Chapter 3. Glittering Prizes for Merit

'I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin, but by the content of their character. I have a dream today.' *(loud cheers)*[*]

3.1 An example of random distribution: University entrance

When a public organisation selects a winner from a group of eligible applicants, there is considerable interest in the rightness of the process. In a world where selection on merit is held as the ideal, it is instructive to encounter a deviation from this norm. Universities are public bodies, in receipt of state-funding. When they have to choose who should be allowed on popular courses, and of course, who should be rejected, how should they decide? These are not trivial decisions—the award of a place at a prestigious university or to train for a well-rewarded profession is the 'glittering prize' that can lead to fame and fortune.[#]

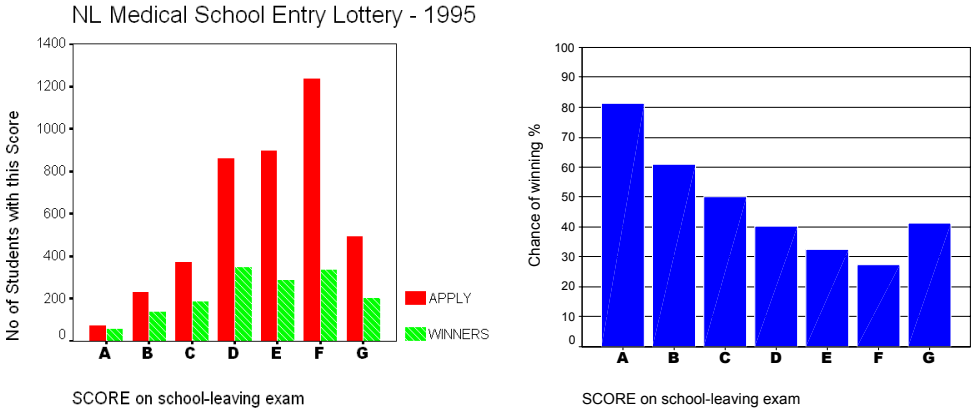
[*] (Martin Luther King speech delivered on the steps at the Lincoln Memorial in Washington D.C. on August 28, 1963. Source: Martin Luther King, Jr: The Peaceful Warrior, Pocket Books, NY 1968)

[#] 'Glittering Prizes' is the title of a 1976 BBC2 serial by Adam Raphael, about a group of young Cambridge graduates winning successful careers in the media.

Example of Random Distribution:

Medical School Entrance in the Netherlands

How the Dutch medical-school entry system operates: Because of a rigid streaming system in Dutch schools, only the top 10% are eligible to apply for university medical courses. Pupils leave school with grades from a nationally-based examination, plus an achievement test provided by their school. Universities are not allowed any other screening devices (interviews, special tests, references, extra-curricular achievements). Where demand exceeds supply, a central committee decides the allocation process. During the 1990s in the case of medicine, about 5,000 qualified pupils applied for the 1,800 places available. The graphs below illustrate the way in which the lottery, weighted by students' grades is used to allocate applicants to courses at the Dutch universities:



The graph on the left shows how the 1995 Dutch Medical School entry lottery worked. The 'SCORE' category relates to results on school-leaving examinations, combining both national tests and a teacher-based assessment. 'A' is the top category, 'F' the lowest, although this is relative: Only pupils from the top 10% are eligible to apply. The category 'G' relates to non-standard entries, such as those with non-Dutch qualifications. The second chart shows how the chances of winning a place vary with the Score achieved. The system allowed students to make repeated applications, and there was provision for appeals.

Source: Report of the Drenth Commission, 1999

3.2 Discussion on the entrance lottery for university courses in the Netherlands

In this case the universities have ceded control of entry to a government agency which allocates students to courses. The use of a lottery as part of this process has been reported by Elster (1992) and others, but the fullest English-language description can be found in a 1999 Report to the Irish ministry of Education, who commissioned Professor Piet Drenth to describe the Dutch system. The system used applies to three courses where demand for places greatly exceeds the supply—medicine, dentistry and veterinary studies. The selection mechanism in use is a hybrid one combining a measure of merit together with a weighted lottery. (This is sometimes called a 'graduated lottery' or 'graduated random distribution'.)

In operation since 1972, the mechanism came under intense scrutiny in 1996 when a very bright student (Meike Vernoy) was rejected for medical school entry, despite gaining near-top grades in her school-leaving tests. Her case became a national *cause celèbre*, and under political pressure the Dutch minister of education set up a commission (*Commissie Toetalting Numerus Fixus*) chaired by Professor Drenth. The Drenth Commission examined and evaluated the existing system and suggested modifications. Its Report published in 1997 stated that the existing system was sound and should not be changed. The Drenth Report provides a useful examination of an existing lottery-based allocation mechanism. The evidence which it collected and presented makes a formidable case for the appropriate use of some form of merit criterion with a weighted lottery being the final arbiter.

Drenth tested the ability of entry scores to predict performance on the course. At the end of the first level, it was found that entry scores gave some indication of time taken to complete the level, and also the success rate. By the time of the finals, this variation had practically disappeared. From this Drenth concluded that the Dutch system is *not* characterised by too many falsely accepted students (who then go on to fail). Rather that far too many students who would have succeeded have been rejected. Drenth, it

seems, would have liked that the lottery aspect was strengthened, with less emphasis on school-leaving scores.

Drenth also examined alternative entry systems which are used in other countries, especially those related to medical school entry: These include greater use of school-leaving scores, special aptitude tests (such as SATs in the US), psychological tests, interviews, references and the use of probation periods. Apart from school-leaving results, none had much useful predictive power, with interviews and references especially useless.

In response to the Drenth report, the Dutch government decided to stick with the basic system, but modify it somewhat. Top-scoring students (A, B and C) were to be automatically given places; the lower scorers would take their chance in a weighted lottery. The politicians had given in to the pressure from parents, rather than heed the considered advice of Professor Drenth. (In private correspondence with Prof. Drenth he tells me that the students, acting through their union, are still keen to promote the use of lottery selection)

The system in the Netherlands is unusual, and it is worth asking how and why it arose. 'The system stems from the pathological Dutch drive for fairness and their intense dislike for making tough decisions' is one (unattributed) quote given by Drenth (in private correspondence). He also informed me that the idea of using a lottery had no champion, no advocate who proclaimed its virtues, nor any academic who demonstrated its worth. The weighted lottery model emerged during the original debates in the early 1970s as a compromise between the leftists who wanted places to be provided for all students to study courses of their choosing, with excess demand settled through a lottery only. The more conservative parties supported by the employers, the medical professions and medical schools favoured selection based on predictors of success, with school-leaving scores the obvious indicator. Since neither side had a parliamentary majority, they compromised, with the use of the lottery, weighted according to school-leaving grades. The system had lasted 24 years without significant complaint, which is a testimony to its effectiveness as well as its robustness. Subsequently, after 1997, that it has largely survived both intense

criticism and ill-informed tinkering is highly encouraging to advocates of random allocation like myself.

Hofstee (1990) who is also Dutch, comments that the adoption of a mixed system of grade scores and a weighted lottery is 'apart from a political compromise, may be taken as testimony to the wisdom of the Dutch authorities.' Hofstee has also conducted research in the Netherlands into the 'acceptability' of lottery selection compared with other methods. Among potential students he found little enthusiasm for single selection mechanisms. In particular, the use of lotteries as a sole means of selection was highly unacceptable. Instead his respondents expressed a preference for mixed methods which involve educational grades, interviews, waiting lists, psychological tests; in short what Hofstee calls 'fuzziness and indeterminacy'. Later a similar questionnaire was administered to 100 Dutch psychology university students. Of particular interest, and in contradiction to Hofstee's earlier study, these students found a lottery to be a most acceptable mechanism for educational selection. As these were second year students, they, or at least many of their school-mates would have been through such a selection process. Their *only* exception to the acceptability of lottery selection arises in employment: For promotions and lay-offs these students thought a lottery mechanism would be unacceptable. Hofstee also refers to an earlier study in 1983 which found that Dutch youngsters preferred a weighted lottery in admission to *numerus clausus* (course with restricted entry) studies rather than either a straight lottery or selection by test scores only.

In the UK there have been some examples of random selection for university entrance reported: Jon Fuller, in charge of post-graduate entry to medical courses at QMC has adopted a lottery (as reported in *The Sunday Times* 14 Sep 2003). At both Leeds Metropolitan and Huddersfield universities students have been selected randomly for physiotherapy courses. (BBC, 27 Apr 2004). Even Schwartz, in charge of the review of entrance procedures in the UK played with the idea, if a report in The Times of 6th Sept 2003 is to be believed: 'Universities to pick students by lottery' was the top headline for that day. In the final report Schwartz (2004a) did not include this as a recommendation.

3.3 Educational Selection on Merit: the Ideal?

Prior to meritocracy, posts and places were awarded by patronage, nepotism, simony and other curious means. (A fuller description of the emergence of 'merit' is given in Appendix A). Parkinson (1958) of 'Parkinson's Law' fame described the introduction of competitive examinations for Civil Service entry as about the best system ever invented for selecting competent employees. It should be remembered that Parkinson was in a position to know, as he was employed as a bona fide management consultant. The word 'meritocracy' was famously coined about the same time as Parkinson's Law by Michael Young in his 1958 social satire 'The Rise of Meritocracy 1870 – 2033: an essay in education and equality'. 1870 was the date when the Trevelyan reforms of Civil Service exam-based entry were introduced. Young predicted that over-reliance on the admittedly highly reliable Intelligence Quotient (IQ) tests to allocate children to different schools, would lead to the stratification of society. In the end, by the year 2033 the proletariat would rise up against their helotry of the stupid. Young was not entirely correct in his interpretation of IQ tests, as I will explain later in the next section. (Although these books were intended to be humorous, they both had a serious intent, and reached a wide and influential audience).

Despite Young's warning, meritocracy is still seen as the model for a better society, where hard-working individuals are allowed to thrive on their merits, rather than who they know, or worse, who their parents were. 'Selection on merit' is widely accepted, particularly in educational circles as the highest ideal. So it is worth examining, firstly how 'merit' is measured, and secondly whether it works—how reliable is measured merit at identifying potential winners and losers.

In practice, selection on merit is a bureaucratic procedure where the element of merit may be determined by objective criteria (a test), but is often left to the discretion of the selectors. The following describes the familiar university selection procedure, which, it would be claimed, is based on merit alone.

Selection on Merit for University Entry: A Gate-Keeping Exercise

---

*Places become available**:*** Every year, University courses have places to fill. There is widespread information and assistance, so any potential applicant should find it easy to discover what is on offer. Minimum requirements will be published, which may cause many applicants to self-deselect.

Initial screening and prioritising: A further barrier may be interposed at this stage: Universities may require more than minimum grades before they consider a candidate, rejecting all below an artificial threshold. Making the applicant sit a test is another special form of screening: Aptitude or intelligence tests such as SATs in the US can be a major determinant of success. Administrators will also sift through the application forms, removing 'unsatisfactory' applications, and highlighting 'promising' ones.

The interview is often seen as the apogee of the selection-on-merit mechanism: Candidates who fulfil entry requirements are interviewed, usually by a panel of academics in the chosen field of study. Winners will be chosen on the basis of judgements made by the interviewing panel, combining assessment of the candidates' performance on the day, information from application forms together with the opinions ('references') of other people who may know something about the candidate.

Greely (1977) describes a particularly elaborate system used for entry to Yale University Law School. Three thousand applications are made for the 325 places available. Each application is read and ranked by three faculty members. It is relatively easy to identify the top and bottom candidates, but the real problem comes in spotting who fits into the 250^{th} to 350^{th} category, where differences in 'merit' will be insignificant. Attempting to pick the 'best' candidates is not just a costly business, it is in Greely's description a 'pretense'. He goes on to point out that random selection would be the fairest and cheapest method.

So can processes like this reliably identify merit? Is merit the *only* basis for awarding the prize of a place on an over-subscribed course?

3.3.2 Measuring Educational Merit: Intelligence and other aptitude tests: a scientific approach

In an effort to establish a more rational basis for selecting and rejecting candidates, and in particular identifying hidden talent, tests of intelligence were developed, starting over 100 years ago, most notably by Spearman. These tests have been in widespread use ever since. The Stanford-Binet test of IQ (Intelligence Quotient), has been widely used, and correlates well with human abilities. Kline (1991) (who is a notable critic of the use of IQ tests) admits that 'the application of psychometrics (IQ testing) is one of the few technological successes in psychology'. He concludes that 'If we take the correlation between intelligence and academic success across a whole range of ability it is likely to be substantial, around 0.5'—that 50% of ability and achievement can be explained by the score on an IQ test. So IQ tests and their close cousin the US SATs tests are valid, quite probably the best, and maybe the only way of identifying those with potential to succeed. The Economist (2005) makes a spirited defence of SATs: 'If universities admitted students purely on the basis of their grades and test scores, as they should, the proportion of successful poor students would actually go up rather than down.' This is not yet 'merit'—in Young's (1958) pseudo-formula he identified Merit as:

Young's (1958) pseudo-formula for M ('merit'):

$$\mathbf{M = I + E},$$

where I is measured IQ and E is effort.

('pseudo' because economists would prefer a formulation $M = f(I, E)$)

Young assumed that measuring both of these would become more reliable over time. In this he was wrong. Measuring IQ has improved a bit, but measuring Effort remains a highly subjective activity, based on human judgement by work-study practitioners.

There are two features of such tests which are often overlooked or mis-understood: they do not provide fully conclusive identification of merit, and (something Young

missed completely) beyond a certain level, tests have very little predictive power in separating out potential failures from those who might succeed.

Tests and error bounds: fuzziness in the measurement

The score on an IQ test is a good indicator of future academic performance. It is easy to think of the relationship as something like this (Figure 3.1):

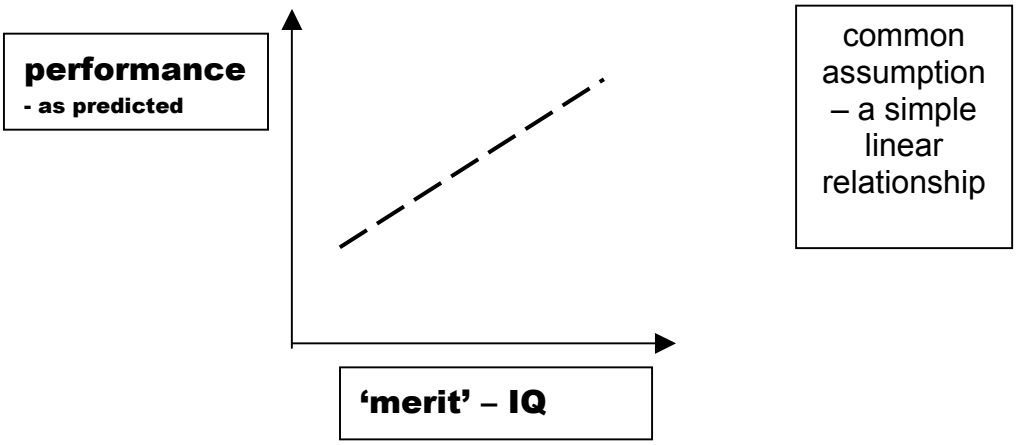


Figure 3.1: Simplistic relationship between measured merit and predicted performance

The graph above suggests that as the IQ score rises, that the Performance rises in exact proportion. Of course, most people are aware that measurement is not an exact science, and there will be fuzziness due to many factors. The relationship shown by Figure 3.1 will then look like:

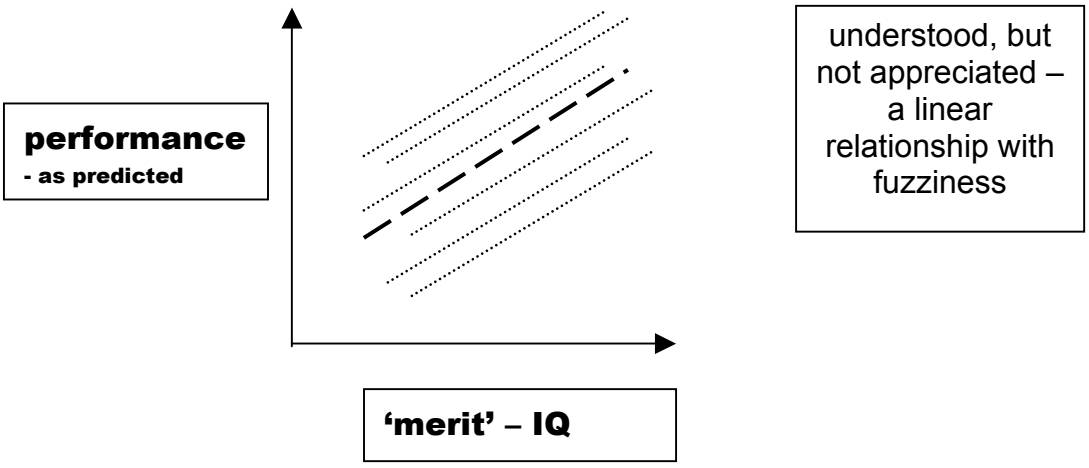


Figure 3.2: Realistically fuzzy relationship between merit and performance

Aware of the likely nature of this relationship, selectors, officials and the organisations will prefer higher 'merit' scores because that indicates a better *chance* of success. This is easy when demand exceeds supply. There is still the risk that because of the fuzziness of the relationship that some failures will slip through. Raising the entry threshold reduces that risk for the selectors. The applicants lose out with many of those, as Drenth indicated, being rejected, despite still having a good chance of succeeding.

To achieve a target quota of entrants, selectors may use an arbitrary score on an IQ test as a dividing line between pass and fail. The old English and Welsh 11+ IQ test was set up to decide who 'won' a place at Grammar School, or who 'failed', and was sent to a Secondary Modern, and was in operation for many years. Typically, the top 25% of scorers on the IQ Test went to Grammars, although the rates varied hugely. According to Vernon (1957), the strict cut-off point meant that many children were sent to the 'wrong' type of school. Because of the uncertainties in the measurement process, it was estimated that 20% of pupils finished up in a Grammar school when they should have been at a Secondary Modern or vice versa. Using more up-to-date information related to university students' performance, Bekhrandia (2002) looked at an entire student cohort, and discovered that there is a significant trend—better entry grades on average predict better final grades. But it is clear that there is much unpredictability in the system: An entrant with 18 points still has a 60% chance of doing as well or *better* than an entrant with 24 points. Elsewhere Bekhrandia (2003) produces evidence to show that pupils from the state sector do much better than those from independent (fee-paying) schools for the same A-level entry points. Independent school pupils need to gain an extra four A-level points to have the same expected degree. This could be taken as an objective criterion to discriminate between applicants.

Non-linearity: more is not always better

If the score on an IQ test or the level of examination grades are sound indicators of future academic performance, then it seems reasonable to assume that the higher the scores or grades, the more likely it is that a candidate will succeed. However in many

cases it is not like this. Performance may generally rise with IQ score, but then tends to level off as shown in Figure 3.3:

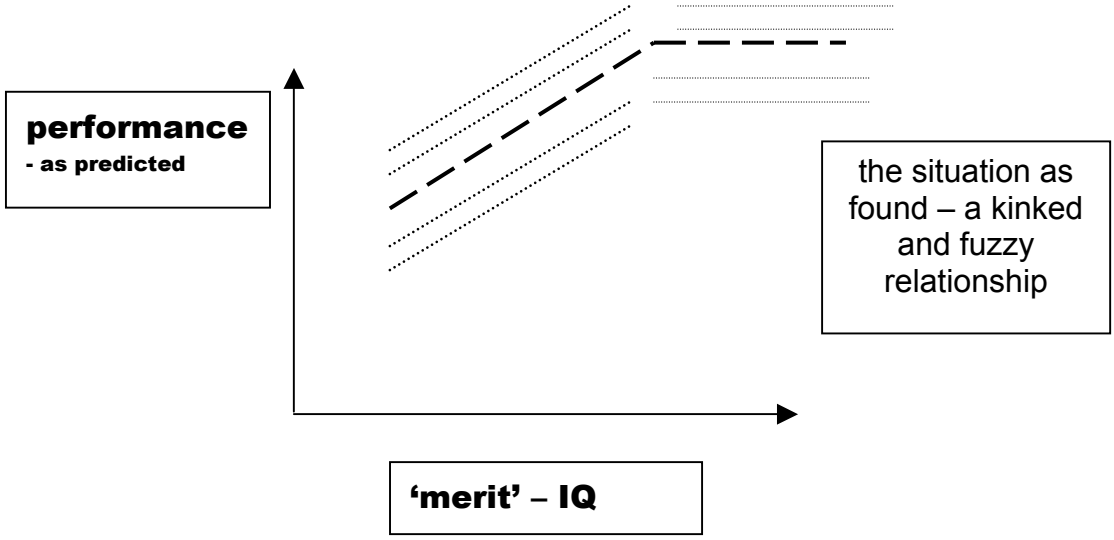


Figure 3.3 Complex reality: IQ scores fuzzily predict performance, up to a point.

There are many examples of entry tests or scores showing a linear, if fuzzy relationship up to a certain level, then flattening out after that:

Pilot training: War time pilot training, like much research based on large-scale military activity shows the non-linear characteristic. Eysenck (1962) showed with a simple graphic (p26) that pilot performance generally increased in line with IQ, but beyond a score of 120 there is scarcely any improvement.

University entrance NL: In the Netherlands Drenth (1999) described the performance of medical students as predicted by their entry grades. Because of the natural experiment provided by random selection, a representative cross-section of eligible students with a range of grades are accepted onto the courses. Drenth concludes that achievement in final secondary school examinations 'does have some, although not very strong, relationship with the study results in the medical studies, especially in the early years (of the course) and if time criteria (time taken to complete) are used. Other predictors have negligible correlations.' Drenth also points out that those in the lowest category for entry qualifications still have a good chance to succeed and finish their studies in a reasonable time.

University entrance UK: Having explained that there is a strong (0.50) correlation between measured IQ and academic performance, Kline (1991) states quite bluntly: (p9) 'if our sample is selected for intelligence (for example at a good university where all students have IQs beyond 120) then the correlation is bound to fail. Everyone has sufficient ability to do the work.' This view is supported by two more recent reports which asked how well A-levels predict final degree classification: Wiliam (2002a, b) studied the results of students graduating from his own institution, King's College. Wiliam concluded that using A-level points to predict class of degree is only slightly better than pure chance. (Since this is an elite university, then this result is in line with what Drenth found in the Netherlands).

The fallibility of human judgment

Since the interview is often the core technique for deciding who has the most 'merit' and should get the prize, the effectiveness of this method should be scrutinised closely. Officials doing the selecting tend to have a high opinion of their powers of judgment. It might be expected that schoolteachers with longstanding knowledge of their pupils could reliably predict their pupils performance. Not so. The predictive ability of the teachers was invariably worse than the 'quick and dirty' 11+ test. (Vernon, 1957). Camerer (1995) adds a much more blunt comment concerning the predictions by experts of post-graduate students' success: 'The faculty's deliberations just add noise'. Simple models, using measurable indices perform well. Adding human expertise seems to make the judgement *worse*. (A fuller extract of Camerer's views is given in an addendum to this Chapter)

Evidence of the ineffectiveness of interviewing as a means of selecting students was given by Steven Schwartz in a submission to the House of Commons select committee on education (2004b): He is quoted as saying '..interviews take place at some of our most ancient universities, and the reliability of these interviews is zero'. He referred to an experiment carried out at Cornell University (Kelman & Canger, 1994) where veterinarian applicants were selected, half with an interview, half at

random. Judging by the results at the end of the course, it was impossible to distinguish between the two groups. ‘To me, it [selecting by interview] is the same as flipping a coin.’ (I feel Schwartz was using this as a rhetorical device rather than as a policy prescription). Claims by admission tutors that their records showed that they were able to pick out high-flyer were dismissed as ‘an illusion’.

A study of the peer review of grant applications (Wessely, 1998) found that overall, the reliability of panels was reasonable: an experiment with a second panel confirmed 75% of the original outcomes. Individual reviewers were far less consistent, showing only ‘slight’ agreement amongst themselves. The amount of rent-seeking activity is also commented on with the reviewers spending an estimated 115 equivalent-years on applications in 1989, plus a much greater but uncalculated amount by the applicants.

A further problem related to interviews and other subjective selection techniques is that of discrimination. It would be wrong, and against university policy if admissions tutors were to actively prefer attractive young white female applicants over others who were equally qualified. This is an agency problem, and however well-intentioned, it is difficult to control this bias. Public Choice theory would assume that selectors would act in this discriminatory way for their own satisfaction, if given the discretion to do so. Even where selectors are acting with best intentions, and even following training to avoid such discrimination, there will still be unwitting bias. Beyond the recognised forms of discrimination on grounds of gender, race, age and perhaps sexual orientation there are many more human traits and features which either help or hinder candidates in interviews. These will be dealt with in more detail in Chapter 5.

3.4 Consumer Choices

Missing, or at least de-emphasised in many of the treatments of public organisations allocating benefits to applicants, is any notion of customer satisfaction. Roth (2002) describing the mechanism of allocation for interns to hospitals, sees the interns as

having awkward selections that need to be satisfied, but are also likely to engage in trickery to dupe the system. Tellingly, the Schwarz (2004a) enquiry which consulted 25 different organisations, only chose one —National Union of Students—which could be said to represent the views of the proximate customers of the universities. Drenth only revealed to me in private correspondence, not his report, that there were positive views on lottery selection by Dutch students. In all these reports, it is the efficiency of the system, primarily on behalf of the producers that matters, choosing who is best for them. Again Public Choice theory has an explanation: This is an example of 'producer capture', where the producers of the commodity run the distribution system for their own benefit, not their customers.

Perhaps this insouciance about customer's wishes derives from the view that 'beggars can't be choosers'—that applicants to universities are being offered a valuable prize, for which they pay well below the market price. All winners have had a boost to their well-being, so why worry if total consumer reward, both of the winners and the losers is not maximised? I will attempt to identify the net consumer benefit from allocation systems like university admissions, but that begs the question of 'Who is the consumer'? for places on university courses. Consumers are usually the ones who pay. Behind most students are families who are required to pay the majority of the (considerable) expense of a university course—although through student loans, this burden is being shifted more onto the student. Professional organisations and employers are frequently consulted about the content of courses, examination standards and admission criteria. In a sense, they are 'customers' for the product of university courses, and employability is a prime concern to applicants. Government, and the politicians who run it are 'customers' in the sense that, using public funds, they provide a large proportion of university revenues. With such an array of powerful, financially significant interest groups involved, it is not surprising that the views of students count for so little, and are virtually ignored in analysis and reports. Yet it is they who spend time and effort going through the process, and they as individuals who stand to gain or lose thereby.

A significant cost to applicants is the need to obtain higher grades in order to qualify for consideration to their chosen course. Putting in extra effort, or spending more time is an example of rent-seeking behaviour. It might be argued that augmenting one's education by gaining higher grades is a good thing, even for those who fail to get on the course of their choice: A better educated workforce can be more productive. Alternatively, it might be said that the extra time spent gaining better grades, would be better spent acquiring life-skills which would be far more useful in later careers. In a survey which I carried out in 2003 on economics students at UWS (details in Appendix B), my tentative conclusion about 'Rent-seeking' was that students had spent on average about two extra months of their life over and above the basic requirement to be adequately qualified for entry. By any calculation this is a significant cost.

Other aspects of rent-seeking might include behaviour likely to put the candidate into favour with the selector. There is a belief that selectors may be signalling some of the secondary criteria that may be taken into account for selection: Out-of-school activities involving charitable works or energetic outdoor pursuits are deemed worthy; they certainly appear on application forms. The good opinion of teachers is also important, because a reference is needed. This may induce conformist behaviour, and suppression of natural exuberance. In extreme competition for coveted university places some candidates may even deliberately sabotage a perceived rival's work.

Satisfaction with the process of selection is much more difficult to ascertain, but the effort is surely worthwhile. If an alternative mechanism, like appropriate merit combined with a weighted lottery, along the lines of the Dutch system were on offer, then it could be studied. Clearly the Dutch system should reduce the wasted effort of rent-seeking. A constant refrain of those who examine university entrance is that of 'fairness'. Exactly what this means in this context—what is fair?, and in fairness to whom?—is not at all clear. I will return to the abstract philosophical notion of fairness later in Chapter 7. There have been some highly significant developments in the literature of economics which may shed light on this.

3.5 Conclusion: What would a valid test of merit be like?

It is clear that simple tests of ability are vital in identifying 'merit', in the sense of possessing potential to succeed. Grades on examinations are useful measures for such merit and should continue to be used. It is to be hoped that researchers will continue to refine such indicators, the better to assist admissions tutors in their selection, although too much should not be expected—100 years of development have not added greatly to the power of such tests. The relationship between the validated indication of a test, and the ability to correctly choose from a pool of applicants is poorly understood: When large numbers of qualified applicants present, it is *not* appropriate to raise the threshold, and demand higher grades. Some such as Astin (1985) take this partial failure of tests to predict reliably as a good reason to do away with selection altogether. Goldstein made a similar comment on an earlier paper of mine (Boyle, 1998). This is wrong: We should apply validated knowledge where it exists, and admit when our knowledge runs out. At this point other criteria may be applied, hopefully in a transparent manner, but apart from a lottery amongst qualified candidates, it is difficult to envisage any alternative, defensible method of discrimination.

Legislating for such a form of selection processes would not be dangerous novelty. Legislators have for long made piecemeal efforts to make the selection process fair. Since all organisations both public and private owe a great deal to the state that nurtured and supports them, rationalising this interference should not be seen as an onerous new burden, rather a clarification.

Addendum to Chapter 3: Camerer on Judgements by Experts:

Colin Camerer (1995) (p 611-2) puts it more directly: 'A body of literature concerns judgments made repeatedly by people (many of them experts) in natural settings where stochastic outcomes depend on some observable predictors (e.g., test scores) and some unobservables. Examples include medical or psychiatric diagnosis (severity of Hodgkins' disease, schizophrenia), predictions of recidivism or parole violation by criminals, ratings of marital happiness, and bankruptcy of firms. About 100 careful studies have been documented so far**.** The remarkable finding in almost all these studies is that weighted linear combinations of observables predict outcomes better than individual experts can (Meehl, 1954; Dawes, Faust, and Meehl, 1989). In a typical study (Dawes, 1971), it was discovered that academic success of doctoral students could be predicted better by a sum of three measures—GRE scores, a rating of the quality of the student's undergraduate school, and her undergraduate grades than by ratings of a faculty admissions committee. (Put bluntly, the faculty's deliberation just added noise to the three measure index.) The only documented exceptions to the general conclusion that models out-predict experts are a few kinds of esoteric medical diagnosis.

In these studies, experts routinely violate rational expectations by using observable information inefficiently (worse than simple models do). The violations have two common forms: (1) experts often add error to predictions by using complicated interactions of variables (weighting grades from low-quality schools more heavily, for example), rather than more robust linear combinations of variables; (2) experts pay attention to observable variables that they should ignore because the variables are not highly predictive of outcomes (personal interviews, for example). These psychological tendencies can be traced to some of the judgment biases discussed above (e.g., Camerer and Johnson, 1991).'

References used in this extract:

Camerer, Colin *Individual decision making* **in Kagel, John H. and Roth Alvin E. eds** (1995). *The handbook of experimental economics,* Princeton: University Press

Camerer, C. F. and Johnson E. J. (1991). *The process-performance paradox in expert judgement: How can experts know so much and predict so badly,* in **Ericsson K. A. and Smith** (eds) *Towards a general theory of expertise: Prospects and limits*.

Dawes, R. M. (1971). 'A case study of graduate admissions: Application of three principles of human decision making' *American Psychologist, vol.* 26, pp. 180–8.

Dawes, R. M., Faust D. and Meehl, P. E.(1989). 'Clinical versus actuarial judgement', *Science,* vol. 243, pp. 1668–74.

Meehl, P. E. (1954) *Clinical versus statistical prediction: A theoretical analysis and review of the evidence,* Minneapolis: University Press.