

Who is missing from higher education?

Stephen Gorard
School of Education
University of Birmingham
B15 2TT
s.gorard@bham.ac.uk

Abstract

This paper discusses the difficulties of establishing a clear count of UK higher education students in terms of the categories used for widening participation, such as occupational background or ethnicity. Using some of the best and most complete data available, such as the annual figures from the Higher Education Statistics Agency, the paper then establishes that there is little evidence of a simple consistent pattern of under-representation within these categories, except perhaps for men, and students of white ethnicity. However, once prior qualifications are taken into account, there is no evidence that potential students are unfairly and disproportionately denied access to HE in terms of occupation, ethnicity, sex or disability. This has important implications for what we mean by widening participation in HE, and how we might achieve it.

Who is missing from higher education?

Introduction

UK Government policy is to increase rates of participation in, and qualification from, higher education (HE), focusing particularly on those aged between 18 and 30. The pressure to increase participation is intended to be directed primarily at those groups previously under-represented in comparison to their population share, especially students from low-income families. This is because the current student body in the UK appears to be stratified in terms of class, ethnicity, and location. At time of writing the government had spent at least £2 billion on widening participation (WP) activities since 1997, and despite this there has been a reported fall in the percentage of young entrants from lower social classes (Sanders 2006). This shows how difficult the task facing WP is.

The widening participation agenda is predicated on the notion that particular social groups, defined perhaps by social class or ethnic background, are unfairly under-represented in higher education. Taking post-compulsory education and training as a totality, there are, in theory, opportunities of some sort available to the entire adult population. These include library drop-in centres, free basic-skills provision, job-seeker training, liberal evening classes, and courses delivered entirely by technologies such as television or computer. Therefore, the continued under-representation of certain groups in these objectively open episodes suggests a pervading problem (Selwyn et al. 2006).

However, unlike the patterns in lifelong learning more generally, it is not clear that this unfair under-representation in HE has been established, and for a very simple reason. HE has not previously been intended to be available to all and is, to a large extent, based on selective entry in a way in which other lifelong educational opportunities are not. The majority of first places at HE institutions in England are allocated on the basis of applicants' prior qualifications. Where these prior qualifications are distributed as unevenly as the opportunities for HE then this both explains the patterns of participation in HE, and also suggests that using prior

qualifications in this way may be unfair. We should not, therefore, consider widening HE participation without a more detailed consideration of who the additional participants are intended to be. Put another way, who is missing from higher education?

We cannot assess the claims of under-representation in HE using figures from HE alone. We need also to track changes in the social class of the population from which HE entrants come, and changes in the distribution of entry qualifications by social class in that population. These figures then have to be combined in appropriate proportions. And even this takes no account of the 'inflation' taking place in class categories, due partly to the feminisation of the workforce. Non-manual occupations have grown in past decades, while both skilled and non-skilled manual occupations have declined, changing the meaning and relative privilege of non-manual occupations. So, for example, an observation that the proportion of students from non-manual backgrounds has remained the same over a number of years could actually be construed as evidence of wider participation in HE.

In order to establish that access to higher education is unfair, we would need to demonstrate that particular social groups are seriously under-represented in universities, and that this under-representation has no reasonable or merited explanation. In a sense, this sounds easy to demonstrate, but it is actually dependent on a sequence of less than perfect analytical steps (Gorard and Smith 2006). These steps include having:

- a suitable definition of, and method of measuring, membership of the social groups involved;
- a suitable definition and characterisation of the relevant population;
- an accurate measure of the prevalence of the social groups in the relevant population;
- an agreed definition of what we mean by participation in HE;
- and an accurate measure of the prevalence of those with higher education experience in the social groups involved.

From the results of these five steps, we can then calculate the difference between the proportion of each social group in the relevant population and the proportion of the same group in HE. If this difference is large and important then we can assume that there is a problem requiring explanation and amelioration. This should form the background to the HE widening participation agenda in the UK. The paper briefly discusses each of these analytical steps, and mentions some of the key decisions and compromises that need to be made even in such an apparently simple calculation. It continues with a presentation of some official data on HE participation that takes these compromises into account.

Defining social categories

To establish that there should be more of a particular social group in HE than there is we would first need to be able to define the group clearly, in such a way that the definition could be used by different people in different places and at different times to mean the same thing. Unfortunately, the categorisation of social groups by occupational class or ethnicity is a matter of judgement over which even experts disagree (Lambert 2002, Lee 2003). The categories themselves are arbitrary, and they interact importantly with each other and with other categories such as sex (Gorard 2003). A further key problem in examining trends in social categories over time is that the variables collected, or the coding used for the same variables, also change over time. Consequently it is often difficult to make genuine and straightforward comparisons over time or between groups.

Significantly for the measurement of WP, it is not clear whether any classification by ethnicity or occupation should be of the potential student or of their parents. For example, it seems absurd to try and base the occupational classification of a student on their own work history when they may never have been anything other than a full-time student in education. But where the occupations of the two parents differ, which is to be preferred? If one or more of the parents has not lived with the student, does this make a difference? It is no less absurd to base the occupational classification of a student aged 45 on the previous occupation of their parents. If, on the other hand, we use two different classification systems for younger and older students, at what age

should the cut-off point be? Should the cut-off be based on age alone or on work experience? Can we reasonably aggregate the classifications based on the two different systems?

As illustrated repeatedly in the first part of this paper, there is no simple answer to such analytical questions. Yet every analysis covering patterns of participation must make, even by default, a bewildering number of decisions just like this, and every analyst might quite reasonably make a different set of decisions. Unless these analytical compromises are clearly reported, there is a danger that debates about what is happening in widening participation will be misinterpreted by commentators as being about issues of substance, whereas they are, in reality, merely about differences in making these analytical decisions.

Defining the relevant population

The next step in establishing that there should be more of any particular social group in HE requires us to assess the prevalence of that social group in the relevant population of those who could be participants. Unfortunately, when researching episodes of post-compulsory learning, it is not clear what this relevant population is. An analyst using figures for all adults is open to the charge that the inclusion of people over the age of 50, for example, is irrelevant since so few of these are currently participating in HE even though they represent a large proportion of the population. Another analyst using only young adults, however, is open to the charge of presuming that WP is only about traditional-age students, and so excluding from the analysis precisely those to whom access could and should be widened. It is not even clear what is the youngest age that should be considered in the population of potential HE students. Some HE institutions admit students at age 16, or even younger on rare occasions. This decision about age is crucial to our results, however, because the characteristics of the birth cohorts in the UK have changed over time in terms of the relative prevalence of ethnic and occupational groups. Using population figures for all ages, for example, may lead one analyst to conclude that working-class students are under-represented while an ethnic minority is over-represented in HE. Another analyst, using the same figures for HE but using population figures only for those

aged 17-21, may conclude the reverse. The opposite conclusion can be drawn from the same HE data by different analysts, because the proportion of working-class families may have been decreasing over time in the UK, while the proportion of an ethnic minority group may have been increasing.

A similarly key decision for an analysis of participation in one country concerns the original domicile of the potential students. In an analysis for England, is only the population of England relevant, or should the analysis include the potential students, and so the population, of the other home countries of the UK - Scotland, Wales and Northern Ireland. Excluding them complicates the analysis because it is then essential to distinguish England-domiciled students from others. But as with the definition of social categories (see above), this leads to issues about whose domicile counts. Is it the residence of parents, of the potential students themselves, or some combination of these? Students from within the European Union are treated, and charged fees, as home students. Does this mean that the populations of all EU countries must also be in the analysis? This leads to several analytic problems – most notably that the population of countries like the Czech Republic is very different to that of the UK. So, even if comparable official population figures from the Czech Republic exist, and this is very unlikely, it means that the population figures in our analysis will be considerably affected by those of a country which, in reality, provides an almost negligible proportion of students for HE in the UK.

Measuring the characteristics of the population

If an analytical decision has been made to exclude overseas applicants, then the population census of the UK provides the most complete coverage to help assess the characteristics of the relevant population. But this census only happens every ten years, making it dated, and some of the most relevant questions for this analysis are only asked of a sub-sample of 10% of the cases, or for the economically active head of the household, rather than for individuals.

Despite it being a legal requirement, not every household actually takes part, not everyone is in a household, and not everyone who takes part responds to the class and

ethnicity questions even when asked them. The categories used for the class and ethnicity questions are not the same as those used in other large data sets – such as the individualised student records (ISRs) held by the Higher Education Statistics Agency (HESA) for all students, the Universities and Colleges Admissions Service (UCAS) database of applicants, or the annual schools census. This makes it difficult to use the population figures as the denominator in the final step of this analysis. For example, the annual schools census does not ask for parental occupation, and the most commonly-used indicator of disadvantage that it provides instead is eligibility for free-school meals which is not recorded for applicants to HE. The UCAS figures for entry qualification to HE exclude the majority of each age cohort who do not apply to HE. Therefore, it is not possible to compare directly the qualifications attained at school by different social classes with the rates of participation in HE. We can only estimate the relevant figures from sample surveys, often with high non-response or longitudinal dropout, and sometimes with incompatible measures of class or qualification. Coupled with the many cases in HE not classified by occupation, the situation for analysis is highly unsatisfactory. Yet, it must be stressed that this is the best kind of evidence available for studies of widening participation.

Defining participation in HE

For the fourth step in our apparently simple calculation we would next need to know the prevalence of the social group we are concerned with that had participated in HE. This step also faces problems in the form of yet more crucial analytical decisions that could swing the results of the analysis either way. We need to know what proportion of the population has already participated in HE (even if they did not receive a qualification). We need to decide whether to include NQF level 4 (HE) courses in FE colleges, level 3 (pre-HE) courses in HE institutions, postgraduate or level 5 students, and those involved in professional training such as postgraduate teachers and social workers, or those involved in short courses such as HE-based continuing professional development. Do we include those taking degrees by correspondence, or via the internet? We need to know (as above) whether we distinguish between home country, UK, Commonwealth, and EU home students. If not, then our prior population figures become more problematic. If so, then some datasets make it difficult to distinguish

between categories of home students. Variation in these decisions over time, or between analysts would be perfectly proper, yet it makes comparisons between their results difficult. As with the general population figures, there will be incompleteness in HE records, and for some years of the data the ‘Individualised’ Student Records are not actually linked to individuals but to courses, so that a part-time student taking two courses in two different institutions does not have a unique identifier, and is in danger of being counted twice (Gorard and Taylor 2001).

Measuring the characteristics of those in HE

The final requirement, before being able to make the relatively simple arithmetic calculation involved in producing the proportionate representation of social groups in HE, is in some ways the easiest since it concerns only those in HE. However, it is worth illustrating some of the difficulties in using the data even for this group to help readers understand the severe limitations of any analysis of patterns of participation. There are no ideal datasets for the analysis of patterns of participation in higher education (HE) in terms of policy changes, or social, economic, or regional disparities. All existing datasets suffer from one or more defects: they include only participants, have incomplete coverage, have substantial proportions of missing data or cases, or are incompatible in range or aggregation with other datasets.

As with the population census, there are cases simply missing from official statistics on participation in HE, and as with the population census we cannot be entirely sure how many cases are missing. The UCAS data on applicants to HE has historically seriously under-represented part-time, mature and distance students. Returns from each university of the number of students actually in place may give a better indication of the overall figures but are generally deficient in terms of key background variables such as ethnicity and occupational class.

A common problem for the relevant large-scale datasets lies in data missing even from the cases that are known about. For example, many of the variables in the HESA datasets are compulsory – i.e. some value has to be reported for each student. But this does not mean that complete data are available for every student. The ‘missing’ data,

which can include not known, information refused, information not yet sought, and 'other' non-completed, often covers a large proportion of the students. One example is that other than 'white', 'missing' is officially the largest ethnic group among students in England. In fact, the unknown cases considerably outnumber all of the minority ethnic groups combined. Some of the minority ethnic groups are quite small, leading to the usual volatility of small numbers when analysing trends over time or differences between groups. Consequently, the high proportion of missing cases in any analysis using this variable could significantly bias the results being presented, even where the overall response rate is high. This means that any differences over time and place, or between social groups, needs to be robust enough to overcome this bias (among many others). The scale of a difference or change must be such that it dwarfs the bias introduced by measurement errors, missing cases, and changes in data collection methods over time. This difficulty is seldom acknowledged by commentators.

Similarly, UCAS applicant figures, and HESA Individualised Student Records (ISRs), have a large proportion of cases with no occupational category. In fact, when non-responses are added to those cases otherwise unclassifiable by occupation (through being economically inactive, for example) then having no occupational category becomes the single largest classification. In 2002/2003, around 45% of first year undergraduates were unclassifiable in terms of their occupational background, according to HESA figures. How then, could we possibly know whether any occupational group was under-represented in HE? Any difference between groups in HE and in the population is simply dwarfed by the missing data. Thus, if students from less-elevated occupational groups were less likely to respond to questions about occupational background, so that a high proportion of the missing 45% of cases were really from less-elevated occupational groups, then these groups might actually be over-represented in HE, even though they may be under-represented among those who answer the occupational question. We just do not know.

Analysis

The analysis required to produce patterns of participation in HE by social groups is relatively simple once the preceding analytic decisions and necessary compromises have been made. For example, the proportion of the social group under consideration (e.g. an ethnic minority) can be compared in HE and in the relevant population by dividing the former by the latter. A result of one (1) shows proportionate participation, and a figure less than one (<1) shows under-representation. The difficulties, possibilities, and consequences of this kind of analysis are illustrated in the next section of the paper using official figures, which represent the best evidence we have. There is not some other, far superior, dataset on which the WP agenda is based. In the light of the foregoing, the figures for the HE participation of particular social groups are presented to help try and decide which students are ‘missing’, and so where the efforts of WP activities might be directed. Some of the figures are new, some come from a HEFCE-funded review of evidence (Gorard et al. 2006), and some from other work, such as that for the Rees Reviews of student hardship in Wales (Gorard and Taylor 2001, Taylor and Gorard 2005).

Patterns of participation

There has been a considerable increase in the number of home undergraduate students of all ages in the UK over the past decade, with an overall growth of around 50% (Table 1). Much of this increase has been in study for qualifications below degree level (according to HESA), including foundation degrees, diplomas, and professional certificates. This distinction is important, because it shows that increasing participation, and the widening of opportunities that accompanies it, has been disproportionately concerned with many of these relatively recent kinds of opportunities. However, the standard economic analyses presenting arguments for expanding and participating in higher education are usually based on more traditional undergraduate degree courses. The arguments are based on the notion of HE as an investment in human capital, both by the individual participant whose lifetime earnings will rise, and by a society whose workforce will be more skilled and competitive in a global economy. The econometric calculations presented as evidence that such investment pays off are necessarily based on participation some time in the past – sufficiently so for the investment to have had a chance to pay off. So, if the

recent growth in HE has been of a different kind – not full-time undergraduate study – then many of the economic arguments for HE participation, based on the traditional first-degree model, are not applicable. Expansion of HE has expanded the system in such a way that it has changed, inevitably, and we no longer have much solid evidence on the consequences for individuals or society.

Table 1 – Number of home students, UK, 1994/95-2004/05

	1994/95	1999/00	2004/05
All undergraduates	451840	525140	673775
Full-time first degree	273586	281780	320865
Part-time first degree	49425	31400	58285

Source: HESA

The ethnic background of students is one area of concern for widening participation, and this paper has already rehearsed some of the difficulties in deciding whether ethnic minority groups are under-represented in HE. Table 2 shows that there has long been a high proportion of students whose ethnic origin is unknown. This proportion has declined from 20% to just under 10% over a decade. However, this still leaves analysts with at least four major problems. First, the decrease in missing data has been disproportionately among the traditional full-time undergraduate body (now only 3% missing), whereas the increase in actual numbers of students is among part-time students and others (still 13% missing). Second, the decrease in missing data itself makes it nearly impossible for analysts to decide whether apparent changes over time are due to widening participation or merely due to changes in form-filling. Third, the ethnic classifications used by HESA have changed over time, and these changes, including the introduction of mixed ethnic categories, mean that the classifications from different years do not nest and can not be converted into each other. So, it is not actually possible to track changes in ethnic participation over time. Fourth, the number of cases in each of the rapidly growing number of ethnic categories is generally small and shrinking as the categories increase. As one may imagine, the number of students self-reported as being of mixed Chinese and Black other origin,

for example, is currently very small in the UK. All of these factors together mean that it is not possible to differentiate robust patterns for specific minority backgrounds from the ‘noise’ generated.

Table 2 – Percentage of students with known ethnicity, UK, 1994/95-2004/05

	1994/95	1999/00	2004/05
All undergraduates	80	90	91
Full-time first degree	83	94	97
Part-time first degree	80	87	87

Source: HESA

The proportion of students with known ethnicity who are non-white has increased slightly over a decade in the UK (Table 3). This decline is approximately in line with, but ahead of, an overall increase in reported ethnic minorities in the population as a whole. The population as a whole is 92% white, according to the 2001 census. Therefore, one could argue that the HE system somewhat over-represents the minority groups, if we can assume that all ethnic groups are equally likely to answer the question about ethnic backgrounds. But the third largest ethnic group in the UK, after white and not known, is Indian and this group is one of those obtaining higher qualifications at NQF levels 2 and 3 than white students, making them differentially eligible for HE acceptance. Nor does this simple comparison between population and HE take into account ethnic differences between age cohorts, or potential inequalities between subjects, institutions, regions, and specific backgrounds. We can say, though, that existing figures give us no reason to assume that ethnic minorities, in general, are under-represented in the HE systems of England or Wales. As far as WP is concerned, we do not, and probably could not, have the robust evidence needed to highlight any under-representation of specific ethnic minority groups (each mostly representing only a fraction of 1% of the population). However, we would have a reasonably strong case in arguing that the majority white group is the most obviously under-represented at present.

Table 3 – Percentage of white students, UK, 1994/95-2004/05

	1994/95	1999/00	2004/05
All undergraduates	87	87	84
Full-time first degree	87	84	81
Part-time first degree	90	88	87

Source: HESA

Another area of concern for WP is the participation of students with some form of disability. In the UK, there has been an increase in the proportion of HE students reporting a disability, with the proportion almost doubling over a decade even while the numbers of students overall was growing (Table 4). It is not immediately clear whether this increase in students with a reported disability is evidence of a widening of opportunities, or more to do with an increase in reporting.

Table 4 – Percentage of students with disability, UK, 1994/95-2004/05

	1994/95	1999/00	2004/05
All undergraduates	3	4	6
Full-time first degree	4	5	7
Part-time first degree	3	3	5

Source: HESA

It is clear that the major part of this increase has been for students with a non-visible disability such as dyslexia (Table 5). In fact, if the figures for dyslexia alone are subtracted from the figures for disability in Table 4, there has been no overall growth in the proportion of other disabled students since 1994/95. The numbers involved here are smaller than for many specific ethnic minority groups, making any claim to under-

representation even more difficult. Disability is not covered by the population census, and other estimates from National Statistics (e.g. <http://www.statistics.gov.uk/cci/nugget.asp?id=795>) present sample data in which serious and mild disabilities are aggregated with long-term illnesses like asthma. Between 15% and 20% of the child population have reported a disability or long-term illness in repeated surveys over the past decade. Of these, the majority (over 40%) suffer from asthma, and others suffer from other illnesses apart from a disability. This means the proportion of children with a reported disability is very similar to the proportion of HE students with a disability (6%-7%). The possible under-representation of the less than 1% of the population with reported ‘serious disabilities’ is too difficult to determine for the same reason as for specific minority ethnic groups.

Table 5 – Percentage of students with dyslexia, UK, 1994/95-2004/05

	1994/95	1999/00	2004/05
All undergraduates	0	1	3
Full-time first degree	1	2	4
Part-time first degree	0	1	1

Source: HESA

One of the simplest classifications, and hence the variable with least missing data of those covered in this paper, is the sex of the students. The figures from HE and from the population census strongly suggest that males are under-represented in HE to an extent that is not true for the available evidence on ethnic minorities, disabled students, and occupational groups (see below). Since 1994/95, the proportion of female first degree students in HE has grown considerably, especially among part-time students in the UK (Table 6). A small part of the explanation may lie in demographics – there have until recently simply been more women than men in the relevant age groups of the population. Part of the explanation lies in the purported underachievement of boys at school (Gorard et al. 2001), and part may be due to favourable attitudes towards continuing education among young women, especially at

age 16. This participation gap is clear (there are fewer missing cases in answer to questions about sex/gender than occupation or ethnicity), robust in the sense that it appears annually, and apparently growing. Yet, ironically, this is one area where WP is not particularly active at time of writing, reflecting a lack of policy concern with the under-representation of males.

Table 6 – Percentage of female students, UK, 1994/95-2004/05

	1994/95	1999/00	2004/05
Full-time first degree	49	53	53
Part-time first degree	53	65	65

Source: HESA

Perhaps the most important target of widening participation activity has been tackling the apparent under-representation of less advantaged socio-economic groups. Table 7 presents a historical breakdown of the student body in the UK by social class (Registrar General's previous scale, Gorard 2003). It shows that students come from predominantly professional and intermediate backgrounds (I/II), with few from part-skilled and unskilled backgrounds (IV/V). This pattern changes very little over the time period shown. The most consistent change has been in the growth of those students of unknown occupational class. It is important to note that occupational groups are not evenly divided in the population, and we would expect there to be many more individuals in HE from class II than from class IV, for example. And this is what we find. The dominance of certain social groups in HE is partly a function of their numerical frequency in the population which changes over historical time, to an extent that is not always made clear in media and policy reports. If the population is becoming more middle-class over time, for example, then we would quite rightly expect students at HE (who tend to be younger than the population as a whole) to be more middle-class than the resident population. In itself, this would not be unfair or even disproportionate in relation to the correct figures for the appropriate age-related population (which we may not have).

Table 7 – Students accepted for home degree in UK, 1994/95 to 1998/99

Class	1994	1995	1996	1997	1998
I Professional	16	15	15	14	14
II Intermediate	41	40	40	39	40
IIIN Skilled non-manual	12	11	12	12	12
IIIM Skilled manual	16	16	15	14	15
IV Partly skilled	7	7	7	8	8
V Unskilled manual	2	2	2	2	2
Not known	8	9	10	11	11

Source: National Statistics (2001)

However, if we compare the proportion of students and the proportion of the population in each social class there seems to be an anomaly. For example, whereas only around 25% of the population is in group II, around 40% of HE student acceptances are in this group. And whereas around 4% of the population is in group V, only around 2% of HE student acceptances are in this group (<http://www.statistics.gov.uk/STATBASE/xsdataset.asp?More=Y&vlnk=313&All=Y&B2.x=24&B2.y=8>). Thus, the key indicator is not the breakdown of the student body into social classes, but the rate of HE participation in each social class. But here we reach the problem of age again, because most students are young and the proportion of each social class in the population is changing over time. This problem is traditionally overcome, to some extent, by using only the figures for traditional-age students, for whom there is superior data.

There has been a considerable growth in overall HE participation, from 12% of the traditional age cohort in 1980 to 35% in 2001. More importantly, however, Table 8 shows that this growth has been disproportionately among social classes IIIM to V (the largely under-represented groups), rather than I to IIIN (the over-represented groups). In 1940 an individual in one of the over-represented social classes was four times as likely to go to HE as one from the under-represented social classes. As recently as 1990 the odds remained at nearly 4:1. But by 2001 an individual from the over-represented social classes was 2.6 times as likely to go to HE – still a considerable difference but at least an improvement. Looked at another way, those in

over-represented social classes are now over six times as likely to continue to HE as they were in 1940 whereas the figure for those in under-represented social classes is over nine times. Since 1990, classes I to IIN have improved their position by a factor of 1.35 (or 50/37) whereas classes IIIM to V have improved theirs by a factor of 1.9 (19/10). Whatever the problem in the current situation is, it is better than it was – for young students and insofar as existing figures allow us to judge.

Table 8 – Age participation index by collapsed social class in UK, 1940-2001

Year	I/II/IIN	IIIM/IV/V	Overall
1940	8	2	-
1950	19	3	-
1960	27	4	-
1970	32	5	-
1980	33	7	12
1985	35	8	14
1990	37	10	19
1991	35	11	23
1992	40	14	28
1993	43	16	30
1994	46	17	32
1995	47	17	32
1996	48	18	33
1997	48	18	34
1998	45	17	32
1999	45	17	32
2000	48	18	33
2001	50	19	35

Note: The API is the number of home-domiciled young (aged less than 21) initial entrants to full-time and sandwich undergraduate courses of higher education expressed as a proportion of the averaged 18 to 19 year old UK population.

Other commentators and analysts agree on this improvement and its scale (e.g. Mayhew et al. 2004). Using a different dataset and approach Raffe et al. (2006) show

a long-term decline from 1986 to 1999 in the odds ratios for managerial/professional class compared to working class students, in terms of attainment levels at age 16, staying on in formal education at age 16, attainment at 18, and subsequent participation in HE. For these two groups in England, the odds ratios of taking a degree fell from 7.9 in 1990, to 5.7 in 1996, and 4.4 in 1999. The near consensus among these analysts confirms that the DfES (2003), who claim that the situation is actually getting worse in terms of social class participation, are simply wrong. In fact, HE participation has been slowly widening for decades.

The fastest growth in widening participation between the two collapsed social class groups in Table 8 occurred in a short period of the early 1990s, reflecting perhaps little more than a rapid increase in the number of places available at HE institutions. However, there has also been a more general increase since 1940 as part of a relatively long-term historical and social trend. This trend is largely undisturbed by specific policies to widen participation. This suggests that the quickest and easiest, and once everything is taken into account perhaps even the cheapest, way to widen participation is simply to increase the number of funded places at HEIs. The HEIs will generally find the students if there are places to be filled, and the historical evidence is that these new students will be disproportionately from the less represented social groups. Interestingly, the more recent initiatives intended to widen participation are associated with the only period in this historical record during which participation has not widened.

Out of interest, Table 9 shows a rather unusual way of comparing the socio-economic composition of HE students and the population census, by including the not known responses as a category rather than eliminating them from the analysis as is normal. The figures represent Welsh-domiciled students participating in HE in 2002/03. The socio-economic classifications derived from UCAS admission data ignore the 38% of all Welsh-domiciled students who did not enter HE through UCAS, and a further 28% of UCAS-entered students were not classifiable or did not respond to the occupational question. This means that socio-economic classifications are only available for 34% of students, and even these will contain errors and areas of subjective judgement (see above). Similarly, the population census only has responses for 71% of complete cases. Presenting the table in this way, making the missing cases visible in both

columns, is unorthodox and somewhat inconvenient for readers. But it should be effective in alerting readers to the scale of the problem. There is not some alternative higher quality or more complete dataset on which the widening participation agenda is based. The data presented here are the best available for these students in this year.

Table 9 – Percentage of Welsh-domiciled students participating in HE by socio-economic classification, 2002/03 compared to population census 2001

	HE 2002/03	Population census 2001
Higher managerial and professional occupations	6	6
Lower managerial and professional occupations	11	16
Intermediate occupations	5	8
Small employers and own account workers	3	7
Lower supervisory and technical occupations	2	8
Semi-routine occupations	4	12
Routine occupations	2	10
Never worked and long-term unemployed	0	4
Not classified or not valid	67	29

Source: HESA

What these figures show above all is that no-one really knows the socio-economic make-up of the population or of HE students in enough detail to make clear claims about the proportionate differences between small groups or about relatively minor changes over time. All socio-economic groups appear to be under-represented to some extent – which clearly cannot be so. The key question is whether there is any bias in the non-responses. If, for example, the non-responses are more common among the prestigious occupational groups (as evidence from surveys would generally suggest) then this could explain the apparent under-representation of these groups in HE. We just do not know. There is some indication that both higher- and lower-managerial

and professional occupational backgrounds are over-represented amongst Welsh-domiciled HE students, and a similar suggestion that working-class occupations, including the long-term unemployed, routine and semi-routine occupations and lower supervisory and technical occupations are under-represented. The picture for only those students aged 18-30 is similar (see Taylor and Gorard 2005). However, it must also be considered, given that a majority of the data is missing for students, that the difference may lie not only in the student population but also in the kind of respondents to the question about occupational background. If, for example, higher managerial and professional occupations were proportionately represented in HE but more likely to respond to the occupational question, the result would be indistinguishable from what we see here. As with the problem of historical changes in the class structure of the relevant population, we simply do not know about the impact of differential response rates. This means, of course, that we do not really know whether and to what extent different social classes are under-represented in HE.

The relevance of prior qualifications

The overwhelming majority of applicants to university are accepted on the basis of their prior qualifications (around 95% according to UCAS 1999), and two-thirds are accepted on the basis of A/AS levels alone. These prior qualifications are strongly associated with social class and, to a lesser extent, with ethnicity, disability and sex. According to the Youth Cohort Study (YCS), 51% of social classes I/II in England obtained the equivalent of two A-levels at age 18-19 in 1993. According to the National Audit Office around 56% of the same group obtained NQF level 3 or its equivalent. On the other hand, the figures are 28% for social classes III/IV/V, 8% for classes IV/V, and 13% for class V (YCS in 1993). This means that we should expect HE places, awarded competitively in terms of prior qualification, to be taken disproportionately by those from the higher social classes. This is what we find. Given that social class I is only one fifth of the size of social class II, their combined weighted average age participation index (API) for 1995 is 53% in 2000 (Table 10). This is almost exactly the same as the proportion of each social class attaining level 3 qualifications for entry to HE (the qualification index). This means that social classes I/II were represented in HE entirely proportionately to their prior qualifications. The

participation rate for social class V in 1995 is 12%, again in line with, or even above, their qualification index from the National Audit Office or YCS.

Table 10 – Age participation rate by collapsed social class, UK, 2000

	UK
I/II	53
III/N/M	30
IV/V	8

Source: Callender and Kemp (2001)

In 1989, the proportion of suitably qualified 18-19 year olds who attended HE was 65%. By 1992 that had risen to 90%, and is now higher again. A recent report by the House of Commons Select Committee on Education and Skills suggested a qualified age participation rate of 97%. This is what the analysis above confirms. ‘Lower academic attainment at age 18 accounts for most of the lower participation in higher education by 18 year olds from poorer social classes’ (National Audit Office 2002, p.11). This summary is in line with that of the DfES (2003), which points out that 18% of people from manual or unskilled backgrounds gain two A-levels by the age of 18, and that this proportion is exactly the same as the proportion in HE. Therefore, the qualified age participation index is at or near 100% (although this official statistic is no longer calculated). At the higher end of attainment, for those gaining 25+ UCAS points (the old tariff system), 97% from higher social classes and 94% from lower go on to HE. Of those with 13-24 UCAS points, the figures for participation are 92% and 88% (Connor and Dewson 2001). This means that we can probably explain any stratification in young peoples’ participation in HE by social class almost entirely by the stratification of their prior qualifications. So, to establish that groups are unfairly represented in HE we have to show either that these prior qualifications are unfairly distributed, or that it is otherwise unfair to use prior qualifications as a basis for access to HE.

Conclusions

This paper has shown the difficulties an analyst faces even when making an apparently simple comparison between the characteristics of individuals in HE and characteristics of individuals in the population. So the analyst is faced with a judgement about whether there is indeed under-representation of specific social groups, of whether the proportionate participation of these groups is far enough below one (1) to trigger a search for the cause. The traditional panoply of statistical analyses, such as significance tests, confidence intervals or standard errors, cannot help here because these address only the sampling variation due to chance. None of the many analytical decisions and compromises summarised in this paper concerns such sampling variation (Gorard 2006). They are much more to do with clarity and good judgement. There is no simple consistent answer to the analytical decisions described in this paper, including decisions about the meaning of HE, the nature of domicile, the classifications used, and the age range and geographical span of the relevant population. Yet every analysis covering patterns of participation must make, even by default, a bewildering number of these decisions, and every analyst might quite reasonably make a different set of decisions. Unless these analytical compromises are clearly reported, there is a danger that debates about what is happening in widening participation will be misinterpreted by commentators as being about issues of substance, whereas they are, in reality, merely about differences in these analytical decisions.

The difficulties rehearsed in this paper often lead analysts to focus mainly on young full-time participants taking their first degree, for whom the data is most complete. The relative quality of data for young full-time participants, in turn, leads some commentators to take these elements for granted in an uncritical way in their own smaller scale work. All of this may bias public perception of HE issues, by apparently marginalising part-time and older students. Yet these are two groups that the available evidence shows are the most likely to help create the widening participation that policy, apparently, requires. The paper has illustrated how difficult it is to decide which social groups are under-represented in HE in the UK, if any, and by how much. In fact, the two groups most obviously under-represented in HE at present – males and whites – have been largely ignored in concerns about WP.

Is HE participation lower in less affluent areas because of poverty of aspiration among the relevant age cohorts? Or is it the generally lower initial school qualifications in poorer areas that precede and largely determine the lower level of participation? The answer is crucial for policy purposes, but this issue is rarely addressed directly with even tentative figures. If the former explanation is true, then policies to persuade children from families in poverty of the benefits of HE such as means-tested grants, interest-free loans and so on, are appropriate. If the latter explanation is true, then such policies can have only limited effects. In this case, a much greater emphasis needs to be put into strategies to prepare students of all ages for pre-university qualifications.

The best available datasets appear to suggest that there is no simple and consistent pattern of under-representation among socially disadvantaged groups in attendance to HE, once prior qualifications for entry are taken into account. Any under-representation is already as much in evidence in terms of the possession of entry qualifications at NQF level 3, and these in turn are based almost entirely on staying-on rates in schools and colleges, in turn based almost entirely on NQF level 2 qualifications, and so on (see Gorard and Smith 2004). This, in turn, suggests that WP activities need to be directed at the earlier-life of potential students more than at the point of possible transfer to HE.

References

- Callender, C. and Kemp, M. (2001) *Students in Wales: An analysis of data from the student income and expenditure survey 1998/99*, Cardiff: National Assembly for Wales
- Connor, H. and Dewson, S. (2001) *Social class and higher education: issues affecting decisions on participation by lower social class groups*, DfEE Research Report 267
- Department for Education and Skills (2003) *Widening participation in higher education*, London: Department for Education and Skills
- Goddard, A. and Utley, A. (2004) Review 'papers over cracks', *Times Higher Educational Supplement*, 9/4/04, p.1

- Gorard, S. (2003) *Quantitative methods in social science: the role of numbers made easy*, London: Continuum
- Gorard, S. (2006) Towards a judgement-based statistical analysis, *British Journal of Sociology of Education*, 27, 1, 67-80
- Gorard, S. and Smith, E. (2004) What is 'underachievement' at school?, *School Leadership and Management*, 24, 2, 205-225
- Gorard, S. and Smith, E. (2006) Beyond the 'learning society': what have we learnt from widening participation research?, *International Journal of Lifelong Education*, 25, 6, 575-594
- Gorard, S. and Taylor, C. (2001) *Student funding and hardship in Wales: a statistical summary*, Report to the National Assembly Investigation Group on Student Hardship, Cardiff: National Assembly for Wales
- Gorard, S., Rees, G. and Salisbury, J. (2001) The differential attainment of boys and girls at school: investigating the patterns and their determinants, *British Educational Research Journal*, 27, 2, 125-139
- Gorard, S., Smith, E. Thomas, E., May, H., Adnett, N. and Slack, K. (2006) *Review of widening participation research: addressing the barriers to participation in higher education*, London: HEFCE, 170 pages, available at http://195.194.167.100/pubs/rdreports/2006/rd13_06/
- Lambert, P. (2002) Handling occupational information, *Building Research Capacity*, 4, pp.9-12
- Lee, C. (2003) Why we need to re-think race and ethnicity in educational research, *Educational Researcher*, 32, 5, 3-5
- Mayhew, K., Deer, C. and Dua, M. (2004) The move to mass higher education in the UK, *Oxford Review of Education*, 30, 1, 65-82
- National Audit Office (2002) *Widening participation in higher education in England*, London: The Stationery Office
- Raffe, D., Croxford, L., Ianelli, I. Shapira, M. and Howieson, C. (2006) *Social class inequalities in Education in England and Scotland*, Edinburgh: Education Youth and Transitions Project
- Sanders, C. (2006) Tories: £2bn access drive is a 'failure', *Times Higher Education Supplement*, 28/7/06, p.2
- Selwyn, N., Gorard, S. and Furlong, J. (2006) *Adult learning in the digital age*, London: Routledge

Taylor, C. and Gorard, S. (2005) *Participation in higher education: Wales, Report for independent study into devolution of the student support system: The Rees Review*, Cardiff: National Assembly for Wales