Modern English Education, Vol. 11, No. 2, Summer 2010

A Review of Journal Article in *Educational Research*: "Challenging Underachievement in Boys"

Jong-Hee Lee
(Kangwon National University)

Lee, Jong-Hee. (2010). A review of journal article in *Educational Research*: "Challenging Underachievement in Boys". *Modern English Education*, *11*(2), 254-277.

This review of a journal article in *Educational Research* demonstrates that the reported research contains a number of crucial shortcomings in both overall research designs and specific research procedures which are deemed to threat the validity and reliability of the widely-recognized international journal. Specifically in the research methods, the authors of the journal article are considered to fall short of the tenable foundations for their mixed-method research in a few key issues; selecting equitable predictor and criterion variables, overcoming false positives in identifying experimental target schools, taking the essential steps for probability sampling, and avoiding biased approaches to semi-structured interviews. As a result, it seems to be increasingly clear that the consumers of any journal article should be wary of its possible defects and misconceptions in assessing the overall generalizability of research findings.

[educational research methodology/multilevel modeling/mixed-method research design/교육연구방법론/다층모형/혼합연구방법설계]

I. INTRODUCTION

The main purpose of this review is to subject the international journal paper entitled "Challenging underachievement in boys" (Lindsay & Muijs, 2006; see Appendix for relevant details) to critical examinations in terms of educational research methodology. Such a review is expected to guide the consumers of journal articles to arouse their critical awareness of potential defects in the validity and reliability of those articles, so that they can monitor and reformulate their routine approaches to processing various academic papers in the fields of educational research.

In doing so, I will first describe the research context and the general approach used by the researchers, then summarizing the main claims made in the journal paper. After that, I

will go on to discuss the strengths as well as the weaknesses of the reported research. Much attention will be given to exploring considerable weaknesses found within the research methods described in the journal paper. With this in mind, after touching on some strengths of the journal paper, I would like to address a number of significant weaknesses found in the reported research. Finally, I will make some suggestions on how the research could be improved.

Let me commence with the initial examination of the general context involved and approach used in the research project. It has been stated that one of the challenges to the British education is the underachievement in boys with ethnic minority backgrounds. The journal paper focused on such gender and ethnic issues shifted from a social class variable, reporting and identifying a number of reasons for which the Local Educational Authority (LEA) and some primary and secondary schools have succeeded in solving the underperformance problems of said pupil groups.

To explore the characteristics of educational infrastructure developed by those successful schools, the researchers took the mixed methods of 'quantitative research' using multi-level modeling for regression analysis among the schools and 'qualitative research' analyzing the data collected from semi-structured interviews with relevant subjects. In line with this, the research project was also supported by both scientists with expertise on research methodology and the hands-on knowledge of key members from the LEA's Scrutiny Commission. The following chapter will provide a summary of the claims in the reported research as well as the critical examination of several main issues found in the journal paper.

II. CRITICAL REVIEW OF JOURNAL ARTICLE

1. Summary of Claims in Reported Research

The reported research was designed primarily to investigate the characteristics of schools that demonstrated success in overcoming the underachievement of boys within particular ethnic groups. The authors of the journal paper argued that the research project was an exemplary action research conducted by the educational experts with the support of key members from the LEA's Scrutiny Commission. Given a mixture of quantitative and qualitative methods, the authors stated that this combination was more popular than others in educational research.

At stage 1, through quantitative methods, they identified three primary and two secondary schools as successful models from the parent groups of schools (i.e., among 58 primary and 9 secondary schools) by using multilevel modeling. In this process, they

argued that the target schools were selected on the basis of academic outcomes above expectation that particular minority ethnic groups attained in key stage 2 results (in case of primary schools) and the total GCSE scores (in case of secondary schools). As a rationale for their final identification of target schools, they stressed that the scores and confidence intervals of three primary schools and two secondary schools were entirely above the zero horizontal line of prediction, and so those schools could be concluded to be adding value for particular ethnic minorities (i.e., black African, black Caribbean, and white UK-born boys). Later one more sample was drawn from a secondary school that stood out consistently. As a result, three primary and three secondary schools were finally selected for visits. At stage 2, through qualitative methods, they attempted to draw up the characteristics of those schools on the basis of semi-structured interviews with head-teachers, teachers and pupils.

According to the authors of the paper, a key finding at stage 2 was that the identified target schools took two strands of approaches to improve school effectiveness. They stated that one approach stressed the insistence on equally high standards for all pupil groups, and the other approach was primarily to target specific underperforming groups. The first approach was that high expectations and academic achievement were emphasized for all groups without any biased treatments. The second approach was that specific support and guidance were concentrated on and provided for particular ethnic/gender groups. This umbrella strategy also incorporated six more concrete factors such as (1) diverse curricular approaches, (2) individual performance monitoring, (3) high and realistic expectations, (4) relatively adequate staffing from ethnic minority groups, (5) strong ethos of the school as a community, and (6) strong partnership with parents.

In light of this, the researchers argued that the characteristics of successful schools were identified from the thematic analysis of the interview data, which played major roles in overcoming underachievement of three particular ethnic groups as the focus of the research: white UK-origin boys, black African and black Caribbean, and white UK-origin boys. Accordingly they stressed that the research project, as a good example of democratic action research, provided evidence for the target schools' successful management in overcoming relative underachievement of boys within particular ethnic groups.

2. Review of Strengths and Weaknesses

1) Strengths of the Reported Research

(1) Advantages of Multilevel Modeling

As a whole, the reported research exhibits some strength, although limited in its scope of application, through using multilevel modeling at the two hierarchical levels of individuals

and school domains in order to identify two groups of primary and secondary schools demonstrating success in overcoming underachievement. The researchers stated that multilevel modeling was undertaken to compare the subjects' achievement outcomes (criterion variables) with those data expected on the basis of control measures (predictor variables); (1) prior achievement, (2) eligibility for free school meals, (3) the Special Educational Needs(SEN) Code of Practice stage, (4) English proficiency, and (5) term of birth. As a result of using multilevel modeling, the researchers could obtain the reliable data sets showing the residual point estimates and confidence intervals of two parent groups of schools they selected for statistical analyses, and identified two sets of target schools demonstrating the scores and confidence limits entirely above the zero horizontal line of prediction.

In this regard, the causal relationships between predictor and criterion variables noted above do not appear to be salient due to the predictor variables which do not reflect a range of disadvantaged factors causing underachievement of ethnic minority boys. Thus, I wish to point out that although such research designs may be *invalid*, the statistical tool – *multilevel modeling* – used to select initial target schools is largely *reliable*, as an advanced technique, to improve the predictive power by giving equal weight on both individuals and their school contexts, which are often hierarchically structured (Kreft & De Leeuw, 1998).

(2) Advantages of Mixed-Methods Approach

The advantages of mixed-method designs rest mainly on allowing researchers to combine the strengths of each approach and to gain more reliable source data of research topics, thus to look for outcomes or products (quantitative) and to explore the process (qualitative) through which particular behaviors or results occur (Kubiszyn & Borich, 2003). The reported research chose one of the common types of mixed-method study – an explanatory design – in which the quantitative method was used to obtain numerical evidence deemed to stem from the major factors involved in the topics above, and on the basis of initial outcomes, the qualitative method was conducted to draw up a range of contextual factors supporting the quantitative outcomes. Thus, it is considered relatively perceptive for the researchers to have used such a mixed-method research design.

In light of this, the reported research used the designs that combine both quantitative and qualitative analyses, which is called 'mixed-methods': (1) the first phase of quantitative research constituted the multilevel modeling-based identification of three primary and two secondary schools deemed to have succeeded in overcoming the underachievement of particular target groups. They selected a total of five schools on the basis of mean scores and confidence intervals which were 'entirely above the central horizontal line of prediction' (Lindsay & Muijs, 2006, pp. 318-319); and (2) the second phase of qualitative research constituted interviews conducted by using semi-structured schedules to elicit

general characteristics to result in 'success' in bucking trend of underperformance. The thematic analyses of the interview data collected from narrative descriptions generated a wide range of contextual factors: curriculum, performance monitoring, high expectations, staff, inclusive ethos, and parents.

2) Weaknesses of the Reported Research

Based on the methods used in the first and second stages of research project described above, I intend to explore a number of weaknesses of the journal paper which are found not only in overall research designs but also in specific research procedures.

(1) Overall Research Designs

① Problems in Research Design Features

The reported research was designed to shed light on the characteristics of successful primary and secondary schools in overcoming their particular ethnic minority groups' underachievement in boys. In the first stage, the researchers identified two groups of target schools on the basis of their respective particular pupils' academic performance (criterion variables) by applying multilevel modeling, and then, through interviews in the second stage, they investigated a range of characteristics of the schools (predictor variables) which made contributions to overcoming the pupils' underachievement. Thus, the researchers should have scrutinized how the predictor variables (evidence of school effectiveness as 'control measures') influenced the criterion variables (evidence of overcoming underachievement as 'outcome measures'); the two discrete procedures of their research needed to be reversed with each other, so that they could implement causal and comparative studies on the two sets of control and outcome measures.

In this reported research, dependent variables are the measured outcomes which involve changes in pupils' accomplishment attributable to the positive influence of the independent variables. In brief, predictions should be made from independent variables to dependent variables. Therefore, the reported research failed to follow the valid procedures of prediction studies. In other words, the researchers virtually attempted to explore a set of 'causes' (positive intramural factors including parent involvement – predictor variables) by means of 'effects' (academic results above the level of expectation – criterion variable), which is the completely reversed order. This is considered to be one of the crucial weaknesses in the overall research designs.

② Criteria for "Success" in School Qualifications

Another key problem with the journal paper lies in the invalid criteria for measuring "success" which is a term used to identify the target schools for visits in subsequent

qualitative research. As described earlier, as measures of achievement, the researchers used key stage 1 results (predictor variable) and key stage 2 results (criterion variable) in case of primary schools, and key stage 3 results (predictor variable) and the total GCSE scores (criterion variable) in case of secondary schools, with four other contextual factors.

In this regard, the researchers seem to have believed that a range of five factors given above are 'key variables known to have an impact on (academic) attainment'. Then they considered three primary and two secondary schools to be *successful* which demonstrated the scores of outcome measures and confidence limits entirely above the zero line of prediction. But it would seem inaccurate and therefore unreliable to conceive only such 'results above expectation' in criterion variables as valid evidence for *success* in school qualifications related to overcoming underachievement.

There are two main reasons for their misconceptions: one is that they did not compare the pupils' attainment results in control measures with those in outcome measures to examine any possible gap between the two separate test scores; the other is that they did not take into account widely verified criteria for 'successful' and 'non-successful' schools in prior research findings related to the improvement of ethnic minorities' underachievement. According to the educational research literature, there are four key criteria used to measure the causal relationships between ethnicity and school performance within a hierarchical structure: (1) racial centrality (how important being Black is to one's identity), (2) public regard (one's belief about how others perceive an individual's racial group), (3) perceived teacher discrimination, and (4) mother's level of education (Thomas, Caldwell, Faison & Jackson, 2009).

However, apart from the four essential criteria noted above, I wish to suggest how to define the standards of 'success' in school qualifications with a focus on improving underachievement. Therefore, it would have been far more valid and reliable for the researchers to select the primary and secondary schools as 'successful' institutions from the parent groups of schools, in cases where those institutions were verified to be distinguishable from non-successful schools in *statistical significance* through the three steps of data treatments with such *predictor variables* (PV) and *criterion variables* (CV) chosen in the research as follows:

(A) Means, standard deviations, and alpha reliability for variables in research; (a-1) key stage 1 results – PV and (a-2) key stage 2 results – CV in case of primary schools, (b-1) key stage 2 results – PV and (b-2) actual GCSE scores – CV in case of secondary schools, and as PVs for both groups of schools, (c) eligibility for free school meals, (d) the SEN Code of Practice stage, (e) English proficiency, and (f) term of birth;
(B) Correlation coefficients among PVs and CVs among total sample; and

(C) Correlation coefficients among PVs and CVs among black African boys, among black Caribbean boys, and among white UK-origin boys.

③ Issues in Generalizability of Interview Data

The final key weakness involves the overall design features where by the researchers attempted to generalize a number of contextual factors through the interviews conducted, thus to look for only positive reasons in successful schools in overcoming underachievement: (a) diverse curricular approaches, (b) individual performance monitoring, (c) high and realistic expectations, (d) adequate staffing from ethnic minority groups, (e) strong ethos of the school as a community, and (f) strong partnership with parents.

In my view, it is not valid to believe that a range of 'successful' school characteristics elicited from analyzing interview data would be distinguishable from those of 'non-successful' schools in aspects of underperformance. The main reason for this is that the researchers did not verify the reliability of the factors noted above through both *deductive and inductive data analyses* based on relevant previous research findings. In the first place, they should have confirmed a general set of prevalent factors involved in causal relationships between ethnicity and school performance in a deductive approach. Then, as an inductive approach, they should have compared a new set of themes elicited from their analyses of interview data with those generated by prior research activities.

For example, through the interviews, the researchers found a broad and diverse curriculum designed in all secondary schools as contributing to success with the targeted groups. These schools used a comprehensive curriculum to reduce possible detrimental effects to ethnic minorities although there were differences in degrees of targeting. They also recognized that to effectively deal with underperforming groups, the schools took a combinatory mode between encouraging students to focus on high academic achievement and orientating the school/staff members to community values. In my opinion, these approaches to schooling effects may not be greatly differentiated from those conducted by certain non-successful schools.

(2) Specific Research Procedures

① Selection of Predictor and Criterion Variables

The reported research contains several important weaknesses in its validity and reliability in the processes of identifying target primary and secondary schools for the subsequent qualitative research based on semi-structured interviews. The researchers used key stage 2 results of English, maths and science (for primary schools) and the total GCSE points score (for secondary schools) as a dependent (criterion) variable, and selected five independent (predictor) variables:

(A) Key stage 1 results of reading & maths (for primary schools), or key stage 3 results of English, maths & science (for secondary schools) as the measure of prior achievement;
(B) Free school meal eligibility;
(C) The SEN Code of Practice stage;
(D) English proficiency; and
(E) Term of birth.

The researchers also stated that the most important of predictor variables was the pupils' prior attainment, and analyzes were carried out on two years of data, 2001 and 2002, in order to check stability of the school influence. In this regard, several requirements should be met to gain the significant level of accuracy in statistical predictions: (a) the reliability of the measures of the predictor and criterion variables; (b) the reasonably short span of time between the two variables; (c) the reasonably simple criterion variables; and (d) the plurality of predictor variables (McMillan, 2004). With these conditions in mind, I would like to explain several significant weaknesses of statistical bases used for reliable prediction as follows:

(a) Problems in Identifying Primary Schools

In the case of selecting primary schools, the predictive power of control measures (independent variables) used to identify the target schools is not adequate enough, due to the individual characteristics of predictor variables described above. These predictor variables seem to be highly correlated with one another in the trend of particular ethnic minorities' underachievement. Here, items (2) and (3) in the predictor variables above seem to be overlapped largely with each other, and in terms of language skills, item (4) is deemed to be similar to reading in item (1).

In general, statistical predictions become more accurate if additional predictor (independent) variables are used and the multiple predictor variables are not correlated with one another. In this reported research, the five independent variables were combined and analyzed as control measures against one outcome measure (dependent variable) – key stage 2 results of English, math and science tests taken at age 11. So, the researchers should have selected a plurality of independent variables which are far less correlated with one another, but are much more correlated with the dependent variable. It is necessary to set up reliable criteria that can distinguish successful schools from non-successful schools in reducing or overcoming underachievement of particular boys from ethnic minority groups.

In doing this, to determine reliable predictor variables, the combination of educational

measures and socio-demographic/background measures having a direct impact on school performance of the pupils would be recommendable. For example, in addition to (1) key stage 1 results, (2) disciplinary actions, and (3) placement in remedial classes, as educational measures, the researchers may be advised to use the following significant variables:

(A) Racial centrality (how important being Black is to one's identity);
(B) Public regard (one's belief about how others perceive an individual's racial group);
(C) Perceived teacher discrimination; and
(D) Mother's level of education, as socio-demographic and background factors.

In this context, it is essential to recognize the key fact supported by many other research findings: the socio-demographic and background measures specified above are correlated highly with academic achievement in comparative studies on the relationships of ethnicity and school performance (Thomas et al., 2009).

(b) Problems in Identifying Secondary Schools

In the case of selecting secondary schools, the predictive power of control measures used to identify them is much weaker than that of control measures used for primary schools. The reason is that the key stage 3 results of English, math and science tests were gained by pupils after three years of being in secondary school. More seriously, the total GCSE points score used as their outcome measure (dependent variable) is also unreliable on the ground that intermediate GNVQs are counted as being worth four conventional GCSEs at grades A* to C, and school courses for the vocational qualifications are widely known as "soft options" (BBC NEWS, 2007). Such a weighting practice is said to make the GNVQs particularly appealing to schools which are eager to improve their league table position because league tables are compiled without drawing a reasonable distinction between GCSEs and GNVQs.

In this context, recognizing the needs for a diverse curriculum to contribute strongly to success with the targeted groups in secondary schools, the authors of the paper state that the option in key stage 4 (14~16-year-olds) of taking GNVQs with vocational orientation, equivalent to a number of GCSE subjects, was used to improve achievement and percentage of five GCSE passes A*–C. They also exemplify the fact that at one school the large majority of the Year 11 group (15~16-year-olds) took GNVQs in Information and Communications Technologies (ICT), and the results were described by the head-teacher as 'outstandingly successful' (Lindsay & Muijs, 2006, p. 320). This supports that the GCSE scores (at age 16) used as the outcome measure to identify successful secondary schools are not valid in the reported research. Therefore, the statistical bases of both key

stage 3 results and total GCSE points score may threaten the reliable prediction of positive school effectiveness in the reported research.

In addition, in the published paper, the researchers do not provide any tangible evidence for the comparative studies on their multilevel modeling analyzes carried out on two years of data spanning 2001 and 2002, although they stated that such data collection and analyzes with respect to the two years were implemented to confirm the sustainability of the school influence.

② Issues of False Positives in Identifying Target Schools

The authors of the journal paper commit False Positives (Type I error) in identifying particular primary and secondary schools from the two parent groups of schools initially examined for their subsequent qualitative research. In other words, they mistakenly thought that there were any significant differences between the target schools selected as samples and their parent population. The finally selected three primary and two secondary schools but the other schools in the two preliminary groups of schools *differ only by chance* in evaluating the overall characteristics of the target schools.

It is generally accepted that (a) if the difference between the mean achievement of the experimental and the control groups is too great to attribute to the normal fluctuations that results from sampling errors, the experimenter may reject the null hypothesis; and (b) when samples are large (more than 30 in size), the t critical value approaches the z score, and in these cases if the z value equals or exceeds 1.96, the researcher may conclude that the difference between means is significant at the 5% (0.05) level of significance (Best & Kahn, 2006, pp. 408-409).

However, in the reported research, the mean scores of 26 primary schools above the zero line of prediction and those of 6 secondary schools above such a criterion line (see two graphs on p. 318 of the journal paper) are not that great, so that the difference can be merely the outcome of sampling error. Further, even though the actual number of particular subjects in each of primary and secondary schools having been investigated is deemed to be less than 30 in the critical sample size, the researchers argue that 'the vertical bars represent the 95% confidence bands and hence statistical significance' (Lindsay & Muijs, 2006, p. 318).

In addition, the authors do not provide any accurate number of pupils having ethnic minority backgrounds in each of the primary and secondary schools they identified to examine their school effectiveness. They explicitly reveal only the number of schools in each group of primary and secondary schools for their initial investigation although how many schools and particular students in each of those schools are clearly recognized as statistical parameters in the reported research. Therefore, the researchers' argument for 'the 95% confidence bands and hence statistical significance' may lose its valid and reliable

foundation for their research.

③ Essential Steps for Probability Sampling

The researchers and the Scrutiny Commission who selected their subjects and schools are considered invalid. In identifying the target primary and secondary schools at stage 1, they did not follow any of the essential steps in probability or random sampling. Typically, to obtain a sample which adequately represents the target population, random sampling is required, which implies that each member of the population as a whole or of subgroups of the population has an *equal chance* of being selected to provide an *accurate description* of the entire population (Reynolds, Livingston, & Wilson, 2006).

Let me specify the five exemplary procedures for random sampling to solve the selection problems found in the reported research:

(A) To define the target population, (e.g., all primary and secondary school boys and their ethnic groups who show underachievement in England);
(B) To identify the sampling frame, (e.g., list of the boys and their ethnic groups from the State Department of Education in England);
(C) To identify the sample size, (e.g., selecting 10% of the boys in such ethnic groups);
(D) To select method of sampling, (e.g., stratified random sampling by the grade levels of the primary and secondary schools); and
(E) To select sample, (e.g., selecting specific boys for sample).

Therefore, I would like to stress, and argue my point, that the invalid sampling methods undertaken by the researchers may undermine the credibility of their research outcomes.

④ Biased Approaches to Semi-Structured Interviews

The authors of the journal paper state that for the type of interview questions, 'the interviews were conducted using semi-structured interview schedules' (Lindsay & Muijs, 2006, p. 317). In general, unlike a structured type, semi-structured questions do not provide interviewees with specific choices designed to be selected by them, but instead, the questions have individual objectives and are open-ended, so that subjects may respond to them in a rather independent fashion. For this reason, such type of questions, as the most common type of interview questions, are designed for eliciting reasonably objective viewpoints from interviewees. For interviews, however, the researchers selected only head-teachers, teachers, and pupils, excluding parents, administrators, and boys from other ethnic groups.

The authors also tried to conduct a thematic analysis of interview data. In the paper,

however, they do not show an actual set of questions used for the semi-structured interviews, and so the readers cannot see how appropriately the questions are designed to elicit objective responses. It is widely held that interview questions should be worded so as not to lead subjects to any intended answers. In examining the positive characteristics of schools, interviewers were deemed to have made the interview easier and were more likely to entice subjects to give biased responses without contemplating the reasons for both positive and negative elements involved. Accordingly, the researchers should have chosen well-balanced interviewees, and carried out an in-depth analysis of the interview data.

III. CONCLUSION

Given the foregoing considerations, the reported research identifies a number of weaknesses arising from validity and reliability requirements in educational research. To conclude this critical review, I have made some general suggestions about how this problematic research can be improved as follows:

1. Overall Research Designs

First of all, to make the reported study valid and reliable, the researchers should have determined 'the independent variables which are not correlated much with one another, but are highly correlated with the dependent variable'. In general, predictions come to be highly accurate if more than one predictor (independent) variable is employed. Multilevel modeling was used in the research to identify the schools whose disadvantaged groups of boys gained certain levels of academic results, which were considered to be made possible by their specific educational strategies. In doing this, the five independent variables were combined and analyzed as control measures against an outcome measure (dependent variable). The critical problem here is that the predictor variables were highly correlated with one another, so that their predictive power was substantially reduced.

This structural defect is so far-reaching that the subsequent qualitative analyses of interview data may lose their valid foundation automatically. In general, direct manipulation of the independent variable (e.g., a particular teaching method or teaching tool) is one of the important characteristics of experimental research. So, any subsequent inquiry into a similar topic to that of the reported research is recommended to step out of this non-experimental design and to carry out such a well-prepared experimental research in which multiple independent variables precede dependent variables.

In summary, the first step to establish valid criteria for identifying successful schools had to be focused on exploring a set of internal and external factors which are correlated to

the relative underachievement of black African, black Caribbean, and white UK-origin boys, and then, based on those factors, the second step should have been centered on generalizing a range of strategies carried out to assist the pupils in overcoming their underperformance. Finally, the researchers should have identified certain schools as successful ones after their strategies of which were verified to be statistically significant in bucking the trend of underachievement.

2. Specific Research Procedures

Above all, for valid statistical treatment through probability sampling procedures, the researchers should have obtained the significant number of subjects and schools randomly according to the '30/30' or '20/50 selection requirements. In this research, at least 30 individual subjects in each of at least 30 primary and 30 secondary schools were required as the significant number of 30/30 law in order to gain the power of 0.90, or of 20 subjects/50 groups for the identification of cross-level interactions in multilevel modeling.

In this connection, it is widely held that multilevel modeling also has inherent limitations just as the other statistical models have. Pollack (1998) states that (1) this modeling technique is not appropriate for research studies with relatively small number of samples; (2) it is not recommended in the case of small differences between the group variances; and (3) it cannot indicate significant values as to whether predictor variables at level 1 influence those at level 2, as it is designed primarily to observe the effects that predictor variables at level 2 influence such level 1 variables within the frameworks of level 2.

Further, the researchers were required to provide any outcome of statistical data analyses which addresses the comparison of the academic performance gaps between key stage 1 and key stage 2 results in primary schools, and between key stage 2 results (instead of key stage 3 results) and actual GCSE scores in secondary schools. In this regard, I have pointed out earlier that the total GCSE points score used in the research might not be the valid measure of successful schools because some of the nine (9) secondary schools selected as a parent group may have had their pupils take GNVQs instead of GCSEs to improve their school ranks in published 'league tables'. The GNVQs may be used as an alternative to GCSEs in compiling league tables, and the GNVQ examinations as vocational qualifications are known to be easier than those of conventional GCSEs.

In addition, regarding the semi-structured interview schedules, the researchers seem not to have provided well-defined questions used for general guidelines for participants to follow. Moreover, the participants had to respond to the reasons for the schools' positive outcomes attained from special efforts to overcome underachievement problems. To work out valid qualitative analyses of their interview data, the researchers should have taken

three essential steps addressing both positive and negative outcomes in organizing, summarizing, and interpreting the data.

3. Suggestions for Future Research

As discussed earlier, the overall crucial weaknesses found in the published paper in question stem from invalidity in selecting predictor and criterion variables, and in determining criteria for successful schools and non-successful schools. These fundamental design problems caused the reported research to scratch the surface of the effective characteristics that enabled the three ethnic groups of boys to overcome their relative underachievement. Such methodological issues are widely recognized as far-reaching criteria for leading relevant tasks to a success not merely in practical research designs but also in approaches to language teaching and testing of language skills (Seon-Yoo Hwang, & Myong-Kwan Lee, 2009; Min-Jong Song, 2007). In light of this, a future research on a topic similar to that of the paper may be designed to explore the interactions among teacher discrimination, public regard, and ethnic centrality on academic achievement for each group of black African and black Caribbean boys in England in terms of widely recognized approaches.

Therefore, with due weight on the widely known causal relationships of ethnicity and underachievement (Thomas et al., 2009), I would suggest the five procedures addressing correlation and hierarchical regression to tackle the key issues, including predictor variable (PV) and criterion variables (CV), found in the reported research as follows:

Procedure 1: Means, standard deviations, and alpha reliability for predictor, criterion, socio-demographic, and background variables in research; namely, (a-1) key stage 1 results – PV and (a-2) key stage 2 results – CV in case of primary schools, (b-1) key stage 2 results – PV and (b-2) actual GCSE scores – CV in case of secondary schools, (c) racial identity, (d) public regard, (e) perceived teacher discrimination, (f) mother's level of education;

Procedure 2: Correlations among predictor, criterion, socio-demographic, and background variables described in items (a-1) to (f) above *among the total sample*;

Procedure 3: Correlations among predictor, criterion, socio-demographic, and background variables described in items (a-1) to (f) above *among black African boys, among black Caribbean boys, and among white UK-born boys, who show underachievement*;

Procedure 4: Hierarchical regression of (a) racial identity, (b) public regard, (c) perceived teacher discrimination, (d) mother's level of education variables,

and (e) their interactions on academic achievement *for the total sample*; and

Procedure 5: Hierarchical regression of (a) racial identity, (b) public regard, (c) perceived teacher discrimination, (d) mother's level of education variables, and (e) their interactions on academic achievement *for black African boys, for black Caribbean boys, for white UK-born boys, who show underachievement.*

REFERENCES

BBC NEWS. (2007). *Four GCSEs from single exam*. Retrieved February 11, 2007, from the World Wide Web: http://news.bbc.co.uk/2/hi/uk_news/ education/default. stm.

Best, J. W., & Kahn, J. V. (2006). *Research in education (3rd ed.)*. New York: Pearson Education .

Hwang, Seon-Yoo, & Lee, Myong-Kwan. (2009). Does test taking experience make a difference?: In strategy use and test scores. *Modern English Education, 10*(2), 23-43.

Kreft, I., & De Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage.

Kubiszyn, T., & Borich, G. (2003). *Educational testing and measurement: Classroom application and practice (7th ed.)*. New York: John Wiley & Sons.

Lindsay, G., & Muijs, D. (2006). Challenging underachievement in boys. *Educational Research*, *48*(3), 313-332.

McMillan, J. H. (2004). *Educational research: Fundamentals for the consumer (4th ed.)*. New York: Pearson Education.

Pollack, B. N. (1998). Hierarchical linear modeling and the 'Unit of Analysis' problem: A solution for analyzing responses of intact group members. *Group Dynamics: Theory, Research, and Practice*, *2*, 299-312.

Reynolds, C. R., Livingston, R. B., & Willson, V. (2006). *Measurement and assessment in education*. Upper Saddle River, NJ: Pearson Education.

Song, Min-Jong. (2007). Teaching reading with an interactive approach: Using dialogue journals as a practice. *Modern English Education, 8*(3), 220-242.

Thomas, O. N., Caldwell, C. H., Faison, N., & Jackson, J. S. (2009). Promoting academic achievement: The role of racial identity in buffering perceptions of teacher discrimination on academic achievement among African American and Caribbean Black adolescents. *Journal of Educational Psychology*, *101*(2), 420-431.

APPENDIX
Challenging underachievement in boys

Authors: Geoff Lindsay, & ;Daniel Muijs; **Affiliation:** University of Warwick, UK
DOI: 10.1080/00131880600992389; **Publication Frequency:** 4 issues per year
Published in: Educational Research, Volume 48, Issue 3 November 2006, pages 313 – 332;
To cite this Article: Lindsay, Geoff and Muijs, Daniel 'Challenging underachievement in boys', Educational Research, 48:3, 313 – 332

Abstract: *Background* Underachievement among boys and particular ethnic groups is a major challenge to the education system. *Purpose* This paper reports on one local education authority's attempts to address this concern by investigating the characteristics of its schools demonstrating success with the groups causing most concern: black Caribbean, black African and white UK-born boys. *Sample, design and methods* Multi-level modelling was used to identify three primary and three secondary schools producing results above expectation. Visits were made to each school to explore the reasons for success by interviews with headteachers, teachers and pupils. *Results* A number of factors were identified, which were similar to earlier research on effective schools. *Conclusions* Schools generally subscribed to an approach that stressed inclusivity and overall school effectiveness, rather than the targeting of specific groups, but the nature of schools' approaches was influenced by the relative proportions of boys from different ethnic heritages.

Keywords: Boys' underachievement; Ethnicity; School effectiveness

Background: Differential levels of achievement between various pupil groups have been a cause of concern in the UK for many years. In recent years the focus has shifted from social class, where working-class children have lower levels of performance than those from the middle class (Rutter & Madge, 1976; Essen & Wedge, 1982), to gender and ethnicity. The present paper reports on the attempt of one local education authority (LEA) to challenge underachievement in boys. As two of the three groups of boys identified as underachieving were from minority ethnic backgrounds, this initiative addressed both gender and ethnicity issues and the interaction between the two. The focus of gender concern has switched from girls to boys during the recent past. Previously, the main concern was the underperformance of girls. For example, in the case of 16-year-old pupils taking public examinations in the UK, the percentage of girls achieving good grades was generally lower than that for boys in the UK General Certificate of Education (GCE) and its replacement the General Certificate of Secondary Education (GCSE), and in the General National Vocational Qualification (GNVQ), designed to be an equivalent examination to GCSE but to access vocational rather than academic programmes. This trend continued at the Advanced-level (A-level) examination, and the percentage of girls staying on in full-time education until the age of 18 also decreased differentially compared with boys (Battacharyya *et al*., 2003). In an attempt to counteract this imbalance, a number of initiatives were developed including Girls into Science and Technology (Kelly *et al*., 1984). The focus has changed since the early 1990s as concern increased about the relative underperformance of boys (Mortimore *et al*., 1988). The debate hit national headlines and seized government attention with particular reference to reading standards (Turner, 1990). However, more analytical investigations indicated that the main change was the presence of a 'long tail', a higher than expected proportion of very poor readers in the 7-8 years age range, comprising mainly boys (e.g. Cato &

Whetton, 1991; Lake, 1992). Department for Education and Skills statistics for 1992/93 through to 2002/03, indicate a consistent superiority of girls over boys. By 2002/03, 58.2% of girls achieved five or more GCSE grades A*-C, or GNVQ equivalent, compared with 47.9% of boys. Also, the achievement gap is widening: since 1992/93 the percentage of girls achieving 5 or more good GCSEs has increased by 12.4 points, while the comparable percentage increase for boys has been 11.1 points (http://www.dfes.gov.uk/trends/ index) (see Connolly, 2006, for a critique of this measure of difference).

The relative success of children from different ethnic backgrounds also became a matter of heightened concern in the UK during the 1980s and 1990s, reflecting similar concerns in the USA (Coleman *et al*., 1966; Jencks *et al*., 1972). Government committees were set up to investigate the reasons for the differential performance: in the Rampton Report (1981) and the Swann Report (1985). These confirmed the presence of differential outcomes but the Swann Report, in particular, identified the need to recognize differences within broad ethnic groups. It has also become apparent that there are significant gender-ethnicity interaction effects. For example, Battacharyya *et al*. (2003) reported more black Caribbean girls achieving at a higher level than black Caribbean boys. Recent reviews on behalf of Ofsted have provided comprehensive evidence for both underachievement, and the interaction between ethnicity and gender (Gillborn & Gipps, 1996; Gillborn & Mirza, 2000). Entering the twenty-first century, these differences between ethnic groups have continued. Differences in levels of achievement between ethnic groups for the national tests for 7-year-olds in England (key stage 1 Reading and Writing), 11-year-olds (key stage 2 English), 14-year-olds (key stage 3 English) and 16-year-olds (GCSE English) have remained relatively constant, apart from African-Caribbean pupils who were the only ethnic category to demonstrate a substantial change in relative position: a fall from 5th out of 8 to 8th position (Battacharyya *et al*., 2003). Also, the spread of percentages of pupils achieving expected levels increased over these four assessments. While over 75% of African-Caribbean pupils achieved the expected level at key stage 1, this reduced to just over 40% at GCSE, a reduction of 35 percentage points compared with a reduction of only 25 percentage points for white pupils and fewer than 20 percentage points for children of Chinese heritage. (. . .)

In response to these persistent differences in attainment between ethnic groups, one line of research has attempted to investigate the reasons for such differences by gender and ethnicity. The former has often been conceptualized as a problem of disaffection, while investigations of ethnicity have additionally included hypotheses regarding both the lower relative intellectual ability of some ethnic groups and the negative effects of racism. Both have been investigated with reference to their interaction with social disadvantage, which has continued to demonstrate a substantial correlation with lower academic achievement (Mortimore & Whitty, 1997; Lindsay *et al*., 2006). However, such studies have suffered from contested definitions of race and ethnicity (Phinney, 1996), intelligence (Neisser *et al*., 1996) and of underachievement. The definition of underachievement is problematic as it may be used to denote differences in absolute levels, or in levels predicted on the basis of certain variables, while partialling out others; for example, the latter approach typically partials out social disadvantage and/or prior achievement as a known correlate with lower achievement. Such approaches require more sophisticated statistical techniques, but also raise questions regarding the appropriateness of the modelling approach (Gorard & Smith, 2004). Another approach has focused on identifying interventions that have shown success in challenging the underachievement shown by boys and certain minority ethnic groups, in particular, black Caribbean and some black African children. For example, underachievement in boys has been addressed by focusing on improving early

literacy development to prevent the high levels of reading failure among boys, particularly those from disadvantaged backgrounds. To challenge underachievement of certain minority ethnic pupils the Department for Education and Skills (DfES, 2003) has produced guidance for schools with an explicit message in its title, *Aiming high: raising the achievement of African-Caribbean pupils*, following an earlier report from Ofsted (OHMCI, 2001). Evidence suggests there may need to be different approaches from different schools as pupils of black Caribbean, Indian and Pakistani backgrounds in mainly white schools have been found to outperform similar pupils in urban multi-ethnic schools at 16 years (GCSE level), although not at 11 years (key stage 2) (Cline *et al*., 2002). The present paper reports on the initiative of one LEA in east London. The LEA covers an area of low socio-economic status where the population had changed from predominantly white to one of increasing ethnic diversity. Following a change in local government structure with the installation of an elected mayor, the local authority had set up a Scrutiny Commission to oversee educational provision provided by its schools and LEA. The Scrutiny Commission had reviewed the achievement of pupils and identified the underachievement of three groups of boys as particularly problematic: black Caribbean, black African and white UK-born boys (the authority's descriptions). The Scrutiny Commission sought to identify good practice that might improve the academic achievement of these groups.

Method: The present study was commissioned by the LEA's Scrutiny Commission. However, rather than commissioning a university research centre to undertake research and report, the present project was envisaged as a joint activity. We acted as advisers on the research design, development of instruments and took responsibility for data collection and analysis. Members of the Scrutiny Commission were actively engaged in discussions regarding design and development work, and also undertook some data collection, as well as receiving our final report. This, in turn, was incorporated into an authority document, in consultation, which provided the background to a day conference for the LEA's staff and special interest groups. The research model adopted may be considered an example of democratic action research. This approach has been criticized by Atkinson and Delamont (1985) and Adelman (1989), for example, because non-researchers lack the skills and knowledge of trained researchers. These criticisms are addressed by this variant as the outside experts (ourselves) had the major role in design and implementation. The democratic element comprised the active interchange between ourselves and the Scrutiny Commission and also their engagement with the research process by data collection. The action element relates also to the use made of the findings to effect change.

Procedure: A two-stage procedure was adopted. Stage 1 comprised the identification of primary and secondary schools for which there was evidence that they were bucking the trend of underachievement in the three target groups of boys. Multi-level modelling was undertaken of the authority's data set of achievement outcomes. Separate analyses were undertaken for boys in each ethnic group (black Caribbean, black African, and white) and for primary and secondary schools. In each case, the analysis compared the outcome results with those expected on the basis of a predictor model. The model took into account key variables known to have an impact on attainment, namely: prior achievement; eligibility for free school meals as a measure of disadvantage; stage on the SEN Code of Practice as a measure of presence and degree of SEN; proficiency in English; and term of birth. On the basis of these analyses, three primary and three secondary schools were selected as having evidence of 'adding value' for one or more groups of boys. These formed the sample of

schools visited in stage 2 of the procedure, when interviews were undertaken with the headteacher, a sample of teachers (between three and six) and a sample of pupils (between six and 12) in each school. The interviews were conducted using semi-structured interview schedules. Visits were made, in some cases, as a group comprising one of the authors together with two members of the Scrutiny Commission and one or two officers of the authority, who provided administrative support to the Scrutiny Commission. All interviews with children were conducted by a researcher alone. In all cases, interviewees were informed of the purpose of the research and assured that no individual would be identified. However, given the small number of schools visited, the headteachers and other adults were informed that it would not be possible to guarantee anonymity. All were content with this, partly because the study focused on the reasons for positive outcomes. Interviews with headteachers typically lasted approximately 45 minutes. Interviews with teachers were typically shorter, while group interviews with pupils typically lasted 15-20 minutes. Interviews were recorded and detailed fieldnotes were made during each interview, and written up immediately afterwards to ensure maximum comprehensiveness and clarity of the record. A thematic analysis was then undertaken of interview schedules by each author. By process of iteration key themes were identified, together with sub-themes.

Results: Identification of schools achieving results above expectation In stage 1 the outcome measures for primary schools were the key stage 2 results (English, Maths and Science) at 11 years. Data taken into account in examining the key variables known to have an impact on attainment were the key stage 1 results (at 7 years), as the measure of prior achievement, free school meal eligibility, Special Educational Needs Code of Practice stage, English proficiency and term of birth. For secondary schools the outcome measure was the total GCSE points score (at 16 years) and the control measures were as for the primaries, except for the use of key stage 3 (at 14 years) rather than key stage 1 test scores as measures of prior achievement. This approach allowed the identification of schools which were performing better than predicted for one or more of the three groups of boys having taken into account these other factors, the most important of which was the pupils' prior attainment. In order to check stability of school influence, analyses were carried out on two years of data, 2001 and 2002. Figures 1 and 2 provide examples of this method. Figure 1 shows the results for each primary school ($N = 58$) for KS2 English in 2001 for black African boys after the other factors have been taken into account. The graph shows the school-level differences from expected performance for black African boys sorted in ascending order. That is, it provides a display of the relative 'value added' by each school. The triangles represent each school's mean (average) score for English. The vertical bars represent the 95% confidence bands and hence statistical significance. At the right of the graph there are three schools where the scores and confidence limits are entirely above the central horizontal line of prediction. These are the three schools that can be concluded to be adding value for black African boys. Figure 2 presents a similar graph for secondary schools for GCSE results of white UK-born boys in 2001. Again, the graph shows school-level differences between the nine secondary schools, with 95% confidence intervals and sorted in ascending order. Two schools in this case are 'adding value'. Results for the black Caribbean and black African boys indicated even greater degrees of 'value added' for the school with the highest mean score in Figure 2, that on the extreme right. This suggests that this school's influence extended across all three of the groups identified for study.

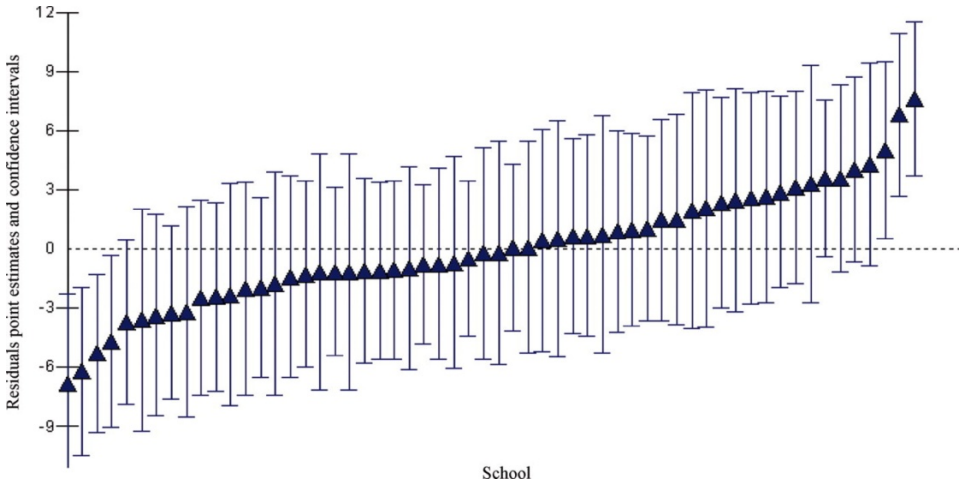

Figure 1. Value added, by primary schools, for 10-11-year-old black African boys in key stage 2 English test.

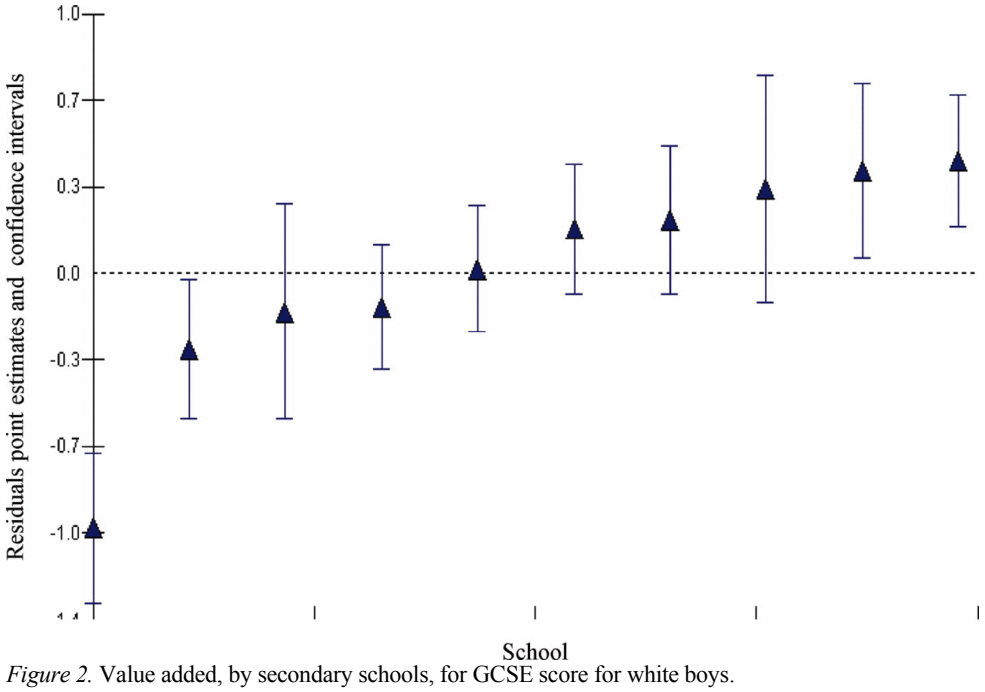

Figure 2. Value added, by secondary schools, for GCSE score for white boys.

A limitation of this approach is that it requires a reasonable number of pupils in each school for the particular group under consideration, otherwise the results are unstable and unreliable. This was a problem at primary school level where it was not possible to conduct this analysis for black Caribbean boys, as the numbers in many schools were too low. Consequently, two analyses of primary schools were conducted: for black African boys alone and also for black Caribbean and black African boys combined. At secondary school level it was possible to analyse the two groups separately, as well as white UK-born boys, as a group.

Findings from interview data: The results of these analyses provided a number of primary schools that appeared to be 'adding value' for one or more groups. From these were selected schools that were stable across the two years and the three curriculum subjects. A similar process was conducted for secondary schools. It should be noted that the primaries were a selection from a larger group that showed value added for one or more groups of boys, but these results were often variable across the three curriculum areas and across the two years. At secondary school level, the sample was drawn from one school that stood out consistently and a small number of other schools that also had positive indicators for one or more groups in at least one year. Three primary and three secondary schools were selected for visits. At stage 2 the thematic analysis of data from interviews with headteachers, teachers and pupils produced one superordinate and eight major themes concerning the practice of these schools in bucking the trend of underachievement among the three identified groups of boys.

No single approach: A key finding is that there was no single approach taken by the successful schools visited. Rather, the schools differed in their approaches to these underachieving groups. One approach emphasized the insistence on (equally) high standards for all groups and not targeting any one specifically. The other was specifically to target specific underperforming groups. The first approach is exemplified by a school that stressed high expectations and academic achievement for all, without providing any specific support for, or groupings of, particular ethnic/gender groups. The ethos of achievement of the school permeated the whole institution and was tangible to visitors even during a short visit to the school. Pupils here specifically stated that there was 'no discrimination' in the school. The second approach was exemplified in one school by the provision of specific support to boys identified as 'at risk', by putting them into one class without girls present. The challenge of teaching these groups was impressively met during our visit, and the boys appeared to benefit from this approach, where they were provided with a fast tempo, group work, many different activities and competition. This superordinate theme is important to bear in mind when considering the following six more specific factors which have been identified from the thematic analysis of the interviews as being helpful in overcoming relative underachievement in the three groups that are the focus of the study: white boys of UK origin, black Caribbean and black African boys. These factors comprise curricular approaches, performance monitoring, high expectations, staffing, inclusive ethos and parent involvement. (. . .)

Discussion: The present study addressed the important issue of the underachievement of groups of boys, selected by one LEA's Scrutiny Commission, causing concern locally, namely black Caribbean, black African and white UK-born boys. The demography of this local authority resulted in a limited

social class range, so it was not easy to identify social class factors as independent or interacting variables. Nevertheless, it is necessary to acknowledge the relationship between overall achievement and social factors (Bynner & Joshi, 2002), and also the interaction between ethnicity, social factors and special educational needs (Lindsay *et al*., 2006). Although our focus was on success, it is necessary to acknowledge this was relative among the schools in one of the country's most disadvantaged authorities. The major themes identified as influential in the schools' success include ethos, curriculum organization, teaching approach and the collaboration with parents and the community. Several themes resonate with earlier studies (e.g. Office of Her Majesty's Chief Inspector of Schools, 1999; 2002a, 2002b), and are also supported in the wider literature on effective schools (e.g. Teddlie & Reynolds, 2000). Similar findings regarding key factors influencing achievement of Black Caribbean pupils, in particular, have recently been reported in a study of another London authority, Lambeth (Demie, 2005).

Ethos has long been considered a key element of school effectiveness and school improvement, related to factors as variable as pupil behaviour, achievement and attitudes, teacher job satisfaction, teacher retention and bullying (Teddlie & Reynolds, 2000). Definitions of ethos are varied, but one that encompasses this range of factors succinctly is: 'the ethos of the school is the set of values, attitudes and behaviours which will become characteristics of the school as a whole' (Rutter *et al*., 1979). In a school, ethos touches on all aspects of its operation but its very pervasiveness means that it is hard to pin down—it can be part of the 'taken-for-granted' aspects of school life. Ethos in this definition is a characteristic of the organization and not of individuals, although all contribute to forming, and are in turn influenced by, the organization's ethos. It is not fixed, and, in this study, schools influenced ethos by addressing teachers' perceptions of traditionally underachieving groups and these pupils' own views of school and learning. The importance of a focus on academic achievement and learning from school leadership downwards (e.g. Creemers, 1994; Muijs *et al*., 2004) was also present. We found a particular emphasis on the importance of making connections with community cultures, also reported by Wrigley (2003) in his study of multi-ethnic schools in Scotland. The combination of an academic 'press' by schools, placing a strong and explicit emphasis on expectation of high achievement and their taking on community cultures, indicates the necessity of a more adaptive mode: to be effective with underperforming communities, schools need to find a balance between socializing students to traditional school norms, while simultaneously re-socializing the school and its staff towards community values. The latter has often been neglected in traditional school effectiveness research.

This issue is clearly connected with the important role of the pupils' carers, and their involvement in the school, factors that have long been considered important to school effectiveness and school improvement, particularly for schools serving disadvantaged communities. Joyce *et al*. (1999) suggest that for true school improvement to occur, schools must become communities, involving parents and local businesses as well as teachers and heads. Parental involvement and community outreach have been found to be important factors in improving schools in a wide range of studies both in the USA (Seeley *et al*., 1990) and Europe (Maden & Hillman, 1993). Achieving this has proved difficult, however, with many schools finding difficulty in engaging parents, many of whom

had negative experiences of schooling, or perceive school staff as culturally alien (Maden & Hillman, 1993). Attempts to address these issues have either included incentives, such as providing transport, childcare or even in one highly impoverished district laundry facilities (Guthrie *et al*., 1989; Leithwood & Steinbach, 2002), or by using more coercive means. For example, one of the sample schools gave parents the impression that some forms of involvement, such as parents' evenings, were compulsory. In another school, participation seemed high, at least in part because the school's high achievement and singular ethos appeared to engender a pride in the school among the community. Again, it may be that both strong academic focus and a greater awareness of and responsiveness to community cultures might be necessary to attain this elusive goal. A key question regarding teaching and learning concerns the extent to which curriculum and teaching require differentiation for different pupil groups and, in particular, whether or not a more vocational or basic skills orientation may be appropriate for pupils from disadvantaged backgrounds (Muijs *et al*., 2004). These schools tended towards greater use of vocational qualifications as a way of raising both motivation and achievement, particularly for boys. While, as in some other studies (e.g. Teddlie & Springfield, 1993), this was found to have positive effects, if taken too far such strategies can discriminate against certain groups, creating new forms of race- and class-based disadvantage or, as in this instance, gender streaming (Hansell & Karweit, 1983).

It is also important to note the lack of a single general model adopted in these schools. We have distinguished two general approaches: an 'effective school' approach, where the focus is optimizing attainment for all pupils, and so is 'ethnicity-free', and 'gender-free', in the sense that specific targeting of boys or minority ethnic groups is downplayed or eschewed. These are problematic concepts that, if treated simplistically, underplay social variables. 'Ethnicity-free' may be compared with the 'colour-blind' approach that Demie (2005, p. 482) has termed one of the 'biggest obstacles to raising Caribbean achievement … which has put the group at a disadvantage'. However, an approach that seeks to improve the attainments of all rather than targeting has support from a study by Cline *et al*. (2002) that showed that pupils from black Caribbean, Indian and Pakistani backgrounds in mainly white schools outperformed similar pupils in urban, multi-ethnic schools, suggesting they shared in the educational advantages available generally within those schools. The second approach does target specific groups for particular attention. Overall, the former was the more common in this study, but across the schools there were varying examples and degrees of targeting. An important question is the extent to which the adoption of the model by any school was specifically determined, or was a general philosophy modified by practical issues requiring an expedient response. This issue can be exemplified by the resistance to targeting generally, but support for targeting literacy development during the early years of the school's phase; or by the resistance to analyse data by ethnicity because of a concern about discriminatory practice, but ease with analysing by gender. Blair and Bourne (1998) indicate the dangers in focusing on the majority group to the detriment of minorities, but these schools' approaches appeared to avoid this by an inclusive, comprehensive orientation. The approach taken also appeared to be shaped by pupil profile. This varied across the six schools, some having high proportions of one or more ethnic group while others had pupils with a large number of different heritages, each comprising small numbers. The former allows targeting in a way that the latter does not.

A second contextual factor concerns change. The LEA's schools had changed from mostly comprising white pupils, as a result of a substantial influx of families from a wide range of ethnic minority heritages. Change could be very quick: one primary school reported a reduction in its traditional white intake from about 70% to under 30% over three years. As the headteacher noted, 'that kind of demographic change is a catalyst for school improvement'. In this school, the strong inclusive and effective school approaches appeared particularly suited to address the changes in pupil composition. This study provides evidence for schools' success in overcoming underachievement in particular groups of boys. The project was a successful example of democratic action research. Although this approach has been criticized (e.g. Hammersley, 2004), especially when undertaken by novices with little support, the project combined the research expertise of researchers with the local knowledge and needs of key local sponsors, the LEA's Scrutiny Commission. The research incorporated both quantitative methods using multi-level modelling to identify schools, and qualitative methods to explore interviewees' views of influences and causality. The use of 'combined methods' is becoming more popular within educational research, partly following criticism of its quality and impact (Hillage *et al*., 1998; Tooley & Darby, 1998). This small-scale study is limited in scope, with only six schools in the sample. However, it does point to some interesting areas for further study with a larger sample of successful schools, and larger numbers of interviews with teachers and pupils, in an area which is becoming ever more salient as schools struggle to work within fast-changing contexts; and in which they have to learn to work with new communities, while not leaving traditional groups behind.

References (For details, see *Educational Research*, *48*(3), 313 – 332)

Jong-Hee Lee
Department of English Language and Cultural Studies
College of Humanities and Social Sciences
Kangwon National University - Samcheok Campus
1 Joongang-ro, Samcheok-si, Kangwon-do, 245-711, Korea
Tel: (033) 570-6652
Email: freshfields@kangwon.ac.kr

Received 1 June 2010
Revised 31 July 2010
Accepted 9 August 2010