Connolly, P. (2008) (Mis)Representing underachievement: A more detailed reply to Gorard and Smith, Unpublished Paper.
<http://www.paulconnolly.net/publications/article2008a.htm>

Introduction

In 2006 I published a critique of Gorard's (1999) notion of the "politician's error" and his associated advocacy of the use of proportionate methods for measuring differences in educational attainment between groups (Connolly, 2006a). This initial critique attracted a reply from Gorard and a brief response by me to that reply (Gorard, 2006; Connolly, 2006b). Following this initial critique, I published a wider critical review in 2008 of various aspects of Gorard's work on gender and educational achievement that he had published with colleagues over the last decade (see Connolly, 2008a). One significant aspect of this work that I focused on in my review was that undertaken by Smith with Gorard (see Smith, 2003; Gorard and Smith, 2004). This more extended critical review also attracted a reply, this time from Gorard and Smith (2008). While I was given the opportunity to offer a brief rejoinder to this reply (Connolly, 2008b), the limits of space meant that I was unable to address any of the specific points raised by Gorard and Smith in any detail. The purpose of this present paper is therefore to address each of the substantive points raised by Gorard and Smith (2008). As such, this paper should be read in conjunction with these other articles.

My critical review (Connolly 2008a) focused on three substantive areas: the use of achievement gaps by Gorard and colleagues; their consideration of outcomes other than at specific thresholds; and their approach to studying the issue of underachievement. As Gorard and Smith (2008) organized their reply around these three areas then it is worth considering each of these in turn.

Achievement gaps (1)

In my original review I suggest that perhaps the most significant contribution that Gorard had made to date in relation to research on gender and achievement was in relation to his identification and critique of what he has termed the 'politician's error' and his advocacy of proportionate methods for measuring gender differences in attainment in order to address this error. In my review I argued that the simple proportionate approaches that Gorard advocated were 'equally unreliable and also prone to producing fundamentally misleading results' (Connolly, 2008a: 250) and referred to a more detailed outline of my arguments in relation to this specific point that I had previously published elsewhere (Connolly, 2006a).

It is interesting to note that in his reply to my earlier 2006 paper, Gorard failed to address the key concerns I raised about the proportionate approach he was advocating (see Gorard, 2006) and now, in this latest reply, Gorard and Smith (2008) simply ignored the arguments altogether. However, as this is a fundamental problem with Gorard's approach to the study of gender and achievement and, for fear that the point might be lost altogether, it is worth outlining the nature of this problem again with reference to some of my own research.

In 2004 I published a book – *Boys and Schooling in the Early Years* – that reported the results of two indepth ethnographic case studies of young boys' experiences of and attitudes towards schooling. The book attempted to contextualise these case studies with some secondary analysis of national data on gender and achievement (see Connolly, 2004). At that time I found Gorard's (1999) critique of the 'politician's error' extremely compelling; so much so that I used the simple proportionate approach that he advocated in this secondary analysis. Perhaps the main example of this in my book is where I used the simple measure

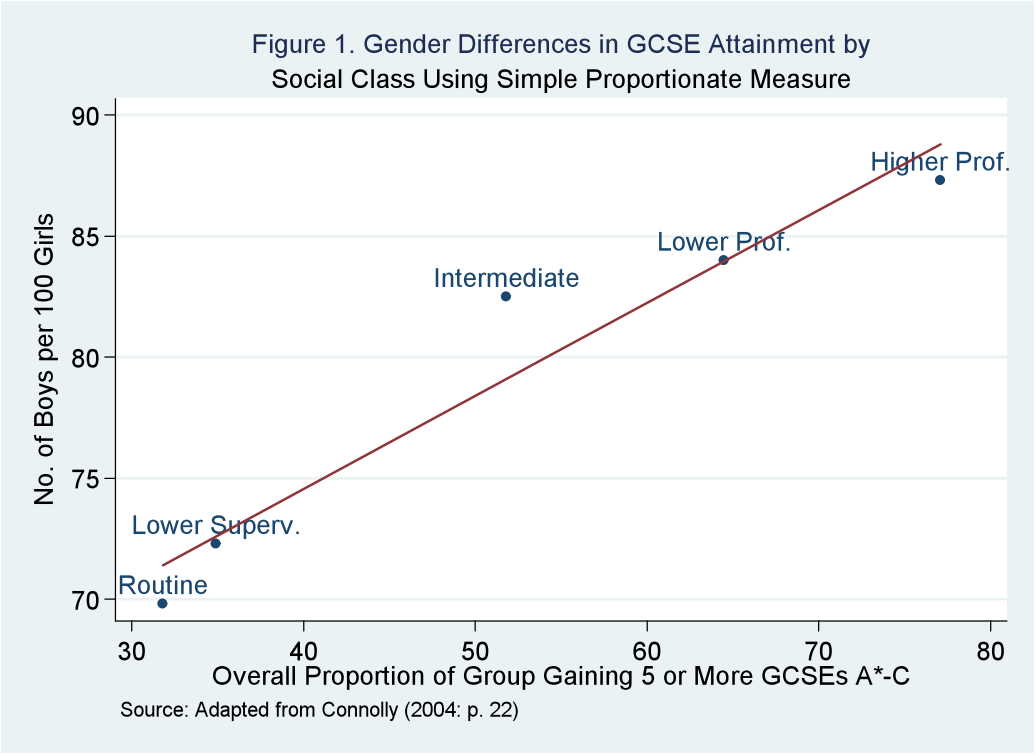
of the number of boys per 100 girls achieving five or more GCSE grades A*-C to explore how gender differences varied across social class categories.[1]

One of the tables I created and included in my book is reproduced below (see Table 1). As can be seen, the higher the social class category, the greater the overall proportion of students gaining five or more GCSE grades A*-C. Moreover, across all groups, the proportion of boys achieving this tended to be less than girls and the gap between boys and girls tended to increase as one went down the social class hierarchy. Thus, as the simple proportionate measure indicated, while 87.3 boys to every 100 girls gained the requisite number of passes among those in the higher professional group, this dropped to just 69.8 boys per 100 girls in the routine manual group. Most intriguingly, this relationship between social class and gender differences in achievement seemed to be almost perfectly correlated as illustrated in Figure 1. Indeed, not only did I include this figure in my book but also reported the correlation between gender differences and overall levels of attainment (r=0.964, p=0.008).

Table 1. Proportions of boys and girls in England and Wales gaining five or more GCSEs grades A*-C in 2000/1 by social class (%)

Category	Total	Boys	Girls	No. of boys per 100 girls
Higher professional	77.1	72.1	82.6	87.3
Lower professional	64.5	58.9	70.1	84.0
Intermediate	51.8	46.7	56.6	82.5
Lower supervisory	34.9	29.5	40.8	72.3
Routine	31.8	26.1	37.4	69.8
Total	51.1	46.1	56.2	82.0

Source: Adapted from Connolly (2004: p. 18)



[1] This is calculated simply by dividing the percentage of boys achieving the required number of passes by the percentage of girls doing the same and then multiplying this by 100. Hence, and from Table 1, for students from higher professional backgrounds we get: 72.1/82.6 = 0.873. Multiplied by 100 gives us 87.3 boys per 100 girls.

As can be seen from reading my book, I uncritically reproduced Gorard's arguments about the 'politician's error' by way of an introduction to using this technique. Moreover, this key finding – that the achievement gap between boys and girls was highest among those from working class backgrounds – provided an important context for the rest of my book and the ethnographic studies that followed. Indeed, so convinced was I of this method for calculating achievement gaps and of the significance of the findings that it generated, that I developed the analysis further, using three successive cohorts of the Youth Cohort Study for England and Wales, and wrote this up as an article for the *British Educational Research Journal* (BERJ) that was subsequently accepted for publication (Connolly, 2006c).

It was only when my article for BERJ sat in line waiting for publication and I was busy preparing to teach an EdD module in advanced statistics, that I noticed the inherent problem in this proportionate approach. Two of the techniques I wanted to illustrate to the students taking the EdD module were logistic regression and loglinear analysis and I thought it would be a good idea to encourage the students to try these techniques out by using the data from my BERJ paper. Given the findings reported above, I thought this would be an ideal way of illustrating the concept of an interaction effect and how this could be formally tested using both techniques. The problem was that when I tried the analysis for myself no such interaction effects emerged. Indeed the findings were clear in terms of indicating that while both gender and social class were significantly associated with the chances of a student gaining five or more GCSE grades A*-C, the effects of these two variables were independent of one another. As such, I decided to recall my paper from BERJ and revise it to reflect these findings. For the interested reader, the findings from the loglinear analysis confirming the lack of an interaction effect between gender and social class can be found in the paper that was finally published in BERJ in 2006 (see Table 4 in Connolly, 2006: p. 13).

This, in turn, led me to explore in more detail the problems associated with the use of proportionate measures of gender differences in the article I published in the *British Journal of Educational Studies* in 2006 where I argued for the need to use standardized effect sizes as a more reliable measure of gender differences (see Connolly, 2006a). While I illustrated the case using the effect size measure, Phi, precisely the same points could have been made using the other main measure of effect size appropriate for data like these – the odds ratio.

To illustrate all of this, Table 2 takes the same data reported in Table 1 and compares the two main proportionate measures of gender differences as advocated by Gorard[2] and the two effect size measures mentioned above. As can be seen, while both proportionate measures suggest the presence of a clear and strong interaction effect between gender and social class (so that the size of the gender gap increases as one moves down the socio-economic categories), no such relationship is evident at all once we make use of either of the effect size measures.

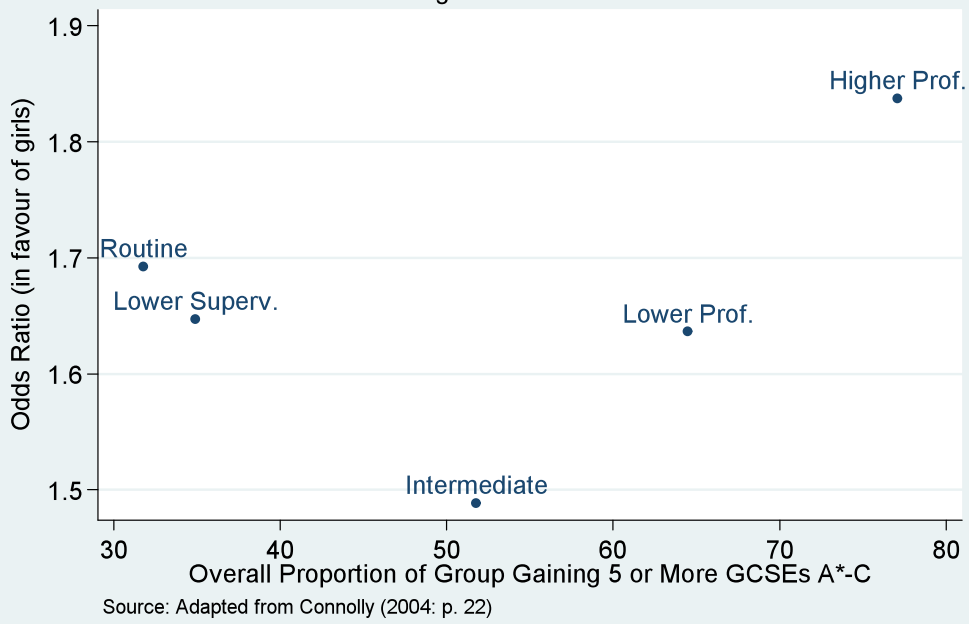
[2] The achievement gap is calculate by subtracting the percentage of girls gaining the required number of passes from the percentage of boys achieving the same and then dividing this by the sum of both percentage figures. As we are working with percentages, this has the effect of controlling for any differences in entry gaps between boys and girls. As an illustration, the achievement gap for higher professional students is:
(72.1–82.6) / (72.1+82.6) = -6.8.

Table 2. Proportions of boys and girls in England and Wales gaining five or more GCSEs grades A*-C in 2000/1 by social class: Comparison of Proportionate and Effect Size Measures

	Total (%)	Boys (%)	Girls (%)	Proportionate Measures		Effect Sizes	
				No. of boys per 100 girls	Achievement Gap (%)	Phi	Odds Ratio
Higher professional	77.1	72.1	82.6	87.3	- 6.8	0.125	1.837
Lower professional	64.5	58.9	70.1	84.0	- 8.7	0.117	1.636
Intermediate	51.8	46.7	56.6	82.5	- 9.6	0.099	1.488
Lower supervisory	34.9	29.5	40.8	72.3	- 16.1	0.118	1.647
Routine	31.8	26.1	37.4	69.8	- 17.8	0.121	1.692

This lack of any interaction effect is also evident from Figure 2 that shows the relationship between the overall proportions of students gaining the requisite passes and the odds ratio in favour of girls for each of the five social class categories. As can be seen, there is no linear relationship this time and the effect sizes seem to vary randomly across the five social class categories.

Figure 2. Gender Differences in GCSE Attainment by Social Class Using Odds Ratio Effect Size Measure



Source: Adapted from Connolly (2004: p. 22)

There is thus a fundamental problem. It is not simply the case, as Gorard (2006) has suggested, that we have different measures and that each has its relative strengths and weaknesses. In this present case, and as illustrated through effect sizes and confirmed formally through the use of appropriate statistical modeling, no interaction effect exists at all between gender and social class. The use of simple proportionate methods to analyse achievement gaps simply give rise to incorrect and misleading conclusions. All of this matters as there are important policy and practice implications that flow from these findings in terms of whether there is a need to target initiatives at boys from more deprived backgrounds or not. The clear evidence in this case is that, if the goal is simply to reduce gender differences in attainment, such targeted initiatives are not warranted and that a more universal approach is required. However, the findings from Gorard's proportionate approach would misleadingly suggest the exact opposite.

Achievement gaps (2)

One of the problems in relation to the above is Gorard's preference for the use of simple arithmetic techniques to address issues where statistical modeling is required. In the above example, the only appropriate way of determining conclusively whether the gender gap varies across socio-economic groups is to formally test whether such an interaction effect exists either through the development of a logistic regression model or a loglinear model. While such techniques do require more specialist knowledge and skill on the part of the researcher, there is no reason why the findings cannot be reported in a clear and accessible manner that are comprehensible to the lay reader (see, for example, Connolly, 2006c).

It is precisely this problem – of opting for simple arithmetic techniques when statistical modeling is required – that is also at the heart of another concern I raised in my review in relation to the attempts, by Gorard and his colleagues, to disentangle the effects of differential entry and differential attainment in the calculation of achievement gaps. As Gorard and Smith (2008: 705) explained:

> The idea of an achievement gap was used … to illustrate the proportionate difference in public examination attainment between males and females, after adjusting for differences in the numbers of each sex entering any examination. Thus, the overall achievement gap is composed of the attainment gap (the proportionate male:female difference in numbers attaining a certain grade) minus the entry gap (the proportionate male:female difference in numbers entering the assessment).

As I explained in my original review, there is much to commend an approach that is able to disentangle the effects of differential entry and differential attainment. Unfortunately, this simple arithmetic approach (of just subtracting the entry gap from the attainment gap) does not do this and actually encourages researchers to reach conclusions from the data that are simply not warranted. This is evident in the following extract from Gorard, Rees and Salisbury (2001: 137) that I quoted in my original review:

> At A level, where students are free from the constraints of the National Curriculum, entry gaps in subject groups tend to be larger, while the achievement gaps, where they exist, tend to be smaller than in equivalent subjects at GCSE. In general, gender appears to play less of a role in assessment at A level.

My key point was that the conclusion offered in the last sentence was not warranted from the evidence provided because differential entry may well mean we are not comparing like-with-like when we consider the attainment of boys and girls at this level. However it seems that, in their reply, Gorard and Smith (2008) have not understood this as they simply go onto re-affirm their original point:

> Our suggested conclusion, which [Connolly] takes exception to, stands. It really does appear that as far as *assessment* is concerned (once entry patterns have been accounted for), there is less of a difference between the sexes at A-level than at GCSE.
>
> Gorard and Smith (2008: 706, *original emphasis*)

The problem here, as before, is that this is simply not a conclusion that one can make without testing this more directly with the use of appropriate statistical modeling techniques. To repeat what I argued in my original review:

> The only way one can validly address the question of whether gender tends to play less of a role in assessment at A-level is to compare the average attainments of males and females at A-level once other potentially intervening variables have been controlled for (such as prior attainment in that subject, possibly some measure of motivation and also the social class and ethnic background of the pupils).
>
> Connolly (2008a: 251)

It should be clear that what I am suggesting here is fundamentally different to what Gorard and his colleagues have done in relation to their own approach to the study of attainment and entry gaps. However as there appears to be some confusion regarding this then it is worth clarifying my point further with a simple hypothetical example of boys and girls' achievement in A-level English. Let us assume, for argument's sake, that while 60,000 girls were entered for this examination only 20,000 boys were. Using the formula outlined and explained by Gorard, Rees and Salisbury (2001), this gives us an entry gap in favour of girls of:

$$\text{Entry Gap} = [(60{,}000 - 20{,}000)/(60{,}000 + 20{,}000)]*100 = 50$$

Furthermore, let us assume that equal proportions of the boys and girls entered for the examination gained a grade A. If we assume that 20% of those entered gained such a grade then this would give us 12,000 girls and 4,000 boys. Again, using the formula provided by Gorard, Rees and Salisbury (2001) to calculate the achievement gap the result we get is:

$$\text{Achievement Gap} = [(12{,}000 - 4{,}000)/(12{,}000 + 4{,}000)]*100 - \text{Entry Gap}$$

$$= 50 - 50 = 0$$

Thus we have a fairly large entry gap of 50 but no achievement gap. I have chosen these figures as they broadly reflect the actual findings of Gorard, Rees and Salisbury's (2001) study. It was precisely findings like these that led Gorard, Rees and Salisbury (2001: 137) to conclude that 'in general, gender appears to play less of a role in assessment at A level'.

To understand why drawing such a conclusion is unwarranted, it is worth looking in more detail at what Gorard, Rees and Salisbury (2001) have actually done when they claim to 'account for differences in entry patterns'. As can be seen, assuming that there are broadly equal numbers of boys and girls in the population then if we went simply by the raw numbers of who achieved a grade A we would conclude that girls were three times more likely to attain this than boys (12,000 compared to 4,000). This is certainly an important finding in itself and draws attention to the under-representation of boys achieving a good grade in English which is an issue that deserves attention. However, this does ignore the differential entry patterns for this subject. In this present example, when this entry gap is taken into account using the method advocated by Gorard, Rees and Salisbury (2001), the differences in achievement disappear.

As the example above illustrates, the reason the differences in achievement have disappeared is that those boys and girls who took the exam were equally likely to gain an A grade. It is precisely with this in mind that Gorard, Rees and Salisbury (2001) feel able to conclude that gender plays less of a role in assessment at A level. Of course, if we were comparing like with like then this would be a valid conclusion to draw. However, and this is the key point, we are probably not comparing like with like. For reasons that Gorard and Smith (2008) acknowledge, as there are far fewer boys electing to take English A level than girls, then it may well be that these boys are, on average, more highly motivated and may also have a higher aptitude for English than their female counterparts. If we assume for one moment that this is indeed the case, then we would surely expect a higher proportion of

these boys to attain a grade A than the girls. The fact that they are still only equally likely to achieve this grade, even though they are more highly motivated and have a higher aptitude in the subject, would suggest that gender does have a role to play either in terms of the content of the A level and the way it is taught and/or in the assessment practices themselves. How else could we explain boys achieving less well than we would have expected of them compared to the girls?

Of course, and in reality, we do not actually know whether the differential entry patterns have resulted in differences between the boys and girls who are taking English in terms of their motivation, aptitude and/or other potential factors. As such, we do not know whether we are 'comparing like with like' when we are analyzing the proportions of boys and girls who gained an A grade. Without knowing this, the fact that those boys and girls who were entered for English were equally likely to attain an A grade, in itself, tells us very little.

This is where the limitations of an approach based on simple arithmetic are once again evident. In contrast, the only way to begin to answer this question of whether gender plays a role in assessment at A level is to use more advanced statistical modeling techniques. In this instance, one way to approach this would be through the use of binary logistic regression models that could be used to ascertain whether there is a difference in the odds of those boys and girls who took English attaining an A grade once as many relevant background factors have been controlled for (most notably their levels of attainment in GCSE English and also their attitudes towards English). This is precisely what I argued in my original review. What I am advocating here therefore, in relation to the need to control for the effects of differential entry patterns, is very different to the limited arithmetic approach adopted by Gorard and his colleagues.

Differential levels of analysis

These points also underpin my concerns regarding the tendency of Gorard and his colleagues to study achievement gaps at each and every level of outcome for any assessment. As I made clear in my original review, there is nothing intrinsically wrong with calculating and displaying the achievement gaps between boys and girls for each and every outcome level for a particular subject. Rather, my concern here is twofold: that such an approach only tends to confuse matters and tells us very little; and, more importantly, it once more encourages misleading interpretations of the findings.

In relation to the former, one of the main contributions that Gorard and Smith (2008: 707) feel that this approach has made to our understanding of achievement gaps is:

> the previously unpublished and unremarked finding that achievement gaps do not exist at the lowest levels of attainment but increase with every grade until A* at GCSE.

This fact that there are few gender differences at the lower levels of attainment is one that is emphasized a number of times in their work and yet it is pretty unremarkable. Indeed, given that the achievement gaps at each level are directly related to one another then one would expect this. To use a simple illustration, let's consider 10 children entered into a race (five boys and five girls). Let us assume that their finishing positions were as follows:

1. Girl
2. Girl
3. Boy
4. Girl
5. Boy
6. Girl
7. Boy
8. Boy
9. Girl
10. Boy

Achievement gaps:

2nd place or above:	100
4th place or above:	50
6th place or above:	33
8th place or above:	0

Clearly, the girls performed much better than the boys (but, as always, there's also a fair amount of variation – for example, there's a boy who came third and a girl who came ninth). However, if we calculate the achievement gaps for some of the thresholds as above then the same picture emerges as Gorard and colleagues found in their own analyses of achievement gaps: the biggest gaps are at the top and the gaps reduce and then disappear altogether when we reach the bottom threshold. However it should be clear from the above that this is bound to happen. Moreover, and in practical terms, is it any solace for boys that they can claim to be doing as well as girls if we count those who came 8th or above?

Again, the problem here is that simple arithmetic approaches to the study of achievement gaps are just not adequate for understanding overall patterns. My own suggestion in relation to the study of achievement gaps in public examinations, that I made in my original review, was that it would be better to focus simply on the thresholds that actually matter (most obviously those attaining a C grade or above at GCSE). However, if Gorard and his colleagues genuinely feel there is value in studying how achievement gaps may vary across differing thresholds then, again, they need to use appropriate statistical modeling techniques. In this case, the use of ordered logistic regression would not only tell us what the odds of girls achieving a higher grade than boys are in a particular subject, and once other relevant contextual factors are controlled for, but could also explicitly test whether such odds vary significantly across the various thresholds identified.

Secondly, and putting all of this to one side, the other key problem with Gorard and his colleagues studying achievement gaps across each and every level of attainment is that it encourages them to draw conclusions that are again unwarranted from the data presented. To illustrate this let us go back to the quotes from Gorard, Rees and Salisbury (2001: 251 & 136) that I used in my original review:

> For all subjects, there are no significant gender differences at the lowest level of any assessment. Otherwise, the gap in attainment between boys and girls rises with every grade or level in an assessment, leading to the conclusion that the problem, if indeed it is a problem, is one mainly facing mid- to high-attainers.
>
> The notion of boys' underachievement being worse at low levels of attainment is simply misjudged. … In terms of the entire age cohort, the achievement gap only appears at high grades, while the proportion of boys and girls 'failing' any examination are remarkably similar, and constant over time. Given the established link between socio-economic conditions and examination success, this does make it highly unlikely the gender gap is primarily to do with disaffected boys from economically depressed families.

The key point to draw out here is not that the main conclusion that Gorard, Rees and Salisbury (2001) make is wrong but that it is based on faulty logic. In this sense, their conclusion that the gender gap is not primarily to do with disaffected boys from economically

depressed families is correct and is also one that I have also demonstrated myself (see Connolly 2006c). However the problem here is with how they arrived at this conclusion.

In this case, when Gorard, Rees and Salisbury claim that the gender gap is not primarily to do with disaffected boys from economically deprived areas, they are essentially claiming that there is no interaction effect between gender and socio-economic background in relation to examination performance. The appropriate way to test this is through the use of statistical modeling and this is precisely what I did in my 2006 article in the *British Educational Research Journal* (see Connolly 2006c). In that article my focus was on the proportion of students gaining five or more GCSEs at grade C or above (popularly known as the GCSE benchmark). In my analyses of the data taken from three successive cohorts of the Youth Cohort Study for England and Wales, I used a combination of logistic regression and loglinear analysis to study the respective influences of gender, social class and ethnicity on attainment and to specifically test whether an interaction effect existed between gender and social class (and between gender and race).

Rather than attempting to test whether such an interaction effect exists or not using appropriate statistical modeling, Gorard, Rees and Salisbury (2001) relied on a (faulty) interpretation of the findings they had generated through their simple arithmetic approach. In essence their arguments proceeded as follows:

- Achievement gaps only appear at high grades and all but disappear at low grades.
- To the extent that a gender gap exists it is therefore among those boys who are high attainers and not among those who are low attainers.
- We know that there is an established link between socio-economic background and attainment such that those from economically deprived backgrounds are more likely to be low attainers.
- Thus, as boys from economically deprived backgrounds are low attainers, and we know that the gender gap is smallest among low attainers, then: “it is highly unlikely the gender gap is primarily to do with disaffected boys from economically depressed families.”

The key problem with the above logic is that it relies upon an acceptance of a simple relationship between socio-economic background and low attainment. And yet we know, as I made clear in my original review, that while an association exists it is far from perfect. Indeed the consistent finding across many studies is that socio-economic background can only account for less than 10% of the variation in educational achievement. In other words, 90% of the variation in achievement between pupils remains unaccounted for and is thus due to other factors. Moreover, when we consider the inherently unreliable nature of the achievement gap measure itself and also the problems associated with interpreting how achievement gaps vary across differing levels of attainment as outline above, then the whole argument comes crumbling down.

Perhaps most tellingly, by the same logic used above Gorard, Rees and Salisbury (2001) should have actually conclude that there is likely to be an interaction effect between gender and socio-economic background but one that is in the opposite direction to perceived wisdom; namely, that the biggest problem in relation to boys’ underachievement is among those from the most affluent backgrounds. After all, so their logic goes, we know that the largest gender gaps appear amongst the highest attainers and that the highest attainers are most likely to come from the highest socio-economic backgrounds.

If this logic held up then Gorard, Rees and Salisbury (2001) could be criticized for failing to identify a key policy and practice implication from their work: that targeted efforts are required after all to raise performance among boys from the most affluent areas. However the logic clearly does not hold up and we are left, once more, with having to deal with the

findings derived from the inappropriate application of simple arithmetic to what is a more complex problem that requires more appropriate use of statistical modeling.

Defining and measuring underachievement

The final key area that I drew attention to in my review of the work of Gorard and colleagues was in relation to their research around the issue of underachievement and, more specifically, their interpretation of the findings of research they conducted in this area. As I made clear in my original review, I have no concerns with the approach adopted by Smith and Gorard in this regard. Indeed I continue to believe that tracking the progress of pupils over discrete time periods and analyzing their relative changes in performance over such periods has many merits in helping to determine the varying effects of gender on children's educational trajectories. The research reported by Smith (2003) in this regard is therefore potentially very useful. As I argued:

> [This method] is very useful, for example, to examine the effects of gender over certain discrete time periods. This, in turn, may well provide some indication of the way in which gender tends to operate differently for differing groups of boys and girls over the course of their schooling careers. [In relation to Smith's (2003) research], for example, it is interesting to note that little (relative) change takes place in terms of working-class boys' levels of attainment between the ages of 11 and 14. This certainly suggests that the key processes that come to establish their particular patterns of low achievement in comparison with other groups occur at an earlier age, certainly in primary years. By contrast it is equally interesting to note that the processes that impact upon working-class girls and that limit their educational attainment would seem not to be complete by the end of the primary years but are still in operation during the first few years of secondary schooling. Clerly, more detailed exploration of such trends, and the analysis of the data upon which they are based, is needed. This certainly indicates an important use for the type of analysis provided by Smith (2003). The problem in this case therefore lies not with the approach taken but with the interpretation of the findings produced.
>
> Connolly (2008a: 255)

My concern here, therefore, is more specific and relates to how they interpret the findings they generate from such analyses. To illustrate this, let us take just some examples of the claims made by Smith (2003) in her article published in the *British Journal of Sociology of Education*:

> Using a stricter definition of underachievement that takes into account a range of background as well as academic variables when predicting examination performance, this paper described the construction of a model for predicting individual performance in future Key Stage 3 examinations. Individuals whose examination performance was then significantly lower than expected were termed underachievers. The composition of this group was examined in light of existing 'moral panics' about underachieving working-class boys. [*abstract*, p. 575)
>
> This paper … is framed around two broad research questions: how can underachievement be defined and measured? And who are underachieving pupils? (p. 576)
>
> The data [for this study] were used to predict academic performance in future National tests at age 14 and to identify those who did not fulfill their 'potential' – the strict notion of under-achievement. (pp. 577-8)

> One of the most striking findings from this analysis was that the variable for gender was ranked seventh in importance and explained only 0.8% of the variance in academic achievement at Key Stage 3. … this result suggests that the impact of gender on overall achievement might not be as significant as was once thought. (p. 581)
>
> [T]here was no suggestion that, in this study, the working-class boys were underachieving. However, the second finding concerns the distribution of girls in the 'underachieving' group, according to their social class. Here there were significantly *more* working-class girls in the 'underachieving' group than were to be expected. There were also correspondingly *fewer* of these girls in the 'overachieving' group. Therefore, while the results indicate that the prediction that the working-class boys would be underachieving cannot be upheld, it does appear that a different group of 'underachieving' pupils has been identified – working-class girls. (*original emphases*, p 583).
>
> The results indicate that the pupils who were identified in this study as underachieving were not predominantly working-class boys, findings that are contrary to popular accounts. However, the results do raise questions about the relative performance of some of the girls. (*conclusions*, p. 584)
>
> That the relatively poorer working-class pupil generally does not do as well in school as their more affluent counterparts may well be the case, but there is little to suggest that this group of pupils was underachieving. (*conclusions*, p. 585)

What should be clear from all of these quotations are the universalist claims that Smith (2003) makes from her data about the performance of working-class boys and girls *in general*. No attempt is made, at all, to contextualize these claims and to explain the crucial point that all of these relate to the boys' and girls' performance over a very specific period (from the age of 11 to 14). It is thus a shame that when Smith and Gorard do finally choose to make use of appropriate statistical modeling techniques they fail to interpret the findings properly.

To avoid any doubt here, let me just re-iterate the point I made in my original review of this work that I have no concerns with the analysis undertaken, the statistical models generated and/or how they were analyzed. Indeed, this article should have stood as an excellent example of how to apply statistical modeling to an important area of concern and to report the findings in a clear and comprehensible manner for the lay reader. If all the quotations above had been properly contextualized in terms of how they related to patterns of achievement specifically between the ages of 11 and 14 then that would have been absolutely fine and valid. However, this was not done and thus it is the interpretation and reporting of the findings that is the issue here. Moreover, this is not just a minor problem – and one that could be presented as unfair 'nit-picking' on my part – but a fundamental error; one that is likely to completely mislead readers. From the quotations provided above, how can one expect the lay reader to conclude anything other than that working class boys are not underachieving at all but that working class girls are?

In relation to some of the more specific responses offered by Gorard and Smith (2008) in their reply to these charges:

- It is simply not good enough to say, as they did, that these contextualizing points were made elsewhere, especially when some of the references used were published some two years or more *after* this original article came out! The basic point is that any article should stand alone in terms of the arguments it makes. In this present case, there is also no excuse in that these are fundamental points regarding the

need to avoid universalist claims and thus to contextualize the findings. It is just surprising that this was not picked up in the peer review process[3]

- I stand by my concerns that, while not the intention of Gorard and Smith, interpreting the findings in this way is likely to encourage others to believe that there is something inherent or innate about working class boys and poor achievement. The simple point is that working class boys are underachieving but that the processes that lead to this underachievement would seem to take place within preschool and primary school so that they are largely entrenched by the time they start secondary school. Simply ignoring this broader picture and claiming that the working class boys' poor performance is not underachievement leaves the reader with little option than to assume that their lower levels of achievement must be somehow to be expected and thus are innate.
- In relation to the failure of Gorard and Smith (2008) to recognize the distinction between sex and gender, this is simply confirmed in their reply when they re-iterate their previous arguments that: 'a basic problem with identifying sex as a factor in attainment is that we cannot change the sex of the student (usually) and so to make practical progress we need to focus on identifying the factors that educators can change (and with which sex might interact)' (p. 710-11). Of course these 'factors' are what constitute 'gender' and can only be identified and understood by studying sex differences and, more specifically, how boys and girls differ in relation to their experiences, perspectives and outcomes. It is simply difficult to know why Gorard and Smith do not appear to understand this.

Conclusions

In their reply, Gorard and Smith (2008) appeared to be genuinely perplexed regarding the reasons why I decided to write a critical review of all of this work in the first place. For them, and as they said, my original review was largely a case of presenting an argument for argument's sake. Moreover, an undertone evident in much of their reply was a deeper suspicion regarding my true motivations for critiquing their work. It is therefore hoped that this more detailed reply helps to clarify the reasons that led me to engage with their work in the way that I have.

In particular, it should be abundantly clear that there is no hidden agenda behind my critiques. Indeed, and as explained above, I was originally very attracted by the approach to quantitative research in education offered by Gorard. There can be no clearer evidence of this than the fact that I unquestioningly used his methods prominently in my own work (Connolly, 2004). Moreover, and if anything, his original arguments regarding the politician's error and the way it had fed into a moral panic regarding boys' underachievement resonated closely with my own feelings around this issue. The same is true regarding Smith's (2003) article and the fact that, when I first read it, I very much welcomed what appeared to be further evidence to undermine the prevailing discourse at the time regarding boys' underachievement. I therefore, genuinely, had no political motive or reason to criticise and undermine their work.

Rather, what has motivated me is the desire to promote the appropriate use of quantitative research in education and to see it done properly. Of course there are plenty of examples of bad quantitative research out there so, it may be asked, why did I specifically choose to focus on the work of Gorard and colleagues? The answer to this is in part due simply to the fact that I became familiar with his work on gender and education precisely because it fell directly into one of my own core areas of interest. Once I had made use of some of the

[3] Which takes us back to one of the concluding points I made in my original review regarding the more general problems associated with the limited capacity within the education research community in the UK in relation quantitative methods.

proportionate techniques he has so forcefully advocated and found them to be unreliable, I was obviously going to be motivated to look in more detail at this and then other areas of his work.

However, and beyond this, perhaps the key motivating factor that encouraged me to engage in a wider critical review were the dangers associated with researchers adopting the type of simple arithmetic approach to the analysis of quantitative data as advocated by Gorard rather than analyzing it properly with the use of appropriate statistical modeling techniques. Given that Gorard is arguably one of the most influential and well-known academics within the UK education research community with respect to quantitative methods, then he certainly has the potential to encourage many researchers down this simple arithmetic road; especially given his forceful advocacy of such methods and often trenchant criticism of others. The dangers inherent in the education research community moving in this direction, as illustrated above, are such that this cannot go unchallenged. It is therefore hoped that this present paper, and the earlier exchanges that gave rise to it, will at least encourage others to reflect upon the direction that Gorard is encouraging quantitative researchers to take. While this present exchange has focused on research in the area of gender and education, the issues raised are much more fundamental and thus deserve a much wider critical debate.

References

Connolly, P. (2004) *Boys and schooling in the early years.* London: Routledge.

Connolly, P. (2006a) Keeping a sense of proportion but losing all perspective: A critique of Gorard's notion of the 'politician's error'. *British Journal of Educational Studies* 54: 73–88.

Connolly, P. (2006b) No new perspectives, just red herrings: A reply to Gorard. *British Journal of Educational Studies* 54: 476–82.

Connolly, P. (2006c) The effects of social class and ethnicity on gender differences in GCSE attainment: A secondary analysis of the youth cohort study of England and Wales 1997–2001. *British Educational Research Journal* 32: 3–21.

Connolly, P. (2008a) A critical review of some recent developments in quantitative research on gender and achievement in the UK. *British Journal of Sociology of Education* 29, no. 3: 249–60.

Connolly, P. (2008b) (Mis)Representing underachievement: a rejoinder to Gorard and Smith. *British Journal of Sociology of Education* 29, no. 6: 715-717.

Gorard, S. (1999) Keeping a sense of proportion: The 'politician's error' in analysing school outcomes. *British Journal of Educational Studies* 47: 235–46.

Gorard, S. (2006) Yet another perspective: A response to Connolly. *British Journal of Educational Studies* 54: 471–5.

Gorard, S., and E. Smith. (2004) What is 'underachievement' at school? *School Leadership and Management* 24: 205–25.

Gorard, S., and E. Smith. (2008) (Mis)Understanding underachievement: A response to Connolly. *British Journal of Sociology of Education* 29: 705–14.

Smith, E. (2003) Understanding underachievement: An investigation into the differential attainment of secondary school pupils. *British Journal of Sociology of Education* 24: 575–86.