current NAGTY procedures. There is no documented *rationale* for the NAGTY approach to identification. A key purpose of this review is to evaluate the appropriateness of the current criteria. In doing so it is impossible to sidestep the issue of definition of what we mean by the "top 5% of ability". Frequently this seems to have been interpreted as a debate about whether it is the top 5%, or the top 2% or top 10%, that should be identified. However these issues are tangential. The fundamental issue is 5% (or 2% or 10%) of what? What is it that we mean by the 'gifted' or 'highly able' and how are we going to quantify it in order to identify the top 5%? This fundamental question needs to be addressed, along with the purpose or rationale for the identification process.

Nevertheless the review is bounded by certain constraints, specifically that:

- NAGTY should be identifying the top 5% of the 11-19 population. What 'the top 5%' means conceptually and operationally is up for discussion, but the overall requirement to provide

the most reliable and defensible basis for differentiating 5% from the remaining 95% is not negotiable. For example the review could recommend to NAGTY that teacher reference is the preferred method for identifying gifted pupils, but we would still have to provide a reliable and valid method for quantifying teacher reference and expressing it on a scale that would allow for selection of the top 5%;

- The review is specifically about the rationale and methods of identification of the top 5%, not the rationale for creating NAGTY. However insofar as the overall purpose and functions of NAGTY will determine the fitness for purpose of any identification procedure, then some debate about the function and role of NAGTY is necessary.

- The review is specifically about the procedures for identifying pupils who are eligible for joining the national academy in relation to giftedness. This is a subset of those who might be identified as gifted by their schools for other purposes, for example as part of the gifted & talented strand of EiC. It also specifically does not include identification in relation to talent in areas such as art, art & design, dance, drama, music or sport.

The development of the NAGTY admissions criteria for identifying gifted pupils

The Loc8or procedure

NAGTY was established in February 2002. The first task relating to identification was to identify pupils for the 2002 summer school. NAGTY formed an expert group on assessment to consider the issue which led to the evolution of the "Loc8or" identification process. The procedure reflected a desire to be comprehensive, flexible and inclusive, and was predicated on a central role for the school and teacher, as reflected in the January 2003 guidance below.

> *"The Academy is looking for academic potential as well as ability and we recognise that teachers are often the best judge of this irrespective of national tests and other formal measures of intellect. In recent years it has become clear from research evidence that there is no single measure of 'giftedness' that can be applied equally to all pupils. Loc8or requires an applicant to assemble a portfolio of information on which we can judge abilities and potential. At least one piece of evidence from each of (a) formal test results, and (b) informal evidence, is required to complete the portfolio. (NAGTY guidance, January 2003)*

This evidence therefore consisted of:

(a) A letter of application written by the student;

(b) At least one formal test result (selected from a disparate range of tests / assessments);

(c) one or more from:

- An original piece of coursework;
- Evidence of interest in a particular subject or topic;
- Recommendation from school or teacher;
- Recommendation from another education professional.

A benefit of this eclectic approach - requiring diverse sources of evidence in a portfolio - was that it side-stepped the thorny issues involved in closely defining the meaning of the term 'gifted'. The process simply collated and passed the evidence and the responsibility for the decision to Academy Assessment Consultants, who reviewed the evidence and made the formal recommendation.

Subsequent revisions

However many schools found this process to be extremely complex and time consuming. It was also resource intensive for NAGTY (c.f. Academy Assessment Consultants) and was slow in recruiting membership. In January 2004 the application process was streamlined to the current system by removing the requirement for the student to write a letter of application and requiring only one piece of evidence, either a formal test result or a teacher reference. However this still includes an extremely diverse range of evidence, including attainment tests, reasoning tests, specialist activities (e.g., UK maths challenge) and teacher reference, and apparently inconsistent thresholds between apparently similar tests (see the minutes of the Expert Group on Assessment discussing the CAT / MidYIS criteria).

Essentially this was an administrative revision rather than a fundamental reflection on the nature of "the top 5%" and what it meant to define this operationally. The current criteria are therefore not dissimilar from those described above. *Appendix 1* contains a copy of the current eligibility criteria for academy membership.

Evaluating the adequacy of the NAGTY procedures

We have started from the premise that the definition of 'giftedness' is inherently problematic and contested. It follows that there can be no such thing as a perfect identification system, and all approaches to identification will be a balance between competing pressures. The aim of the review is to evaluate the current approach with reference to the criteria of:

- Fitness for purpose
- Reliability and validity
- Practicality and manageability
- Equity & access
- Cost (including time and resources).

Fitness for purpose

Arguably, the key criteria in the evaluation of any assessment or identification procedure is it's "fitness for purpose". What is the purpose of the identification procedure and how well do the criteria meet this purpose?

To date, the NAGTY identification procedure has reflected a degree of flexibility in the type of evidence that can be considered. For example while it is typical that the evidence consists of a test result, these are drawn from a wide range of tests and assessments and it is possible for the evidence to consist of a teacher reference. This academy has not wanted to impose a monolithic test or single criterion, preferring instead a situation in which schools reflect on the nature of giftedness and have some flexibility in developing their own criteria. This has been a strength during the development of this new national initiative. The search for a precise definition and a focus on measurement might have distracted schools from finding ways to identify and meet the needs of their most able children. It has been argued that "it is only in schools where identification is embedded in school-based systems that it has significant impact on classroom practice, otherwise the link between identification and provision is tenuous" (Eyre, 1997). This has also fitted well with the optional nature of pupils' involvement with the National Academy. The eligibility criteria are clearly laid out and pupils who meet these criteria can choose to 'opt-in' or otherwise to membership.

This success is indicated by the fact that there has been has been relatively little overt opposition from schools to the process of identifying gifted students. There is a degree of discomfort associated with the discourse of gifted and talented in England. Ideology plays a

part, in that making special provision for gifted and talented pupils is commonly construed as elitist (Bourdieu, 1998; Ball, Bowe & Gewirtz, 1996) by reinforcing the advantages that the professional classes might gain for their children in competition for admission to high status universities. The absence of significant opposition from schools should not be underestimated. A degree of ambiguity in the definition of 'giftedness' has therefore been useful in easing the way.

A new policy context

However a fundamental change to this principle has arisen in the context of debate around the White Paper (2005). The Specialist Schools and Academies Trust (SSAT) has proposed using the national end of Key Stage 2 (KS2) tests completed by all 11 year olds in England each year. The total marks across all tests would be used to identify the top 5% of pupils in terms of total test marks[3]. These pupils and their secondary schools would then be approached directly by the academy.

The people at Warwick have agreed that, instead of relying on teacher recommendations, they will get 30,000 names of 11 year olds each year. This would build over five years to a national register of the country's most gifted 150,000 children aged 11-16". (Times, 5th Dec 2005).

The argument has been galvanised by the assertion that "40% of secondary schools have never recommended any child to attend NAGTY" (Times, 5th December 2005). The proposal seems to imply that identifying the cohort is easily achieved, whereas in reality it is very difficult. Some problematic issues are outlined later. However the principle appears to have been accepted within the White Paper:

"We will develop a national register of G&T pupils. This will allow us to invite all who fall within the top 5% to join NAGTY… we will also use the register to help provide the right local opportunities to extend G&T pupils studies" (White paper, par 4.24).

Arguably the academy has contributed to this demand through altering its eligibility criteria in January 2004 in order to increase membership. Thus membership has moved beyond the initial DfES target of 7,000 members by 1st September 2004 to the current nominated membership of over 70,000. It has been perhaps a predictable result of this that central government has asked for the 'full cohort' to be identified.

---

3. There is some ambiguity at present as to whether this refers to all three tests including science, or just to the English and mathematics tests.

Some would argue that the concept of a 'National Register' is fundamentally flawed - the organisation cannot cater for the needs of 200,000+ individual pupils from the national centre. Even laying on summer schools for 1,000 pupils reaches less than 0.5% of the target population. In practice the needs of this group can only be met through a distributed model where the academy works through schools to improve the provision for their gifted students and to focus on activities to raise the representation of disadvantaged groups. This is the model underlying the gifted & talented strand of Excellence in Cities (EiC). OFSTED (2005) have reported that provision has been promoted and developed well in a large majority of the schools visited, with weaker provision only in schools where the G&T strand is not embedded in the mainstream curriculum and is perceived by staff as an 'add on'. More generally, provision for gifted and talented pupils is now rated as good or better in two-thirds of the schools inspected in 2004/05.

However, there is now a clear and explicit direction for NAGTY to create a National Register of the top 5% of students as a key component of national policy. The identification procedures needs to be reviewed in this light.

Implications of a national register for the identification process

A loose analogy can be drawn currently between NAGTY and MENSA as membership organisations. It is accepted that some people in MENSA may not be highly intelligent, depending on how you define the construct of intelligence. More importantly it is also the case that you can be highly intelligent but not be in MENSA, because you may not have chosen to apply. However inherent in the concept of a national register must be that it is inclusive and will identify all gifted students, otherwise what is the rationale for the register? By this argument, pupils who are not on the register are not considered gifted. This 'raises the stakes' of the identification procedure and has implications for the assessment evidence employed in the admissions process.

There will be an increased need for highly reliable and objective measures. A standardised approach needs to be taken across all schools, so that pupil are given the same opportunity to show their ability and are measured consistently against national standards. Without this there is no meaningful basis for comparison. This raises questions about the possible need for a NAGTY admissions test, and clarification of the role of teacher reference. It also requires a clear specification of the assessment domain - should giftedness be considered multi-

dimensional or general 'all-round' ability? The manageability & cost of various options, and the equity implications, also need to be addressed. We assess the implications in detail below.

Reliability

Reliability is a key concept in educational assessment. The reliability of an assessment is the extent to which it can be said to be accurate and not influenced by, for example, the particular occasion on which the assessment is completed or the person who is administering the assessment. How reliable are the different assessment criteria for eligibility for entry to the academy? In the case of tests this can be assessed by giving the same test twice, using parallel forms of the test, or splitting the test randomly into two halves and comparing the scores. In the cases of tasks, observations or checklists the equivalent procedures are to compare the rating of the same event or construct by two independent raters.

Teacher reference

Evidence of reliability is a key component in the development of tests, and most of the test listed have published evidence of their reliability.[4] Of greater concern is the reliability of teacher reference in the identification process. Teacher reference undoubtedly gives rise to rich data on students' abilities. However there are concerns about the objectivity and reliability of the evidence.

There can be wide variation between teacher judgements and 'objective' measures such as test scores. Bennett et al (1984, p215) found that 40% of potentially high achievers had been underestimated by their teachers. Tempest (1984), using IQ as his criterion of giftedness, reported that out of 72 six-year olds identified by their teachers as gifted, only 24 had IQ scores of 127 or above and seven had IQ scores of under 110. Additionally two children with reading ages six years in advanced of their chronological age were not nominated as gifted. He concluded that teacher recommendation alone was not reliable. Finally, Nebesnuick (1993) showed a significant discrepancy between the assessment of able pupils by their primary teachers and by their subsequent Y7 teachers when they transferred to secondary school. The primary teachers often chose pupils as able by their way of working, rather than their cognitive ability. If different teachers identify different pupils from among the same group of pupils, then the implications for reliability are obvious.

---

4. The single exception is actually the national KS2 and KS3 tests, where traditional reliability indicators are not published.

Teacher checklists

Checklist approaches have attempted to improve the reliability of teacher identification by more closely specifying the behaviours expected of gifted students. However checklists of the supposed characteristics of highly able children vary considerably and some items can be confusing. The following is taken from Freeman (1998):

A child asking lots of questions can be seen as gifted or as attention-seeking, or perhaps lives in a home where questioning is encouraged rather than one where children are encouraged to work things out for themselves…One list may ask the teacher to look for dedicated seriousness, while another suggests a keen sense of humour… While one list may point to a tendency to perfectionism (and thus procrastination), another will describe the highly intelligent as speedy decision makers. Some see introversion as typical feature of the high IQ child, although there is no reliable evidence of personality features being associated with IQ (Freeman, 1998)

There are also issues of technical quality raised with regard to some of the checklists such as the Renzuli scales. Inter-rater reliability is low and measures of predictive validity are extremely limited. Fundamentally, teachers judgments are very much made in relation to local norms and the teacher's previous experience. It can be difficult to judge performance against national standards, divorced from the 'typical' range of ability that a teacher encounters in the context of his/her particular school. The lack of national norms for the Renzuli and other checklists is therefore a particular problem.

Implications

There are concerns that middle-class families may be disproportionately represented if the criteria allows a range of evidence including teacher and parental reference rather than objective test data. Middle-class parents tend to be assertive and may make a convincing case for their children when more objective test data might have excluded them. This would run counter to the desired outcome of increasing the numbers of gifted & talented pupils from disadvantaged areas. For further discussion refer to Power et al (2003).

However the number of pupils being forwarded by the non-test route is actually relatively small. An analysis of the current 36,845 enrolled members of the academy indicates that, of the 25,000 where the basis of the nomination is recorded, only 1,164 or just under 5%, have been nominated through non-test routes. The analysis further shows that this route is used equally by

state and independent schools. Our recommendation at this stage is to continue to allow this route because of the richness of the data on students that can emerge (see following section on validity). However, the academy should monitor the situation so that no more than 5% of members are admitted through this route. Additionally, the academy should monitor closely the subsequent engagement and performance of students admitted through teacher reference, compared to those admitted through the test based routes.

Validity

The validity of an assessment refers to what is assessed and how well this corresponds with the behaviour or construct that it is intended to assess. There are many forms of validity, but perhaps construct validity is key here, or the extent to which the assessment measures what it purports to measure. This requires a clear and detailed definition of the construct being measured. For example a full definition of reading as a construct would include not only reading aloud, but also reading comprehension, accuracy and enjoyment of reading. For further discussion please see Stobart & Gipps (1997).

The problem in assessing the validity of the identification process is the range of definitions and the lack of agreement about the construct of giftedness. By foregrounding issues of reliability there is inevitably a trade-off with the richness of the information arising from other sources. Relying solely on test evidence will increase reliability by using methods that have less error, but necessarily lose some of the richness of the concept inherent in other approaches.

The most valid method of identification would allow for triangulation between these different assessments. The consensus across the literature is that secure identification should be based on multiple criteria (Montgomery, 1996; Freeman, 1998). A weakness of the current process is that is not multi-methods based. While multiple criteria for eligibility are documented, the applicant only has to select one of the criteria. This is particularly problematic where candidates can be selective about what evidence they choose to disclose. Also, inevitably, certain criteria will be easier to meet than others.

There is no simple resolution to this dilemma between reliability and validity. However we have also to consider issues of manageability and cost. The initial Loc8or procedure did indeed require both evidence from a formal test result and teacher reference. However this was widely viewed by schools as complex and time consuming and it was also resource intensive for NAGTY. Consequently test evidence is the most practical basis for considering students as potentially among the gifted cohort. However it does not provide a definitive judgement.

One-off assessment

A fundamental issue that has not been addressed is that the eligibility criteria are only ever applied *at one point in time*. If a student successfully meets the eligibility criteria then s/he gains membership of the academy in perpetuity, or at least until the age of 19. There is no system for reappraising performance or determining whether a student remains eligible after any period of time. There is no analogy with regard to SEN at the other end of the spectrum, since pupils with statements of SEN have an annual review of the statement. It is easy to think, for example, of situations where a student may gain exceptional high scores in the KS2 tests at age 11, but, for a wide range of reasons, may not maintain that trajectory. Assessing whether pupils are more able at one particular point in time is not necessarily a secure way of deciding what their potential is, or what educational experiences would best suit them.

Criteria based on ability tests are probably least vulnerable here, since empirical evidence shows that ability test scores tend to be more stable over time than scores from attainment tests, school grades or other measures of performance. However even for ability tests there is evidence of quite extreme changes in scores for a minority of students (e.g., Strand, 2004). Periodic reassessment, say after a period of three years could be considered. Such a procedure might be an effective lever to raise levels of student involvement in the academy, if allied to a requirement for active participation in the activities offered by the academy. However it would involve further financial costs and would reduce the flexibility of the system. There might also be negative consequence for students' self-esteem of being effectively ejected from the academy. However the weakness of the assumption implicit in the current identification process (that 'giftedness' is a stable trait over time), needs to be recognised explicitly.

Conclusion

It must be publicly recognised and made explicit that the procedure *identifies a student as eligible to join the academy using a particular piece of evidence at a particular time.* It does not define a pupil as gifted. The basis for eligibility is that the student has demonstrated some degree of exceptional performance that suggests they *may* benefit from the distinctive range of provision offered by the academy. To confirm a pupil as gifted would require a more detailed and comprehensive assessment, triangulating across a range of assessment tasks in a range of contexts; utilising a range of modes within the assessment, and a range of response formats and styles. This is beyond the scope of the current procedures. We therefore recommend

retaining the current teacher reference route because of the richness of the information it provides, subject to the conditions previously outlined.

Is giftedness multi-dimensional or general 'all round' ability?

In the past, NAGTY has argued that 'giftedness' is multi-dimensional. Using a single criterion to define eligibility would identify too narrow a range of skills. There are two problems with this position. First, while there is no universally accepted definition or measure, it is clear that "high intelligence, as measured by IQ, is by far the most popular criterion for defining children as very able or gifted" (Freeman, 1998). There is also a general consensus across secondary schools that gifted pupils are those "academically able across a range of subjects" (Pocklington, Fetcher-Campbell & Kendall, 2002). There is also a practical issue. Suppose we have multiple criteria with each criteria designed to identify approximately the top 5% of the population *on that measure*. Inevitably, because there are several different ways to qualify, more than 5% of the population will be identified. This makes problematic the conception of a simple "top 5%".

Multiple subject based criteria

This issue is most apparent if we look at the QCA and DfES websites that support the gifted and talented strand of the Excellence in Cities (EiC) initiative. The QCA website contains checklists for identifying the gifted and talented student in science, the gifted and talented student in design and technology, the gifted and talented student in modern foreign languages and so on for every one of the subjects of the National Curriculum (English, maths, science, design & technology, ICT, history, geography, modern foreign languages, religious education)[5]. Inspection of the lists of criteria shows some communalities, but also many divergences (see Appendix 2 for copies of selected checklists).

There are even divisions within many of these subjects. Thus the guidance for identifying gifted pupils in Design & Technology states: *"The pupils who are gifted in design and technology may be a very different group from those with gifts and talents in other subjects... Some of them will have abilities in a specific area - for example working with food, using computer-assisted design (CAD) or high quality making - but not in others".*

5. *Art & design, music and physical education are also included but are not listed here since they cover the area of talent.*

Similarly in ICT it is argued *"it is important to remember that they may not be gifted in all aspects of the subject. For example, some pupils may be able to use high-level programming skills to solve problems, but may not be as good at constructing and investigating databases".*

We are not critical of these subject based approaches to the identification of gifted pupils for the purpose of teaching and learning. This is entirely appropriate to help teachers in different subjects consider the learning needs of gifted pupils in their classes and departments and the implication for differentiation, extension and enrichment. Indeed *"in schools where the G&T strand is well established, a positive trend has been the tendency of departments to formulate their own, subject specific definition of ability" (OFSTED, 2001)*. However it is hard to escape the conclusion that what are being described are a range of specific skills and interests. For the separate and distinct purpose of identifying pupils eligible for membership of NAGTY, then criteria that identify the top 5% in any separate subject will result in a large and heterogeneous gifted and talented group. This is not inherently problematic, indeed it may well be viewed as a strength by some, but it makes problematic the conception of a simple "top 5%".

Thus one secondary headteacher quoted in the House of Commons select committee report (1990) states: *If children are able in every subject area then, in my own school, we are talking about 2 per cent. If you take the multiple intelligences (which is the approach we take)… and then identify the children, then you are talking of a group in excess of 30 per cent of the school population - sometimes as high as 40" (House of Commons, 1999, par 11).*

Multiple intelligences

A rationale for this plurality is often rooted in the ideas of Gardner (1983, 1993) and his conception of multiple intelligences. Gardner's thinking has been useful in encouraging educationalists to look beyond traditional academic boundaries when considering factors that impact on learning, for example in considering the importance of intrapersonal influences (e.g., motivation, self-esteem, sense of values). It has also been useful in highlighting and valuing a broader range of student outcomes, including the interpersonal, kinaesthetic and naturalistic spheres. However as a theory, 'multiple intelligences' has many weaknesses. For example it is not clear to what extent the 'intelligences' are supposed to operate separately or interconnectedly. The fact that the existence or otherwise of an intelligence is not testable experimentally, and cannot be accurately psychometrically assessed, is also critiqued (Klein, 1997). Klein (2003) makes the reasoning clear:

"Consider a Y6 pupil who selects and applies mathematical operations to solve fractional problems. Different assessors might construe this as evidence of logical-mathematical intelligence, general intelligence or procedural and declarative knowledge about fractions. A teacher or researcher who chooses to ascribe the students performance to intelligence (general or multiple) is construing this performance as part of an underlying ability that is relatively stable over time and moderately general in scope. Generality and stability are both ascriptions that need to be verified. However the forms of assessment that Gardner and other MI theorists recommend do not include checks for such generality and stability".

Reviewing Gardner's research on assessment, carried out as part of Project Spectrum, Klein (2003) concludes:

"most pairs of activities that were supposed to assess the same intelligence showed different results… Moreover, some pairs of activities that were supposed to assess different intelligences yielded similar results… Furthermore, these were 'one-shot' assessments, and Gardner has not published any research showing that the results of such assessments are stable over time. More generally, none of the Multiple Intelligence (MI) surveys currently in circulation has any published evidence for its validity" (Klein, 2003, p56).

In short there is no substantive body of work relating to measurement of the 'intelligences'. Gardner frequently simply extrapolates from individuals' specific performance to their 'intelligences' with little justification. For many authors, Gardner's intelligences are better described as talents, personality dimensions or cognitive styles (e.g., Morgan, 1996; Klein, 1997).

Verbal, quantitative and non-verbal abilities

While we may reject the notion of 10 intelligences, many tests of reasoning ability do make distinctions within broad domains. For example the Cognitive Abilities Test (CAT) is composed of three sets of tests. While the tests aim to assess a common construct (reasoning ability) they do so through distinct stimulus modalities, specifically thinking with words (Verbal reasoning), thinking with numbers (Quantitative reasoning) and thinking with shape & space (Non Verbal reasoning). A similar division across these three broad areas can be seen in the MidYIS tests (vocabulary, maths and non-verbal).

While performance on the three CAT batteries is positively correlated, some individuals show a marked strength or weakness in reasoning in one or other of the three modalities. For example,

some students have an exceptional facility in thinking with words, but are relatively weak in the other areas. Equally other students may have a great facility with thinking in shape and space, or with number, but be relatively weaker in the verbal area. It is tempting therefore to consider a highly able student as one who scores in the top 5% (SAS of 126 or above) on any one of the three tests. However while each separate battery identifies 5% of the population, the proportion who score 126 or above on *any one* of the three batteries is around 10% of the population. Similar evidence has been provided by the CEM centre in an analysis of their MidYIS tests, where use of the three separate vocabulary, maths and non-verbal reasoning tests would raise the proportion to 9% (Lyth & Coe, 2005).

General ability

It remains the case that a 'g' based factor hierarchy, with a general ability factor at the apex and various specialised abilities arrayed below it, is the most widely accepted current view of the structure of intellectual abilities (Neisser et al., 1996). Thus we have the CAT3 mean score, the MidYIS mean score or other indicator of general cognitive ability. General cognitive ability remains the best single predictor of academic success and job performance (Neisser et al, 1996). This is not to say it is the only predictor, for example successful school learning depends on many personal characteristics other than ability such as persistence, interest in school and willingness to study, as well as broader factors such as a supportive home environment and high quality teaching.

Options

We can consider two options. One is to select on the basis of global or all-round ability, using a single mean standardised score of 126 or above, on the basis that:

1. The overall proportion will be restricted to around 5%;
2. On all tests, the best indicator of a pupil's future performance is their overall test score;
3. Most research shows that 'islands of exceptional ability' are usually just that, exceptional;
4. It is conceptually the simplest criteria to understand.

Alternatively, we may want to continue to allow the use of any of the three separate reasoning abilities but raise the cut-score from 126 to 129. This would:

1. reduce the proportion identified on each battery to around 3%, but restrict the overall proportion identified to around 5%;
2. recognise that within general cognitive ability there can be strengths in reasoning with words, with numbers or with shape and space, but still remain grounded in the assessment of students' reasoning abilities rather than their attainment in school subjects;
3. address equity issues which show that pupils from some social and ethnic groups or speakers of English as an Additional Language (EAL) often demonstrate particularly low scores on verbal reasoning tests depressing their overall mean score, even though they demonstrate very high ability on quantitative non-verbal reasoning tests (see equity section for further discussion).
4. In practice it is the simplest criteria in that the cut-score for the majority of tests will be 129 or above (see Appendix 3) and therefore easier to apply.

On balance the second option above is favoured, particularly because of the equity issues. Further discussion with the test publishers, particularly CEM, will be necessary to determine the most reliable criteria for identifying the top 5% on their tests. Tests differ in the number of items they contain, their reliability and whether performance is more or less speeded (time allowed), and it is inappropriate to simply apply the same cut-score to all tests regardless of these factors.

The use of tests of attainment such as the national KS2 and KS3 tests

National end of Key Stage 2 and Key stage 3 tests have historically formed one of the form of test evidence for admission to the academy. However the criteria have usually been based on individual subjects. For example in relation to KS3 tests the criteria include a level 8 in the KS3 mathematics test[6] or a level 8 teacher assessment in any other subject. These criteria fall foul of the issue about multi-dimensionality we have argued above. If a level 8 is accepted in English, maths, science, history, geography, ICT and modern foreign languages then the total proportion is likely to rise above 5%. However an overall measure can be derived from summing performance across all the component tests which then allows identification of the top 5% on the basis of general attainment. Indeed a recent analysis of the KS2 tests by the SSAT has done just this (Jesson, 2005). This was discussed earlier in the section on the new policy context.

6. Level 7 is the highest level that can be awarded in the KS3 English and Science tests.

We should recognise however that the use of the national tests is not unproblematic. It will identify those who have exceptionally high-level attainment, however it may be relatively weak in identifying those with potential for high attainment. A fundamental concern is that the national tests are measures of attainment, not indicators of ability or potential. We can achieve an overall measure by totalling test marks across the English, maths and science tests at KS2 (or at KS3, or across all GCSEs at age 16). However school performance is much more susceptible to factors external to the child such as the extent of curriculum covered in class or the quality of teaching received. Performance in an attainment test, particularly a poor performance, may therefore have little to do with a pupil's ability. These concerns may be particularly pronounced when it comes to the KS2 and KS3 science and mathematics tests, which are most closely allied to the school curriculum, and for GCSE subject grades. For example, Strand (2004) showed that the school level accounted for up to 16% of the variance in some GCSE subjects, but only 2% of the variance in pupils' non-verbal reasoning test scores[7].

It is important to stress that the use of reasoning tests should in no way be seen to imply assessment of innate ability or fixed capacity. In their contribution to the Swann report, Mackintosh & Mascie-Taylor (1985) reporting the following:

"If a child has been deprived of intellectual stimulation or educational opportunity, it is small wonder that his intellectual performance will reflect this fact. An IQ test is no more able to gauge a child's true innate potential regardless of the circumstances of his upbringing than is a pair of scales to measure his true potential weight regardless of what he has been fed" (Mackintosh & Mascie-Taylor, 1985, Swann Report, Annex D).

However reasoning tests do differ from attainment tests in at least three regards. First, reasoning tests focus on very familiar content, basic elements such as simple words or sentences, numbers or number operations, or shapes and geometric forms, while the content of attainment tests is drawn from a direct sampling of what is taught to any particular age group in school. Reasoning tests therefore minimise the effect of specific curricular experience by drawing on general, not specialised, knowledge that individuals in a particular age group could have acquired from a broad variety of experiences in or out of school. Second, reasoning tests emphasise the perception and manipulation of relationships using these basic elements, abilities which should transfer or be applied to a wide range of tasks, while attainment tests

---

7. In a multi-level model for over 10,000 pupils in three consecutive cohorts attending 25 secondary schools, the school level (as opposed to year group or pupil level) accounted for 2% of the variance in non-verbal reasoning test score, 3.7% in quantitative or verbal reasoning score, 6% in GCSE total points score and English and science grade, 8% in design & technology and history, 9% in geography, 11% in mathematics and French, and 16% in art.

tend to measure specific outcomes of learning and instruction. Third, national tests are 'high stakes' for schools and published in performance tables while reasoning tests are 'low stakes' tests, used for management purposes within the school. A particular problem with high stakes tests is the pressure that can accrue to 'teach to the test'. For further detail see Strand (in press).

Therefore it is important that if attainment tests are used, reasoning tests are also used. Identification of gifted students would not be served well if identification was solely based on national tests.

There are some technical issues to resolve. For example, what tests should be included, should it be English and maths only, or should the science tests also be included? Second, what should be done when pupils are missing data from some of the tests? Around 2% of pupils national are absent for at least one of the KS2 tests, possibly due to sickness, and this rises to over 6% for the KS3 tests (Autumn package, 2004). Third, is it necessary to include some form of age adjustment so that older pupils are not identified to the exclusion of younger students.

KS2 tests are not one of the tests specified in the current eligibility criteria. This is a significant omission. We propose that KS2 tests should be included, using total test marks to identify the top 5% of pupils in the age group. While the tests are limited to two particular age groups (age 11 and 14 years) other elements of the framework include tests that can be applied throughout the age range 11-19. Including attainment tests alongside reasoning tests is likely to raise the total proportion of pupils identified above 5%, since some pupils may do well in one kind of test but not in the other. However this may be a price worth paying. Publishing the cut-point for KS2 total test marks will allow secondary schools to identify which of their Y7 pupils have performed exceptionally well in the KS2 tests.

As with any test we should be cautious about the apparent precision of KS2 total test marks (ranging from 0 to 280 marks) since any line drawn in a continuous distribution is subject to threshold effects. It would be non-sensical to consider a pupil above a certain mark 'gifted' but a pupil with one mark less as 'not gifted'. This reiterates the point made in the section of validity that the eligibility criteria do not define giftedness. They are only indicators that the student may have the potential to benefit from some of the distinctive provision offered by the academy, but are no more than indicators.

Equity issues in identification

Most research on equity issues in assessment has focussed on tests and examinations. Concerns are often expressed that test results may be biased or unfair to certain groups. However it is important to remember that objective testing has historically been seen as an instrument of equity. "*Examinations were the obvious method of attacking patronage, the hitherto dominant mode of recruiting to all forms of government (Sutherland, 1996, p16)*.The notion of the standardised test as a way of offering impartial assessment is powerful, though if equality of educational opportunity does not precede the test, then the fairness of the test is called into question.

Tests versus teacher assessment

Less research has been done on equity issues in relation to teacher assessment. However there is evidence that teacher assessment can be systematically biased in relation to student characteristics, such as gender, ethnicity, social class or behaviour, and that test scores may be less likely to suffer distortion from irrelevant extraneous pupil characteristics.

For example in relation to gender, Delap (1995) compared teachers predicted 'A' level grades for 7,000 students from approximately 450 schools against the grades actually achieved. Results indicated that teachers significantly over-predicted the grades achieved by girls relative to boys in three subjects (biology, geography and mathematics). Other research (e.g. Spears, 1984) has shown that written work can be differently assessed by teachers according to whether the student is known to be male or female, or from an ethnic minority. As a result of such concerns, examination boards now routinely remove names from scripts to prevent possible downgrading of girls and ethnic minority candidates.

Behaviour can also have a distorting influence. Thus Bennett et al (1993) report that teachers perceptions of student's behaviour constituted a significant component of their academic judgements. In other words, students who were perceived as exhibiting bad behaviour were judged to be poorer academically that those who behaved satisfactorily, even after controlling for test score and gender. The authors conclude there is a "need to supplement teacher judgments with other objective evidence of academic performance when important decisions about students are made" (p353).

Finally, Thomas et al (1998) analysed the impact of a range of student characteristics including entitlement to free school meals and English as an Additional Language (EAL) and Special

Education Needs (SEN) on teacher assessment at age 7, after controlling for the performance of the pupils in the standard tests. Even after test score had been accounted for, each of these student characteristics still had a significant impact on the teacher assessment. Teacher assessment significantly underestimated the pupil's attainment relative to test scores.

Special educational needs

There is a growing debate over the identification of highly able children with special educational needs (SEN) and the concept of multiple exceptionalities or double disadvantage. The issue is complex since there is debate about the definition of both high ability and many types of SEN. For example there is much scepticism around conditions such as Attention Deficit / Hyperactivity Disorder (ADHD) and some researchers would consider that cognitive impairment contradicts high ability (Winstanley, 2005).

However by including both attainment and ability tests within the criteria, the NAGTY framework allows for some flexibility. Consider for example dyslexia. Dyslexia is itself a highly contested condition, as witnessed by the recent debate around the work of Julian Elliott at the University of Durham (2005). However a discrepancy profile, with reading scores significantly below ability test scores, is still the most widely used criterion for identifying a discrete group of 'dyslexic' students as distinct from 'common or garden' poor readers (Siegel, 2003). Indeed it is the basis of the legal definition of Specific Learning Difficulties (SpLD) in many US states. Since reasoning ability is admissible as one of the criteria for membership, then such dyslexic pupils (so defined) will not be disadvantaged by their relatively low reading or language scores.

The debate is complex and well beyond the scope of this paper, readers are referred to Winstanley (2005) for further discussion. It is clear that further research is needed in this emerging area.

Monitoring

The upshot of this discussion is that it is important for the academy to monitor the nature of the take-up and the composition of the academy membership. A recent paper by Campbell et al. (2005) provides a detailed geo-demographic analysis of the students admitted to NAGTY in 2002/03 and 2003/04. The analysis shows that the National Academy, whilst having a student membership skewed towards groups with high levels of cultural and economic capital, had nonetheless reached significant numbers of students in the poorest areas, something over

1000 students, and 9% of the student members. This is comparable with the proportions of such students achieving the highest ‘A’ level grades and the proportion entering university.

It is important that this monitoring and evaluation continues. Of particular importance is to monitor the subsequent achievement of students identified through different types of evidence. For example, those identified through the various different test routes and those identified through teacher reference. Jesson (2005) reports that only 3% of the 30,000 pupils in state primary schools identified in his analysis of KS2 test scores were entitled to FSM, against an average for secondary schools of 14%. We need further research on the extent to which differences associated with ethnicity, sex, entitlement to FSM, EAL or age exist for the cohort identified through KS2 test scores.

We also need to know whether such group differences are similar across other criteria such as reasoning ability tests, and for teacher reference. NAGTY should commission research to investigate these issues. The availability of national data from the Pupil Level Annual School Census (PLASC) makes this practical in ways which were previously impossible and represents a unique opportunity to better understand these complex issues.

Cost and resources implications

The academy has in the past considered the option of creating its own entrance test for academy membership. However a single national test would be costly to develop. This would require not only development of a customised test, probably on an annual basis, but also scoring and marking services and a national network of test centres. It is arguable whether such costs could be justified, since funds would necessarily have to be diverted from delivering actual provision and activities to members. Nor would it be desirable to require families to pay for access to the test. As soon as access becomes determined by the ability to pay then all sorts of equity issues arise. Of course it could be argued that ability to pay already operates for some criteria, and of course subsidies could be provided for pupils from schools located in disadvantaged areas. However the total cost and logistical issues would remain, whether or not access was subsidised for some groups.

For the above reasons the creation of a single NAGTY test is not recommended. A framework approach, whereby variously widely used tests are endorsed for the purpose of identification, subject to the submission of appropriate technical data from the test publishers (see Appendix 3) is recommended. It is important to note however that the framework approach might be undermined if the tests became widely perceived as ‘high stakes’ tests, for example if

significant social or educational advantage were perceived to accrue from academy membership *per se*. In such circumstances schools might come under sustained pressure from some parents to secure nomination. In such circumstances a secure NAGTY test, created anew each year, might be needed to prevent direct teaching to the specific tests included in the framework. The teacher reference route would also have to be eliminated. The role of NAGTY membership within the University Admissions process must be clarified. If NAGTY membership is a box to be ticked on the UCAS form, and this is perceived to improve the chances of admission to university or to an elite institution, this will undermine the validity of the assessment process.

Conclusions

We recognise that no approach to identifying gifted students will be perfect. The aim is to create the most appropriate procedure balancing the competing criteria of:

- Fitness for purpose
- Reliability and validity
- Manageability
- Equity and access
- Cost (including time & resources)

The proposed process meets these criteria by utilising a range of tests already widely used in schools, and national end of key stage tests administered to the whole age group. This has a number of advantages.

1. First, it requires no extra testing. The criteria include a range of tests already widely used in schools, as well as national end of key stage tests administered to whole age groups. This is important so as not to raise the test burden on pupils which has become a concern in recent years (Assessment Reform Group, 2002).

2. Coverage is good. For example almost all secondary schools in England are using one or other of the leading reasoning tests, CAT or MidYIS. For the very small minority of schools not using either of these tests, then there are of course the national tests. A very small number of schools, drawn exclusively from the independent sector, or the small proportion of pupils educated otherwise may not complete the national tests. However there are currently a separate set of guidelines for independent schools of for

parents/guardians of pupils educated otherwise. These individual would be free to seek an independent psychologists report if they wished to apply to NAGTY.

3. This process, using school based evidence, is likely to be firmly embedded within the schools own process for identifying their top 10% of gifted pupils. Schools are not required to use these tests when determining their own top 10%, but many do. In any case considering the relationship between test scores and other measures is developmental in itself, probably leading into deeper insights into the strengths and weaknesses of test or non-test information in isolation and building capacity within the profession. There is therefore a coherence across the school identification and NAGTY admissions criteria.

4. It allows for flexibility as to the time of identification since many of the tests span the whole of the age range 11-18.

5. Because the schools are already using these tests there are no additional funding implications, in contrast to crating a single NAGTY test as detailed above.

Recommendations

Specific recommendations were given at the end of the executive summary.

References

Assessment Reform Group (2002). Available on the World Wide Web at www.assessment-reform-group.org.uk.

Ball, S., Bowe, R. and Gewirtz, S. (1996) 'School choice, social class and distinction: the realisation of social advantage in education,' *Journal of Education Policy,* 11(1) 89-112.

Bennett, N., Desforges, C., Cockburn, A., & Wilkinson, B. (1984). *The quality of pupils' learning experiences*. Hove: Lawrence Earlbaum.

Bennett. R. E., Gottesman, R.L. Rock, D.A., & Cerullo, F. (1993). Influence of behaviour perceptions and gender on teachers' judgements of students academic skill. Journal of Educational Psychology, 85, 347-356.

Bourdieu, P. (1998). T*he State Nobility: Elite Schools in the Field of Power*. Cambridge: Polity Press.

Campbell, R.J., Eyre, D., Hewston, R., Muijs, R.D., Neelands, J. & Robinson, W. (2004). The social origins of students identified as gifted and talented in England: A geo-demographic analysis - Occasional Paper 3. NAGTY: University of Warwick.

Delap. M. (1995).Teachers' estimates of candidates performance in public examinations. Assessment in Education, 2, 75-92.

Eyre, D. (1997). Able children in ordinary schools. London: David Fulton.

Freeman , J. (1998) *Educating the Very Able: Current International Research.* London: OFSTED.

Gardner, H. (1983). *Frames of Mind: the Theory of Multiple Intelligences*. New York: Basic Books.

Gardner, H. (1999) *Intelligence Reframed: Multiple Intelligences for the 21st Century.* New York: Basic Books.

Hany (1993). How teachers identify gifted students: feature processing or concept based classification. *European Journal for high ability, 4, 196-211.*

House of Commons, (1999) Education and Employment Committee, Third Report. *Highly Able Children.* London: HMSO.

Jesson, D. (2005). Identifying gifted and talented pupils - issues for all schools. *Paper presented to the Special Schools and Academies Trust annual conference, November 2005.*

Klein, P. D. (1997). Multiplying the problems of intelligence by eight: A critique of Gardner's theory. *Canadian Journal of Education, 22, (4), 377-394.*

Klein, P. D. (2003). Rethinking the multiplicity of cognitive resources and curricular representations: alternatives to 'learning styles' and 'multiple intelligences'. *Journal of Curriculum Studies, 35, (1) 45-81.*

Lyth, A. & Coe, R. (2005). Paper presented to the Expert Group on Assessment, September 2005.

Mackintosh, N. J., & Mascie-Taylor, C. G. N. (1985). The IQ question. In: *Education for all: The Swann report, 126-163, Cmnd 9453. London: HMSO.*

Montgomery, D. (1996). *Educating the able*. London: Cassell.

Morgan, H. (1996). An analysis of Gardner's theory of multiple intelligence. *Roeper Review, 18, 263-269.*

Nebesnuick S. (1993). *The secondary transfer of able pupils*. MEd dissertation, Oxford Brookes University, UK.

Neisser, U., Boodoo, G., Bouchard, T.Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51, (2) 77-101.*

OFSTED (2001). *Providing for gifted & talented pupils: An evaluation of Excellence in Cities and other grant-funded programmes*. London: OFSTED.

OFSTED (2005a). *Excellence in Cities: Managing associated initiatives to raise standards (HMI 2595).* London: Ofsted.

OFSTED (2005b). *Standards and Quality 2004/05.* The Annual Report of Her Majesty's Chief Inspector of Schools. London: The Stationery Office.

Pocklington, Fetcher-Campbell & Kendall (2002). *The gifted & talented strand of Excellence in Cities. Report 04-1/2002*. Slough: NFER.

Power, S., Whitty, G., Edwards, T., & Wigfall,V. (2003).*Education and the middle class*. Buckingham: Open University Press.

Renzulli, J.S. (1977) *TheTriad Enrichment Model. A Guide for developing defensible programs for the gifted and talented*, Mansfield Centre, Cn: Creative Learning Press.

Siegel, L. S. (2003). IQ-Discrepancy definitions and the diagnosis of learning disabilities: Introduction to the special issue. *Journal of learning disabilities, 36, (1), 2-3*.

Spear MG (1984) Sex bias in science teachers' ratings of work and pupil characteristics. *European Journal of Science Education* **6:** 368-377.

Stobart, G., & Gipps, C. (1997). *Assessment: A teacher's guide to the issues. (3rd Edition)*. London: Hodder & Stoughton.

Strand, S. (2003). *Getting the Best from CAT: A Practical Guide for Secondary Schools.* London: nferNelson. ISBN 0-7087-0376-3.

Strand, S. (2004). Consistency in reasoning test scores over time. *British Journal of Educational Psychology, 74, (4), 617-631*.

Strand, S. (in press). Comparing the predictive validity of reasoning tests and national end of Key Stage 2 tests: Which tests are the best? *British Educational Research Journal, in press.*

Thomas, S., Madaus, G. F., Raczek, A. E. & Smees, R. (1998). Comparing Teacher Assessment and Standard Task Results in England: the relationship between pupil characteristics and attainment. *Assessment in Education, 5, 213-246*.

Tempest, N. R. (1974). *Teaching clever children 7-11*. London: Routledge & Keagan Paul.

Times Education Supplement (2005). *Hunt is on for gifted pupils*. 2nd December 2005, p7.

Times (2005a). "*Brightest are failed by state schools*". Times, 25th November 2005.

Times (2005b). "*Best pupils fast-tracked in university talent quest*". Times, 5th December 2005.

White Paper (2005*). "Higher standards, better schools for all: More choice for parents and pupils"*. London: HMSO.

Winstanley (2005). *Children with multiple exceptionalities*. Occasional Paper 3, NAGTY: University of Warwick.

APPENDIX 1: Current NAGTY Eligibility Criteria 2005/06

Evidence submitted should be one of the following:

Test Evidence	**Non-test Evidence**
Cognitive Ability Tests – a score of 126 or above in one battery or 120 or above in two batteries. MidYIS – a mean score of 126 or above. Yellis – scores of 78% or above in Year 10 or of 86% or above in Year 11, or a standardised score of 126 or above. Key Stage 3 SATs – an examined Level 8 in Mathematics or a Level 8 teacher assessment in another subject at the end of Year 9. UK Maths Challenge – Gold Award. World Class Tests – Merit or Distinction. GCSEs – a points score of 58 or above in the best eight subjects (where A* = 8, A = 7, B = 6 etc.). This translates to 428 points on the new QCA points system (where A*=58, A=52 etc). Other standardised test of general ability, e.g. other nferNelson test – a standardised score of 126 or above. US Scholastic Assessment Tests (SATs) – outstanding performance when age-adjusted.	Reference from a teacher or other education professional. This may be for a particular subject or where perceived potential is not demonstrated in assessment scores. References must be on headed paper, signed, and should clearly indicate the criteria on which the judgement is based. **Please refer to our Guidelines for References before submitting a written reference.** Evidence of outstanding achievement in an academic-related activity pursued outside school, e.g. success in a national-level chess or debating competition or outstanding performance in a master-class, Aim higher or gifted & talented programme. Independent assessments identifying ability in the top 5%, e.g. an educational psychologist's report.

APPENDIX 2: Examples of the subject specific guidance from the Qualifications and Assessment Authority (QCA) for the identification of G&T pupils in science, modern foreign languages and ICT.

Pupils who are gifted in science are likely to:

- be imaginative
- read widely, particularly science or science fiction
- have scientific hobbies and/or be members of scientific clubs and societies
- be extremely interested in finding out more about themselves and things around them
- enjoy researching obscure facts and applying scientific theories, ideas and models when explaining a range of phenomena
- be able to sustain their interest and go beyond an obvious answer to underlying mechanisms and greater depth
- be inquisitive about how things work and why things happen (they may be dissatisfied with simplified explanations and insufficient detail)
- ask many questions, suggesting that they are willing to hypothesise and speculate
- use different strategies for finding things out (practical and intellectual) -- they may be able to miss out steps when reasoning the answers to problems
- think logically, providing plausible explanations for phenomena (they may be methodical in their thinking, but not in their recording)
- put forward objective arguments, using combinations of evidence and creative ideas, and question other people's conclusions (including their teacher's!)
- decide quickly how to investigate fairly and manipulate variables
- consider alternative suggestions and strategies for investigations
- analyse data or observations and spot patterns easily
- strive for maximum accuracy in measurements of all sorts, and take pleasure, for example, from reading gauges as accurately as possible (sometimes beyond the accuracy of the instrument)

- make connections quickly between facts and concepts they have learned, using more extensive vocabulary than their peers
- think abstractly at an earlier age than usual and understand models and use modelling to explain ideas and observations. For example, key stage 3 pupils may be willing to apply abstract ideas in new situations; key stage 4 pupils may be able to use higher-order mathematical skills such as proportionality, ratio and equilibrium with some complex abstract ideas when offering explanations
- understand the concepts of reliability and validity when drawing conclusions from evidence
- be easily bored by over-repetition of basic ideas
- enjoy challenges and problem solving, while often being self-critical
- enjoy talking to the teacher about new information or ideas
- be self-motivated, willingly putting in extra time -- (but they may approach undemanding work casually and carelessly)
- show intense interest in one particular area of science (such as astrophysics), to the exclusion of other topics.

Pupils who are gifted in modern foreign languages are likely to:

have a strong desire to put language together by themselves
they apply principles from what they have learned to new situations, transforming phrases and using them in a different context, often with humour

show creativity and imagination when using language
they often extend the boundaries of their knowledge and work beyond what they have learned, not wishing simply to respond and imitate, but to initiate exchanges and to create new language

have a natural feel for languages
they are willing to take risks and see what works, knowing instinctively what sounds right and what looks right; they are acutely and swiftly aware of the relationship between sound and spelling

pick up new language and structures quickly
they may have excellent aural and oral skills and may be able to cope with rapid streams of

sound and identify key words at an early stage; they may also display outstanding powers of retention, both immediately and from one lesson to the next

make connections and classify words and structures to help them learn more efficiently
they are able to evaluate new language critically, recognising the grammatical function of words

seek solutions and ask further questions
they may test out their theories and seek to solve linguistic problems, sometimes challenging the tasks set and trying to understand their relevance to the language-learning process

have an insight into their own learning style and preference
they may say how they like to learn vocabulary or structures; they are clear about the type of tasks they like doing; they may show or display an ability to work independently, without supervision, and to make effective use of reference material

show an intense interest in the cultural features of the language being studied
they may use idiom in the language itself and explore the history and the traditions of the language; some pupils may wish to share their knowledge with their peers

Pupils who are gifted in design and technology are likely to:

Demonstrate high levels of technological understanding and application

- display high-quality making and precise practical skills
- have flashes of inspiration and highly original or innovative ideas
- demonstrate different ways of working or different approaches to issues
- be sensitive to aesthetic, social and cultural issues when designing and evaluating
- be capable of rigorous analysis and interpretation of products
- get frustrated when a teacher demands that they follow a rigid design-and-make process

- work comfortably in contexts beyond their own experience and empathise with users' and clients' needs and wants.

Teachers may identify pupils who are gifted in design and technology by:

- performance at an unusually advanced national curriculum level for their age group
- the outcomes of specific tasks
- evidence of particular aptitudes
- the way pupils respond to questions
- the questions that pupils ask themselves.

It is important for teachers to allow time for personal interaction with pupils. By observing the techniques and strategies that pupils use to tackle problems, teachers may pick up on gifts that do not come to light through more formal assessment procedures. It is important to acknowledge that these pupils may wish to hide the extent of their gifts.

The pupils who are gifted in design and technology may be a very different group from those with gifts and talents in other subjects. The breadth of designing and making means that some of them will have abilities in a specific area -- for example working with food, using computer-assisted design (CAD) or high-quality making -- but not in others.

For details of the guidance of identifying gifted and talented learners for all subjects see the following webpage (accessed on 28/02/06): http://www.nc.uk.net/gt/index.html

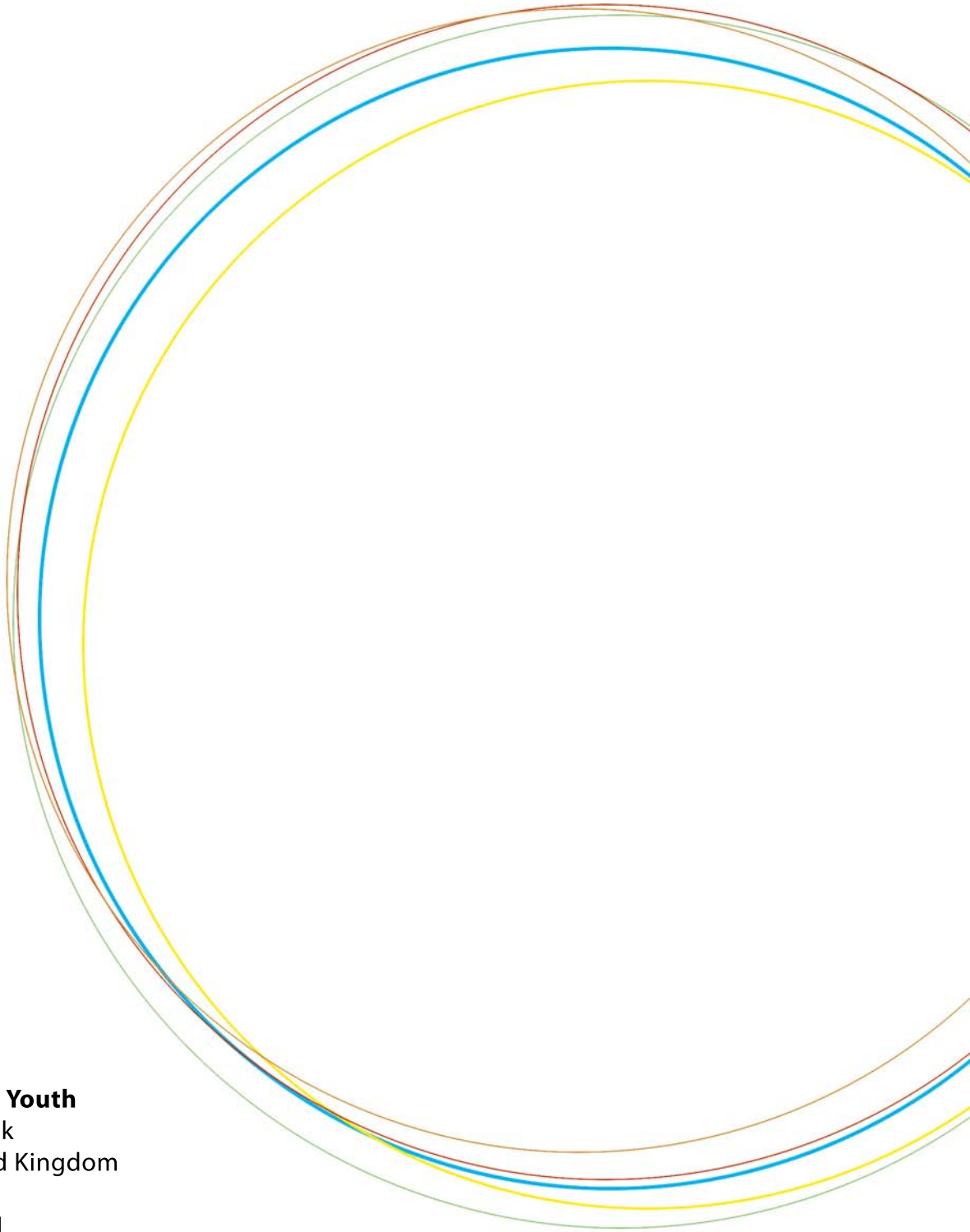
APPENDIX 3: Proposed identification criteria for NAGTY membership from Sept. 2006

Test	age range	Criteria
KS2 total test marks (decimalised levels)	11	determined annually
KS3 total test marks (decimalised levels)	14	determined annually
CAT3 verbal reasoning, quantitative reasoning and non-verbal reasoning tests	7:6 - 18:00	standard age score >=129
MidYIS vocabulary, maths and non-verbal reasoning tests *(a)*	11-12	standard age score >=129
Yellis mean score*(a)*	14-15	standard age score >=126
World Class Tests (Problem Solving)	9-14	merit or distinction
SAT		tba
NFER VR series and NVR series	7:00 - 14:00	mean standard age score >=129
Hodder & Stoughton VR series and NVR series	6-13	mean standard age score >=129
Test of general cognitive ability administered by a chartered educational psychologist (e.g. WISC, Stanford-Binet, BAS)	7:00 - 18:00	BAS GCA standard score >=126 Stanford Binet and WISC performance at or above the 95 centile
GCSE\GNVQ best 8 capped points score	16	58 or above on 'old points' or 428 or above on the new QCA points score
Other ability test allied to university entrance (e.g., TGA)	16-18	performance at or above the 95th centile
A/AS UCAS points	18	428 or above

(a) pending further discussion with CEM about the reliability of Yellis sub-tests.

All test publishers to provide documentary evidence of:

- reliability coefficient of at least 0.90 or above for the test;
- details of the sample employed in the standardisation process, giving evidence that it is nationally representative;
- a standardisation in the last 10 years, or evidence of equating to more recent norms.



The National Academy
for Gifted and Talented Youth
The University of Warwick
Coventry CV4 7AL United Kingdom
Tel: +44 (0)24 7657 4213
Fax: +44 (0)24 7657 4221
Email: gifted@warwick.ac.uk
Web: www.nagty.ac.uk

February 2006
ISBN Number: 1-905877-10-2

THE UNIVERSITY OF WARWICK

department for education and skills