

## *Too Many Students? On the Purpose, Legitimacy and Adequacy of University Entry and Orientation Tests*

Philippe Van Parijs – Université catholique de Louvain, Belgium

### I. INTRODUCTION

For the first time, the Ethical Forum of the University Foundation is revisiting a theme it discussed before. Its fifth edition, held in November 2006, asked the question: “The End of Free Entry? Can university admission tests and *numerus clausus* provisions make higher education more cost-efficient and more socially responsible?”<sup>1</sup> This revisiting does not reflect a lack of imagination. It reflects rather the unprecedented salience of this question in public debate in Dutch-speaking Belgium and the availability of the first results of relevant real-life experiments, in particular the SIMON test at the University of Ghent.<sup>2</sup> This made the present forum more fine-grained and more fact-fed than its predecessor a decade ago, and helped us think afresh about the diversity of tests implemented or imagined and the diversity of goals they might be pursuing.<sup>3</sup>

### II. TYPES

Conceivable tests vary along several logically independent dimensions.<sup>4</sup> First, they can be about *entry* or about *orientation*. In the former case, they aim to indicate whether it is advisable or permissible for a student to enter a particular programme or institution or set of such programmes or institutions. In the latter case, they aim to indicate which, among a set of programmes or institutions, are best suited to the student taking the test. Second, tests can be *compulsory* for a given category of students (for example all those wanting to register at a particular university) or they can be *optional*. Third, they can be *binding* (in the sense that success at the test is required for access) or *non-binding* (simply providing information to the student). An *admission test*, as usually understood, can be characterized as a compulsory and binding entry test. The *positioning test* (*ijktest* in Dutch) now in place in all Flemish engineering faculties can be characterized as an optional non-binding entry test.<sup>5</sup> Though logically possible, it is hardly plausible for an orientation test to be binding (students would be forced to enter the programme or one of the programmes the test shows them to be best for) and for an entry test to be binding without being compulsory (students who did not take the test would be admitted, while those who took it and failed would not), but other combinations are perfectly conceivable.

Fourth, tests can vary according to how demanding they are, i.e. according to how hard it is for student to reach the pass mark. Obviously, this dimension is most relevant for binding entry tests, but its relevance extends beyond them. The probability of false positives (students who do well at the test but badly afterwards) decreases and the probability of false negatives (students who do badly at the test but well afterwards) increases with the

demandingness of the test. Demanding tests aim primarily at identifying (and picking) those quasi certain to succeed later on. Undemanding tests aim primarily at identifying (and excluding) those quasi doomed to fail later on.

Fifth, tests can try to detect the students' capacities and/or their interests. If the focus is their capacities, these can be either raw capacities, 'talents', hardly affected by the quality of schooling, or they can be competences that can be acquired only at (good) schools. The more it is the former that are captured, the more the test can hope to neutralize the impact of socio-economic background. The more it is the latter that are captured, the more the test can hope to determine how well the students are prepared for higher education. If it is rather – or also – the students' interests that are being tracked, these could relate to the content of the studies or to the sort of jobs they lead to. The more it is the former, the better a predictor it is of academic success. The more it is the latter, the better the advice it gives for a fulfilling professional life.

Sixth, there is the timing of the test. The SIMON test is currently organized by the University of Ghent at the end of September. It has to wait until students have registered at the university. It would undoubtedly be better if the test could be organized in June, so that prospective students could ponder the various options open to them in the light of the results. But this would only be possible if the test were organized at community level rather than by each university separately.<sup>6</sup> More radically, Jonathan Hooft (from the student organization VVS) proposed that such a test should be organized at earlier stages of secondary education, so that pupils could gradually find out more both about their own preferences and capacities and about the studies that best match these. Moreover, 21<sup>st</sup> century higher education must stop focusing so exclusively on a learning block at the start of adult life. It must increasingly become lifelong learning. Why not set up a well designed online orientation test that could be taken free of charge by anyone interested in further studies from age 15 to 95?

This is realistic only if the test is of the multiple-choice variety, and hence machine-corrected. Which leads to the seventh and final dimension along which tests could conceivably vary. Obviously, a multiple-choice test does not track the sort of skills that are needed in order to write a well argued essay, or a master thesis. It tracks even less the verbal skills required to make good oral presentations. Designing tests that could track these better is not difficult – think of the interviews of the candidates short-listed by Oxford colleges. But any such test would unavoidably be far more expensive, in terms of examiners' time, than the current multiple-choice tests. The designing of a good multiple-choice test is no simple matter, but – especially if it can be implemented on a large scale – its cost is necessarily very modest compared to written or oral tests that would require individual evaluation by human beings.

### III. GOALS

What kind of test should be used depends on the goals that are being pursued. The most obvious one is to spare a number of students one or more demoralizing

and costly years spent doing things they do not like and/or they are not good at. This goal is not controversial, nor is, therefore, a test suitably designed to achieve it. Such a test can be made compulsory, but the students insisted that it should not be binding: it may have been taken on a bad day, say on a day one did not feel well or could not concentrate because of emotional issues. If the test is not binding, this can be borne in mind when students and their advisers interpret the test and draw their conclusions.

A second objective is to make our educational practice more efficient. One direct way in which tests should help is by getting students more quickly into programmes that are right for them, in terms of both motivation and capacities. Other things being equal, we teachers shall be able to do a better job if our classes are smaller as a result of a reduction in students having to repeat a year, and if they contain more students who are studying what they like best and are best at.

There is, however, a second way in which tests could make our education practice more efficient. This effect can be expected especially, perhaps even exclusively, from tests that are binding and that track capacities that can be acquired only through (good) schooling, rather than innate talents or interests. As outlined by Dirk Van Damme, it operates through the incentives given to both individuals and institutions. If doing well at the test is important, pupils will study harder, and secondary schools will have more motivated pupils. If, moreover, average results at these tests are made public, schools will have a stronger incentive to better prepare their pupils.

In a country that has no centralized final examination scheme for secondary education, this may provide a useful way of revealing the extent and shape of inter-school inequalities in the competences acquired by their pupils, and thereby of pushing and guiding schools towards doing better. Of course, a test usable in this way may meet the same resistance as a centralized examination: the concern of the (Belgian) Catholic network for its autonomy and the distrust of teachers unions were mentioned as major obstacles, with different weights in the north and the south of the country. Moreover, experience in other countries justifies the fear that results will be taken as direct indicators of the quality of a school's job (irrespective of the characteristics of its recruitment) and the associated fear that schools will try to improve their scores by getting rid more easily of the less promising pupils. This fear need not be a decisive obstacle to the introduction of a centralized examination or test, but accompanying measures would need to be taken in order to assuage it.

A third possible goal, suggested by the title of the Forum, is the regulation of the total number of students and graduates. Note that tests need not have the effect of reducing numbers. If well designed, orientation tests may well have the effect of increasing the number of students who complete a programme successfully. And even a compulsory and binding entry test to some programme may have the effect of boosting the number of students who dare to start it because of the confidence derived from success at a test believed to be a good predictor of success. Especially if it is binding or strongly dissuasive, however, an entry test will often have as its goal and effect to decrease the number of students registering for the programme.

There may sometimes be a specific local justification for such a goal. For example, some programmes, like veterinary medicine or physiotherapy, are drowning under applicants from neighbouring countries that have themselves a binding entry test or a *numerus clausus*. Given the legal impossibility of discriminating against EU citizens may leave no option but to introduce an entry test if the cost to the Belgian taxpayer is to be kept under control (assuming that most of these foreign students return to their home countries after graduation). Similarly, whether founded or not, the concern that in the medical domain supply creates demand has justified the imposition of a *numerus clausus* for medical doctors.<sup>7</sup>

The question whether we should reduce the number of students and graduates overall is more contentious. Dirk Van Damme and Jean-Paul Lambert believe that there is no reason for doing so. One argument they used, on the basis of different data sets relying on different criteria of what counts as relevant higher education, is that graduation rates place Belgium in the middle (Van Damme) or at the bottom (Lambert) among OECD countries.<sup>8</sup> This is not a decisive argument, as a statistical average is not a normative ideal. Dirk Van Damme added a comparative survey suggesting that the proportion of workers overqualified for their job is around 16% in Flanders, i.e. lower than in practically all other countries for which comparable data are available.<sup>9</sup> Since the smooth working of a sophisticated labour market can be expected to require some degree of ‘frictional’ over-qualification, there seems to be little ground for believing that there are too many students, that there is overconsumption of higher education, overproduction of graduates.

One consideration that may shake our confidence in this claim is the fact, repeatedly documented, that in Belgium ethnic origin is a more powerful predictor of unemployment among higher education graduates than among people with lower qualifications.<sup>10</sup> If this is the case, it is plausible to conjecture that graduates with a non-EU background, and more generally graduates with a less favourable socio-economic background, will also be overrepresented among those in jobs for which they are overqualified. In stylized fashion, what this suggests is that producing more graduates than the market can absorb has the effect of shifting the social-origin-based inequality of opportunities from educational achievement to job fetching. There is no need to assume any deliberate discrimination for this to happen: those with a more favourable socio-economic background can simply count on more accurate information, on a more valuable network and on the occasional gentle *nudge* to get the internship or job offer that will prove crucial to their professional success.

Is this not too short-termist a reasoning? Is the demand for university graduates not going to increase to such an extent that all those we produce will be absorbed by a labour market that needs the skills their studies equip them with. In the United States, according to the National Center for Educational Statistics, the percentage of people aged 25 to 29 with a bachelor’s degree or more went up from 22.5 to 34% between 1980 and 2014.<sup>11</sup> But according to a report for the Bureau of Labor Statistics, the percentage of job openings requiring a bachelor’s degree or more was 20.0% in 2010, and is expected to increase to 20.5% by 2020.<sup>12</sup> This does not exactly get rid

of the possibility that we may be overproducing graduates and that those left without the jobs they studied for will be overrepresented among those with less advantaged backgrounds. Several empirical assumptions need to hold for this possibility to materialize. They certainly need to be checked and qualified. Yet they may capture enough truth to justify the following warning: a broader access to higher education offers no guarantee of greater social equality in professional opportunities. It may simply postpone the point at which social background produces its causal impact.

#### IV. TWO BIG ISSUES

How well a particular type of test achieves any of the goals listed above depends on many circumstances. These need to be investigated empirically on a case-by-case basis. It would therefore be presumptuous to offer a sweeping conclusion. Instead, I want to broach very briefly two major issues unavoidably raised when discussing entry and orientation tests.

The first, discussed with some passion at the Forum and afterwards, is whether such tests should ideally aim to yield results that show no correlation with socio-economic background. William Guillet (from the student association FEF) showed impressive figures that provide substantial food for thought. In Belgium's French-speaking community, engineering faculties are the only ones that impose an admission exam. It turns out that engineering students have a somewhat higher success rate at the end of the first year (46.5 instead of 40%) but that they are on average from a significantly more advantaged background: 63.8% have at least one parent with a university degree (compared to an average of 43.9%) and 19% are entitled to a student grant from the community (compared to an average of 25%).<sup>13</sup> Now, the causal ascription of such differences to the test needs to be made with some caution. These data do not say anything about whether the correlation between admission and success in engineering faculties on the one hand and socio-economic background (parents' education, entitlement to grant) on the other would have been lower – or higher – in the absence of the test. And it is this difference in correlations, not the sheer existence of a correlation, that needs establishing for the causal claim to be substantiated. The question I want to raise is deeper, however: in an ideal world – in which, in particular, inequality between the pupils' schools would not replicate inequality in their parents' wealth – should we hope for a complete neutralization of the impact of socio-economic status on the results of entry and orientation tests, whether in engineering or in any other subject?

Wouter Duyck reported with some satisfaction that so far the SIMON test he helped develop at the University of Ghent yielded success rates at the test that are only slightly sensitive to socio-economic background. As he acknowledged, however, this relative insensitivity is no doubt largely due to sample bias: those taking the tests are students registered at the University of Ghent, not a representative sample of 18-year-olds from East Flanders (where Ghent is located). Because those with low SES scores

who take the risk of going to university are most likely to display other (unmeasured) favourable characteristics, the SIMON test can hope to combine high ‘social neutrality’ and high predictive power. However, Wouter Duyck also added that he would not be particularly disturbed if his data did show correlations between performance at the SIMON test and socio-economic status.

This should be obvious if the sample subjected to the test were drawn from the whole adult population. In a society in which educational achievement is meant to prompt material rewards (themselves reflected in higher SES scores), performance at a test that aims to be a good predictor of educational success must be expected to correlate strongly with the SES score of the subjects themselves later in their lives. The next step is trickier. Should we also be pleased when observing a correlation between performance at the test and the SES scores of the subjects’ parents, typically their mother’s educational level? Wouter Duyck argues that we should, on the ground that intelligence is, to a significant extent, genetically determined. In an ideal world, therefore, when whatever is objectionable about the impact of material condition on educational success has been neutralized, one should still expect performance at SIMON-type tests to be correlated with the parents’ SES scores, at least when entire cohorts are considered (not only university applicants) and especially when their education level is taken as the relevant index; less strongly correlated than in the real world, no doubt, but still significantly. Unsurprisingly, this conclusion did not meet with unanimous agreement. Behind possible factual disagreements about the measurement and inheritability of IQ lurks a fundamental ethical question: is a null correlation between social background and academic success a necessary condition and is it a sufficient condition for social justice in higher education? My own answer is: neither.

The second big issue was raised by Jeroen Huisman when he emphasized that higher education is not only meant to prepare people for jobs, but also to provide ‘the university experience’. Along the same lines, both student representatives mentioned the critical mind, the spirit of questioning as inherent in the student condition. I have great sympathy for this view. I just gave an address at a graduation ceremony at the Université St Louis. Although I took the greatest interest in my own studies there – so much so that I ended up with a degree from each of its three Faculties – I confessed to the new graduates that the most memorable element of my St Louis experience was when I ended up in a police cell as a result of organizing an unauthorized demonstration. Providing a university experience of some sort, and one that is closer to unauthorized demonstrations than to binge drinking, is and must remain an important role of higher education.

To help it play this role, I doubt that entry and orientation tests can be of much use. If properly designed, however, they can definitely be of great use to help some students waste less time. They can probably also help make our education system more efficient and help regulate upward or downward the number of students and graduates so as to better match the demands of the labour market. But the university’s function is also to help produce the critical and imaginative minds our democracies need. This does not justify funding unlimited years of self-indulgence for spoilt youth disproportionately

stemming from well-off families. But it does justify a fair level of slack. The best university system, all things considered, will not be one with a perfectly tuned sorting gate at the entrance, one that allocates each student to the single track that fits best his or her talents and tastes consistent with the economy's needs, one that processes as fast and efficiently as possible the inputs it collects from secondary schools into competent and docile factors of sophisticated production. The best university system will need to preserve space for the university experience. And this arguably requires that one should allow many students to follow many tracks that the best orientation tests of the world would have advised them not to follow.

## NOTES

1. See [www.universitairrestichting.be/en/content/fifth-forum](http://www.universitairrestichting.be/en/content/fifth-forum), and the published synthesis : Philippe Van Parijs, "The End of Free Entry? Some Synthetic Considerations and Personal Reflections," *Ethical Perspectives* (14/2, 2007: 193-203).

2. SIMON (*Studievaardigheden en interessesmonitor*, [www.simon.ugent.be/project](http://www.simon.ugent.be/project)) was launched at the University of Ghent in 2011-12. It now has a data base with 4,500 student profiles. From 2015-16, it is also being used at the VUB and some *hogescholen* (university colleges). LUCI (*Levens Universitair Competentie-Instrument*, [www.kuleuven.be/luci](http://www.kuleuven.be/luci)) was developed more recently at the KU Leuven and made available in March 2015 to all its students.

3. I am very grateful to Jacques Willems, Eric De Keuleneer, Hilde Garmyn and the staff of the University Foundation, who hosted the event with their usual kindness and efficiency, to Patrick Loobuyck (UA), who master-minded the preparation of the event, chaired it and introduced it, and to all speakers – Wouter Duyck (UGent), Jeroen Huisman (UGent) and Dirk Van Damme (OECD); William Guillet (FEF), Jonathan Hooft (VVS), Jean-Paul Lambert (USt Louis), and Joos Vandewalle (KU Leuven) – for many fruitful insights. In addition, I am most grateful to Wouter Duyck, Jean-Paul Lambert and Joos Vandewalle for extensive subsequent feedback on a first draft of this personal synthesis, and to an anonymous referee for many insightful remarks and useful suggestions. I tried to accommodate these as much as is appropriate for a piece that does not aim to provide a novel contribution to the specialized literature, but only to share some of what could be learned from a particularly fruitful interdisciplinary exchange.

4. My purpose is here to explore the space of logical possibilities by way of concrete background for the subsequent discussion of goals. A recent overview of the schemes actually in place in the European Union is provided in *Higher Education, Entrance Qualifications and Exams in Europe: A Comparison*, European Parliament: Directorate-General for Internal Policies, 2014.

5. This test qualifies as an admission test rather than an orientation test because it is targeted at students choosing a particular programme. It was introduced in 2013 in the Dutch-speaking community, ten years after the abolition of the binding engineering admission test that dated back to the 19<sup>th</sup> century. It consists of 30 multiple-choice questions to be answered in four hours without a calculator (see [www.ijkingstoets.be](http://www.ijkingstoets.be)). According to Joos Vandewalle, its predictive value for study success turns out to be very good, with a rate of participation of up to 85%.

6. Collaboration is on its way between Flemish universities. The Flemish government aims at having a non-binding test in place for all Faculties by 2018-2019.

7. For an in-depth discussion of the putative over – or under – supply of graduates from the various sectors of Belgium’s higher education, see Jean-Paul Lambert, “Choix des orientations d’études et besoins de la société,” *Reflets et perspectives de la vie économique*, 2015.

8. Dirk Van Damme used data for rates of first-time entry into higher education (2013) and for rates of graduate output (2012). Jean-Paul Lambert used data for rates of entry into tertiary education of type A (at least 4 years). Both pointed out that rates of increase have been low in recent years, compared to most other countries.

9. The rate of over-qualification is defined as the percentage of workers whose highest qualification is higher than the qualification they deem necessary to get their job today. Regrettably, these data (drawn from the *OECD Skills Outlook 2013*), like many other OECD data, are available, as regards Belgium, only for the Dutch-speaking community. Convergent data for Belgium as a whole can be found in *Skills mismatch in Europe : Statistics Brief*, International Labour Office, September 2014. Thanks to Jean-Paul Lambert for this useful additional reference.

10. See, for example, Albert Martens and Nouria Ouali. *Etnische discriminatie op de arbeidsmarkt in het Brussels Hoofdstedelijk Gewest*. KU Leuven & ULB, January 2005; Nouria Ouali and Pasquale Cennicola. *Etude sur la discrimination en Belgique et en Wallonie*, Namur: IWEPS, January 2013, 37-38; Stijn Baert and Bart Cockx. “Pure Ethnic Gaps in Educational Attainment and School to Work Transitions. When Do They Arise?” *Economics of Education Review* 36 (2013, 276-294).

11. See Digest of Education Statistics, Table 104.20 ([https://nces.ed.gov/programs/digest/d14/tables/dt14\\_104.20.asp](https://nces.ed.gov/programs/digest/d14/tables/dt14_104.20.asp)).

12. C. Brett Lockard and Michael Wolf, “Occupational Employment Projections to 2020,” *Monthly Labor Review* (January 2012, 84-108; 106, Table 6).

13. The data are drawn from Catherine Vermandele *et al.* “Profil des étudiants entamant des études universitaires et analyse des choix d’études,” *Les cahiers de recherche en éducation et formation*, UCL: GIRSEF, 10/78 (2010).