



Professional Development and Student Achievement Gains: A Look at the *Te Kōtahitanga* Report

New Zealand Journal of Teachers' Work, Volume 2, Issue 1, 20-27, 2005

ROY NASH

Massey University

INTRODUCTION

The publication of *Te Kōtahitanga: The experiences of year 9 and 10 Māori students in mainstream classrooms* by Russell Bishop and his colleagues at the University of Waikato has not gone unnoticed (Bishop, Berryman, Tiakiwai, & Richardson, 2003). The Ministry of Education has made its findings widely available to teachers – there is even a video – and to a receptive news media. Those attending the well-publicised biennial conference run by the Teacher Education Forum for Aotearoa New Zealand (TEFANZ), for example, were informed by the Minister of Education that, '[by] helping teachers confront their deficit thinking and making a serious attempt to understand and relate to students as Māori, and plan programmes accordingly, Professor Bishop has helped teachers make extraordinary differences to students' outcomes' (Mallard, 2004). That professional development programmes can persuade teachers to raise their expectations for students and enable them to transform their classroom practices so that their Māori students show a dramatic increase in achievement must seem good news. Were it so, it would be good news, too. The implication is that for the cost of a few million dollars spent on professional development the Māori – non-Māori disparity in educational achievement could be eliminated. Those of us who would greatly like to see that happen should, therefore, take a look at this research by Bishop and his colleagues.

The claim that student achievement has been raised as a result of a professional development initiative is crucial. This finding is actually but a small part of the Bishop *et al.* research programme, and this commentary does not examine their discussion of student narratives or their analysis of classroom observation and school attendance data. There is a limit to what can be done in a critique of acceptable length, and as the strong professional interest in the Bishop report is undoubtedly stimulated by the claim that participating teachers have generated substantial achievement gains, there is good reason to concentrate on that.

ELEVEN TEACHERS: THE EVIDENCE FOR ACHIEVEMENT GAINS

The research involved 11 teachers in four schools. It should be a simple matter to examine the pre-test and post-test scores of these teachers, compare them with those of non-participating students, and report the statistical significance of the mean differences observed. If the tests are nationally standardised, as we should expect, then the information should be particularly

robust. Of course, there would still be the selection bias problem, for to ensure that teachers who volunteer for such projects are not just those whose students tend to make more progress than others, is always a problem. The small sample of teachers would also raise questions, and doubts about halo or Hawthorne effects would need to be overcome, but significant gains in achievement would be indicative, at least, and merit serious consideration. The reported achievement differences associated with involvement in the project are, then, the focus of our attention. The task of reporting these findings, however, is nowhere near as straightforward as it should be. There are no tests before the intervention; researcher-designed tests at the beginning and near the end of the project were administered in only two of the schools; and the analyses of achievement progress are generally made on the basis of in-school assessments the exact nature of which is only loosely described. The analyses of changes in the relative assessments of students made at different occasions are made sometimes with raw data, sometimes in percentages, and sometimes in 'value added move' deciles (relative change between scores is measured in tenths). This does not make the task of exposition at all easy. There is, in fact, little else to do but list the teachers involved and present in summary the nature of the evidence for their effectiveness in raising student achievement:

Teacher 1. The researchers used an 'added value move' calculated by comparing school entry on non-standardised assessments with credits awarded for, among other things, 'doing homework regularly, participation in sporting and cultural activities, attendance, cooperation and participation in class' (p. 152). The data are reported in the form of 'decile movements' and, particularly considering the extra-academic component, are impossible to interpret. No statistical significance level is reported.

Teacher 2. The same 'added value move' is used as for Teacher 1, but in an aggregate form making it even less intelligible. The authors acknowledge that in this class the 'achievement gain is not as clear', and note that 'the Māori students in the 9H have not reduced their academic performance as much as have other Year 9 Māori students' (p. 160). In other words, the students in this class declined when compared with their intake assessment level.

Teachers 3, 4 and 5 are difficult to separate. They worked at the same school, and taught a Year 10 class of 24 Māori and 4 non-Māori students created for the purposes of the research. The report dedicates many pages to the achievements of class 4XT, mostly in English, and the authors must consider these findings among their most important. The school was streamed by ability and it seems that on non-standardised school-produced tests, the achievements of 4XT were generally similar to those of two 'high stream' classes and to those of at least one of the 'academic' examination-based classes. It is also noted that in Social Studies the target class achieved around 48-49% compared with the 45% expected of a B stream class, but this is not a convincing difference. The interpretation of the English test results, unfortunately, is obscure as it is impossible to know what standard the 'high' and 'academic' streams had actually achieved. If 4XT had gained scores around the mean on nationally

standardised tests one might be impressed, but as things are, one can only regret that what might have been an opportunity to demonstrate the effects of a successful intervention was lost.

Teacher 6. The authors state that there are no appropriate data for this class.

Teacher 7. Differences are reported between Māori and non-Māori students on a 20 item cloze test and a 10 minute writing exercise carried out at two different times. The data are unconvincing: it seems, for example, that 10 Māori students got 10 words right at Time 1 and 14 at Time 2, compared with 17 non-Māori who got 11 words right at Time 1 and 13 at Time 2. No weight can be put on such results. Note that although the report claims that targeted Māori were compared with non-targeted Māori, this was not done in this case.

Teacher 8. 'It is clear from these data that the achievements of students in 10D were no greater than that which could be expected given their placement in the streamed hierarchy' (p. 183).

Teacher 9. The report attempts to make the best of the data by noting, 'a greater spread of grades in the target class which we would attribute to the improved attendance by some students and a fall off by others' (p. 184), but there is nothing to show that the performance of the students in the class is in any respect distinguishable from those in other classes.

Teacher 10. 'It is clear from these data that the achievements of students in 9D were no greater than that which could be expected given their placement in the streamed hierarchy' (p. 185).

Teacher 11. 'These data indicate that the students in the target class did not improve their achievement level beyond that which the school had predicted when streaming the students' (p. 185).

This evidence does not warrant the conclusion that the academic achievements of targeted Māori students were raised in comparison with others. The status of 4XT taught by teachers 3, 4, and 5, is particularly anomalous. It seems to have been reconstituted in 2002 for the purposes of the research and, as a comparison of Figures 6.13 and 6.14 will show, excluded those students with the worst attendance record in the previous year. Attendance data for 2001 and 2002 are given for only 14 4XT students in that class in both years: their mean unexplained absence rate was 6.5% in 2001 and 3.8% in 2002 (These figures are calculated from information extracted from Figures 6.13 and 6.14.). The difference between these means is statistically non-significant and, contrary to the argument of the Report, there is thus no reason to believe that the research intervention improved the attendance of these 14 students. Inspection of the data shows that for 6 of these 14 students the absence rate was exactly the same, or changed by no more than one percentage point, and the overall improvement is mostly due to students 3 and 5, cut from 8% to zero, and student 25, down from 14% to 7%. The 2001 corresponding mean absence rate

for 13 students *not* included in 4XT in 2002 (but shown as in that class in 2001), can be calculated from Figure 6.13 as 10.5%: the unexplained absence rate for students selected for the research class was thus higher, although not significantly so by the Mann-Whitney U test, than that of those not selected. It is certainly noticeable that the worst offenders, students 4, 20, and 28, whose 2001 unexplained absences ranged from 20-29%, were not included, and the possibility of bias in the selection of students for the target Māori class cannot be dismissed.

The report acknowledges, as we have seen, that a meaningful difference in the attainments of target class students and other Māori students were not observed in School 4. This is attributed to the fact that the four teachers involved, all of them Māori as it happens, were not able to participate fully in the in-school professional development sessions. The school was a six-hour drive from Hamilton; visits proved more difficult to arrange than had been anticipated, and the teachers thus 'missed some co-construction meetings and on-going, in-class support and shadow coaching' (p. 195). The authors are careful not to blame these teachers for the failure of their targeted Māori students to demonstrate achievement gains: 'it must be stressed that these data do not indicate that these teachers were somehow inadequate' (p. 195). The teachers concerned will doubtless be pleased to hear this, but as Bishop *et al.* maintain that the existing level of Māori achievement is a consequence of their teachers' low expectations, the grounds for this exoneration are obscure. Whether the classroom practices of these teachers changed as a result of the four-day professional development *hui* and other contact is not specifically mentioned, but there is no statement to the contrary. If their classrooms did change in the direction favoured by the research workers then the explanation of the target students' unexceptional achievements seems *ad hoc* and unconvincing. It is always possible, and perhaps more plausible, to suppose that the school's assessment procedures were just more objective than those in the other schools. That would not have been difficult.

THE STATUS OF RESEARCH INFORMATION

'It is important to note', the report duly cautions, 'that this study was not of an experimental design and the data need to be read and interpreted as such' (p. 194). This advice is all well and good, but just how is it fitting to interpret data generated by a non-experimental design? As the authors' give little guidance on that point the matter may be worth some reflection. It must depend rather on what the information is used for. A non-experimental study of the efficacy of a drug, for example, with no controls and no significance statistics, would hardly satisfy a regulatory authority. Of course, education is not as important as health; after all, if medical research was done to the standards of educational research our life spans would be a great deal shorter, and students do not actually die of a second-rate education. This may sound out of script, but there must be some reason why educational research is allowed to generate low quality data and it is difficult to think of a more plausible one. Of course, it is not a good enough reason, and should certainly not be countenanced, but such is the way things are. This critique is, in fact, motivated in large measure by the hope that it might improve the quality of educational research.

The point of an experiment is to isolate the mechanism by which an event or a process under investigation is generated. This fundamental procedure of science is familiar to everyone. For example: we have reason to think that low-

dose aspirin is effective against heart attacks and strokes; we divide a suitable sample into two and give half aspirin and half a placebo; we then collect data on the occurrence of heart attacks and strokes and analyse the data using appropriate tests. The method is very powerful because it has the potential to isolate the causal mechanism responsible and give reliable information. If things have been done properly we can put considerable trust on information produced in this way, and it is always possible to re-examine the data and replicate the study. But how should we interpret scientific information that is *not* from an experimental design? We should simply put less trust in it. There are no accepted rules of procedure that allow one to interpret the findings of a non-experimental study. If a research study tells us that a drug was effective on half of the patients who took it, but cannot say whether or not they would have got better anyway, we would be sensible to put it aside. At best, it might be information enough to warrant a properly conducted experimental study, but to initiate a health policy based on such information would be bordering on the criminal.

The authors of the report actually advise that, 'these results should be read very carefully in that this study is of a very small number of teachers over a very limited period of time, and as such these results are really more in the order of hypotheses in need of further testing' (p. 194). But these cautionary remarks are unheeded by the authors themselves and would negate the practical utility of the research were they taken at face value. The conclusion states unequivocally that, '[t]he results of this study show that it is feasible, within a relatively short period of time, to improve Māori students' educational achievement' (p. 198), and this is how the research has been presented by the Ministry of Education and reported by the news media. Such bold and unqualified statements are not limited to the generalisations that might be expected in a summary conclusion. The text asserts of Teacher 1, for example, that 'supporting the teacher by way of professional development had an impact on raising student achievement in this class' (p. 155), when the evidence does not warrant such confidence. Causal assertions of this kind are made with such frequency that it would be tedious to quote the text extensively to substantiate this point. There is certainly no hint in the Minister's speech to the TEFANZ conference that the 'extraordinary differences' are hypotheses in need of further testing, and it would have been more consistent with good scientific practice to conduct an investigation with appropriate rigour in the first instance.

The text is marked by glaring contradictions in this area. When we are told that, 'the lack of control groups must limit the causal effects' (p. 199), the concept of 'cause' seems to be entirely metaphorical. This comment is ill-formulated in any case because whatever happens in the world has a cause and the absence of a control group makes no difference to that, and causal mechanisms are a lot harder to discover when the events to be explained are of the type under investigation and no controls are in place. The authors' own model of what interpretation to give their non-experimental research (with its lack of controls and 'limits to causal effects') is hard to distinguish from that which could legitimately be given to controlled experimental research. Indeed, the statement from a research group that, 'professional development had an impact on raising student achievement', would convey to most readers that this finding had, in fact, the status of reliable scientific knowledge. The fact that it does not is even rather shocking. It is not considered good scientific practice to assert that something is so and let the reader discover by 'careful reading' that

this is an 'interpretation' given to findings from a non-experimental study and 'more in the order of a hypothesis'.

But let us follow the authors' instruction that 'these results should be read very carefully'. Some typical data from School 4, with 4XT as the star performers, are presented in support of the general claim that Māori students involved in the project achieved more than other students. The details have to be worked through. It is reported that 21 target Māori students gained an average of 5 words on the researchers' non-standardised 20 item cloze test at Time 1 and 6 words at Time 2 (Table 6.25). This compares with 8 words and 7 words respectively for 6 non-Māori students. The authors then comment: 'those Māori students who attended regularly enough to complete items at both points had closed the discrepancy between them and non-Māori students by 10%' (p. 182). At Time 1 the mean group difference is 3 words, and at Time 2 it is just 1 word. The sample is so small and the difference so slight that it is not worth the trouble to calculate its significance level, and perhaps because 2 words is 10% of the total number of marks available, the gap is said to have been reduced by that amount. It is surprising the report does not argue, as the difference is down from 3 words to 1 word, that the reduction amounts to 67%. Page after page is devoted to analyses of this kind and conveying a sense, because it has the 'look and feel' of a scientific text, that something important is being presented. But this carefully staged impression is fundamentally misleading. Twenty trivial and non-significant analyses do not sum to one grand significant trend.

DEFICIT THEORISING

Bishop *et al.* are convinced that, 'deficit theorising by teachers ... is the major impediment to Māori students' educational achievements', and assert that, 'all other influences play a subordinate role' (p. 190). In their view, '[t]he major influence on Māori students' educational achievements lies in the minds and actions of their teachers' (p. 198). The evidence of improvement in the achievements of targeted Māori children exposes, they say, the error of 'deficit theorising', because 'students and their families did not change (as those who promote deficit theorising would suggest was necessary)' (p. 199). But there seems to be a logical mistake here. It is one thing to say that, given the existing state of the educational system, the origins of social disparities lie in the home, and another entirely to argue that this requires policies to change family practices. On the contrary, it is essentially Bourdieu's (1974) position that a 'universal pedagogy' might offer a way to interrupt the School's reproductive function and reduce the effects of the home. The logic of this argument is not at all out of the ordinary. Short sightedness, for example, is often inherited, but it is corrected by lenses or by laser surgery, not by genetic modification. It might be argued that the degree of non-corrected myopia in a population is due to lack of access to spectacles or surgery, but that is another matter, and if by some chance there are genetic tendencies towards that condition in certain populations, as there are, then that information is worth having. Genetic modification is not, in fact, ruled out in principle as a means of affecting the emergence of any property developed in accordance with DNA structures, but it is certainly not the only method of intervention. The relationship between what can be known about the origins of social disparities in education and what can be known about how to reduce them is more complex than many commentators imagine. All those models that attempt to partition the variance in educational achievement between 'home' and 'school' factors, for example, have only a

limited value as far as the development of policies aimed at the reduction of social disparities is concerned. It is not at all the case that what is caused by agents at a given site must be remedied by those agents at that site.

Bishop *et al.* are not entirely consistent in their account of how their interventions were effective. Their conclusion hints that things are, after all, a little more complicated than the fulminations against 'deficit theorising' imply:

The problem of a large multivariate study of this sort is that it is very difficult to identify any one specific variable that is the 'Silver Bullet' for change. Indeed, we would suggest that the very multivariate approach would suggest that there is no such single solution, but rather there is a complex set of interacting variables that contributes to change in Māori students' educational achievements and it is more useful to continue to weave together this complexity rather than to tease apart the complexity looking for single or simple solutions. (p. 197)

This requires a little decoding. It might be read as an argument for multivariate statistical analysis – known as 'causal modelling' – but that would be an error. The passage is not about the processes of causality – the determination of *effects* – but about the processes of intervention and change. The report includes no multivariate statistical analysis, indeed, it contains nothing that could really be called statistical analysis of any kind other than the organisation of cross-tabulated data. In this passage, the reference to 'a complex set of interacting variables' is actually to the structures, dispositions, and practices that constitute the social environment and that must be included in any analysis of the mechanisms that generate change. We should talk of things happening not because of 'interacting variables' but as a result of concrete social processes.

If Bishop *et al.* mean that the actions of parents and policymakers in homes, schools, and government offices, have a complex and joint effect on educational achievement and on social disparities in access to education, then they agree, after all, with our family resource framework (O'Neill & Nash, 2005). But it seems difficult to reconcile this belated recognition of complexity at the level of intervention with the assertion that 'deficit theorising by teachers ... is the major impediment to Māori students' educational achievements ... all other influences play a subordinate role'. These authors allow causal complexity at the level of *interruption* but not at the level of *production*. Indeed, so wedded are they to the mono-causal, deficit-minded, teacher agency theory that even when constrained to acknowledge that 'changes to in-school structures (number of teachers, type of classes etc.)' might have made a difference to targeted Māori students, they suggest that these 'were as a result of the changes in classroom relationships and interactions between the targeted teachers and their Māori students.' (p. 199). In that case, what cannot be attributed to 'changes that had occurred in the classrooms of the targeted teachers'? If teachers modify their views on the educability of Māori students and decide to reorganise classes, provide more teachers, allocate more resources, and so on, then what exactly is the mechanism of change?

CONCLUSION

This commentary has expressed serious reservations with the Bishop *et al. Te Kōtahitanga* report as far as its claims to have demonstrated achievement gains are concerned. Mallard's view that Bishop's research 'has shown what a huge difference that can be made for Māori students,' when teachers 'reject the notion that culturally different students have deficits that prevent their achieving', may now be put in context. The rest of the research is, I suggest, at least as problematic as its account of achievement gains, but there is only so much one can do at a time. As there is no good reason to accept that claim it is pointless to engage in a detailed critique of the broader hypothesis that the disparity between Māori and non-Māori in education is fundamentally the consequence of teachers' low expectations for Māori students, their uncritical acceptance of 'deficit theorising', and their reluctance to adjust their pedagogy to recognise Māori culture. These authors put my own research in the 'deficit theory' category, which I reject, although it would require another commentary to respond adequately to that charge. Let it suffice to point out that a family resource framework able to integrate social structure, acquired dispositions, and customary practices, within a numbers and narratives methodology, offers the possibility of a sophisticated model of the complex mechanisms that generate social disparities in educational achievement (Nash, 2003a; 2003b).

The limited purposes of this discussion will have been met if some teachers are enlightened by the knowledge that there may be processes at work to reduce the achievement of Māori students more powerful than teachers' low expectations, and very much more is required to improve Māori achievement than a little professional development from Bishop and his colleagues.

REFERENCES

- Bishop, R., Berryman, M., Tiakiwai, S., & Richardson, C. (2003) *Te kōtahitanga: The experiences of year 9 and 10 Māori students in mainstream classrooms, Report to the Ministry of Education*. Wellington: Ministry of Education.
- Bourdieu, P. (1974) The school as a conservative force: Scholastic and cultural inequalities. In J. Eggleston (Ed.), *Contemporary research in the sociology of education*, trans. J. C. Whitehouse (pp. 32-46). London: Methuen.
- Mallard, T. (2004) Speech to biennial conference for Teacher Education Forum for Aotearoa New Zealand (TEFANZ), Auckland College of Education. email@executive.govt.nz, subject: Hon Trevor Mallard: focusing on quality teaching, 6 July.
- Nash, R. (2003a) Numbers and narratives: Further reflections in the sociology of education. *British Journal of Sociology of Education*, 23(3), 397-412.
- Nash, R. (2003b) Social explanation and socialization: On Bourdieu and the structure, disposition, practice scheme. *Sociological Review*, 51(1), 43-62.
- O'Neill, A-M., & Nash, R. (2005; in press). Family explanations of educational attainment. In P. Adams, R. Openshaw, & J. Hamer (Eds.), *Education and society in Aotearoa New Zealand* (2nd ed.). Melbourne: Thomson/Dunmore.



About the Author(s)

New Zealand Journal of Teachers' Work, Volume 2, Issue 1, 2005

ROY NASH

Massey University

Roy Nash is a professor of education at Massey University College of Education. He is a sociologist of education with a speciality in the explanation of social disparities in educational achievement. Recent papers can be found in *New Zealand Journal of Educational Studies*, *Journal of Curriculum Studies*, *Waikato Journal of Education* and *International Studies in Sociology of Education*.

Professor Roy Nash

Department of Social and Policy Studies in Education

Massey University College of Education

Private Bag 11 222

Palmerston North

R.Nash@xtra.co.nz