

Audit and evaluation of pedagogy: towards a sociocultural perspective

Julian Williams, School of Education, University of Manchester, U.K.

Abstract

In this paper I will (a) outline a perspective on audit and evaluation based on Power's critical perspective, and thus present a sociocultural analysis of the forces at work on audit; (b) draw on some empirical studies done recently with doctoral/post-doctoral students of teachers pedagogical content knowledge, in particular their understanding/knowledge of their students' mathematical knowledge. In summary, these studies show that (i) the teachers we studied sometimes mis-judge their students knowledge; (ii) their judgments are influenced by their own mathematical knowledge, and by their teaching experience; (iii) their knowledge of their students can be strongly 'task-situated' and 'tool-mediated' rather than 'conceptual' and 'in the head'; and (c) argue that this work suggests the need to examine our methodology for studying teacher-knowledge with due recognition of the danger, (or is it an opportunity?) that 'teacher knowledge' is a particular, politically and socially mediated construct of audit methodology. I further propose a sociocultural, cultural-historical activity theoretic (CHAT) view of pedagogy: if pedagogy is seen as a social, cultural and historically situated and distributed activity – in the activity–theoretic sense – then this opens up the field of inquiry/research and intervention in what may be helpful ways.

Key words: audit, evaluation, assessment, teacher' knowledge, sociocultural/cultural-historical activity theory

Introduction

I begin the paper with an account (!) of 'audit' from a critical social point of view, following the work of Power (1997) and Strathern (1997; 2000) among others.

Then I argue that the problem of auditing and evaluating teachers' knowledge for teaching requires us to answer some basic questions, and use the preceding discussion to help problematise these.

- What is the purpose of audit and evaluation of teachers' knowledge?
- What kinds of knowledge do we want mathematics teachers to 'have' or 'display' (of mathematics/children/learning and teaching)?
- How can we audit/assess teacher knowledge: what 'realistic' tools/technologies do we have?

In the latter part I summarise some of our group's research and then to move to a discussion of the relevance of socio-cultural and CHAT perspectives on knowledge and practice. The main citations to the corpus on CHAT are usually Vygotsky, Leont'ev, 1978, 1981; Bakhtin, 1986; and the Western developers Cole, 1996; Wertsch, 1991; 2002; Engestrom, 1987, 1991, and many others. But note that this leaves out many authorities, including Luria, Ilyenkov, Davydov from Russia, and alternative modern interpreters such as Stetsenko. See Roth & Lee, 2007, and Roth et al., 2005, and Williams et al., for reviews.

It is important to mention that this tradition has influenced another significant neo-Marxist current well known to maths educators in the UK, that of situated learning in communities of practice, (Lave, 1996; Lave & Wenger, 1991; Wenger, 1998). However, the CHAT theory has an extensive history and more extended repertoire of concepts; much of which we will be drawing on here. I will argue these perspectives raise new horizons in regard to the questions of audit and evaluation of teachers and teacher knowledge.

Accounting for the dialectic of audit

The critical sociological literature on ‘audit’ suggests that we face an audit explosion in all public sectors of the economy from health and education to policing. We all know how dysfunctional this can be, and the literature is not without passionate, strongly politically -positioned critiques of its often deleterious, often unintended impact on practice. (See several chapters in Strathern, 1997; 2000).

On the other hand auditing practices are no doubt here to stay, and in some senses seem to go from strength to strength: in education, despite powerful critiques, one may conclude it is stronger than ever before. The managerial classes need audit to protect themselves from their own lack of accountability and potential accusations of bad judgment, indeed of having made any personal judgment that can thereby be criticised (Power, 1997).

With TIMSS and PISA, one sees league tables going global: it is not difficult to imagine potentially homogenising effects on education internationally to suit the international labour market, and much EU policy seems directed along these lines. Indeed, the nation state in these circumstances may come to have less and less room for manoeuvre themselves! (Williams, 2005).

Yet in the education literature it seems our theoretical understanding of auditing practices is slim. Michael Power (I believe he is the first person in the UK to call himself a ‘Professor of Critical Accountancy’ but I may be wrong) made an important contribution of the conceptions of ‘de-coupling and colonising’ that I want to introduce here. He argues that ‘audit’ is driven by a degree, perhaps a healthy degree, of ‘mistrust’, and of the need for ‘accountability’ and some degree of transparency of procedure. In particular, he has shown how the tensions involved in audit arise from contradictions between audit from the ‘bottom-up’ and the ‘top-down’, and has used these notions to ground insights into contrasting, empirical case studies, in finance, in health and in Higher Education. My purpose here is to use his approach to re-conceptualise ‘formative’ and ‘summative’ aspects of audit or evaluation, and thus re-inforce the controversial insight that both are necessary components of audit. (See Williams & Ryan, 2000, which builds on Black & Wiliam, 1998a, 1998b).

Power shows that the entire history of audit involves a problematic: the purpose of audit is to reduce the necessity of relying on the validity of local custom and practice, or on subjective judgment. And yet the audit practice is *itself* unauditible, i.e. it relies in the end on the judgment of experienced auditors, and this judgment essentially includes their subjective evaluation (based on small sets of data and impressions) of the people they are auditing. In practice there is a ‘gap’ between what audit promises

and what can realistically be ‘known’ (with finite resources a significant part of the problem).

In addition, the demand for audit creates an insatiable need for new measures, i.e. for measures of primary products of practices. In some areas of education this is relatively unproblematic (exam grades) but how do you measure the value of a university’s CVs, or the effectiveness of the management of a government department? This has led to second order measures, whereby the processes of management are measured instead of their products (TQM). Thus, it seems, auditors can be made happy, can be persuaded perhaps, to use second order measures as long as they can be made credible and can be counted. What auditors need is a politically acceptable system that can be credibly said to hold the system being audited to ‘account’ for its outcomes, systems or costs.

It emerges from Power’s account that auditing *in practice* always needs to engage with its auditees and their practices. Increasingly the auditors expect (and on grounds of efficiency this is inevitable) the auditees to actively ‘comply’ with the audit, and this provides room for manoeuvre if auditees are collect data, even construct measures etc. Indeed, in persuading doctors to collect measures, Power recalls that one of the first moves of audit was to use the evaluation data that doctors already used to monitor practice for formative purposes.

Of course there is a huge tension in the purposes of audit/evaluation, as Wiliam and Black (1998), and others have pointed out (an account of this is in Williams and Ryan, 2000). When the bottom-up evaluations break free from top-down audit pressures Power calls this ‘de-coupling’. Our own efforts to develop formative and diagnostic work in connection with summative assessment was done as a means of offering de-coupling possibilities. The use of the ‘audit’ instrument by Ryan and Williams in the service of teachers’ metacognitive evaluation is a pertinent example. (See also Ryan later this session).

On the other hand when evaluators on the ground find themselves using Ofsted instruments to observe each others’ lessons (see eg Williams et al., in press; Corbin et al., 2004) then one sees the opposite, what Power refers to as ‘colonisation’ when the auditing practice ‘takes over’ the evaluation practice on the ground. The account of this in Williams et al. (in press) however reveals that this colonisation is always under contest: teachers can develop surprising resources for turning accountability systems to their own purposes, turning audit into evaluation. Thus managers expected to audit their teacher’ compliance with the three part lesson, when they saw a ‘great lesson’ that didn’t conform, went straight out and told everyone about it.

An important conclusion from the understanding of auditing practices is that the tension referred to is actually a ‘contradiction’ between opposites that actually feed off each other: audit MUST engage with local practice to be credible, and WILL try to colonise local practices. On the other hand this engagement with audit offers opportunities for subversion, the local effects of audit can be de-coupled and made useful in evaluation precisely because of auditors’ need for credibility.

If this analysis is right then what we are engaged in is always a political struggle over audit, its credibility, its power bases etc. To understand this is – I think – essential to

understanding where we are now in education. For instance, to attempt to ‘deny’ audit is to refuse to engage with powerful social forces, to leave the field open to their victory so to speak. The three questions above, however, do begin to set an agenda for engagement.

What is the purpose of audit and evaluation of teachers’ knowledge?

The CHAT analysis of this very particular tension – we prefer contradiction – referred to above has its social roots in the contradiction between two activities: those of the auditors (teachers’ managers, certifiers, accreditors, maybe even educators) and those of the auditees (the teachers and student teachers themselves, but maybe also their educators, though it is not always clear whose side teacher-educators are on). Teachers’ knowledge, for the purpose of the activity of teaching, has a different ‘meaning’ and ‘purpose’ from that of teacher knowledge for their auditors: it can be considered a boundary concept, and when reified in audit practices becomes a boundary object at the interface between the two.

Who are the auditors, or more significantly the commissioners of audit, here, and what does ‘teacher knowledge’ mean for them? I mention a number of groups that may each have distinct intentions, and between which there are yet more contradictions and tensions.... Politicians, their officers and teacher educators may have need of data to monitor/record the success (to account to their own public audiences) of their policy objectives or practices, of ‘improving the quality/qualifications of teachers’ / the ‘teaching force’ . Note the slippage here between ‘quality’ and ‘qualification’ on the one hand and ‘teachers’ / ‘the teaching force’ that offers possibility of engagement in critique: are they the ‘same’ thing, as auditors must rhetorically assume?

For these groups, some measure of teacher knowledge may provide essential exchange value, in meeting their social need for accountability. But in addition, in order for these measures to be credible, there is a need for auditors and their own audiences to buy into the notion that the measure does represent something ‘real’, some use value in teaching: this can only be determined by some articulation of a relation to the practice of teaching. Thus, the ‘percentage of teachers who are graduates of mathematics’, say, is only a viable audit measure if there is a credible relation between this measure and ‘teaching’ or potential teaching quality. This all offers much disputable terrain, but the contest over credibility is not essentially an academic one, it is in everyday political discourses, and discourses of common sense and the like, that the battle is won.

Who are the auditees here, and what is knowledge for them? It may be the teachers, for whom their knowledge is both use value (knowledge needed for them to be able ‘to teach’) and exchange value (the means for them to stake a claim to professional status, possibly accreditation). Another contradiction in the commodification of knowledge.

My conclusion: the essential, primary tensions and contradictions of audit reside in the contradictions ‘within’ the objects of the activities of teaching and auditing, but also (perhaps secondarily) between these practices. The first implicates the

'qualities/qualifications' and the 'use/certification' contradictions. The latter implicates the colonisation/decoupling contradiction.

What kinds of knowledge 'should' (mathematics) teachers 'have' or 'display' for/in (mathematics/) teaching?

This is the favourite territory of dispute for the mathematics teacher educator and many a researcher: 'we' all like to say that we want more than 'just' mathematical knowledge, facts and skills to 'pass on' to children/students, while for the public common sense suggests that this is just what teachers should know and do. This disjuncture between the teacher-educator discourse and that of the general public is ultimately what gives audit so much room for colonisation in the practice of teacher education.

Can we ask, what does a teacher need 'before' and 'when' they teach? Note in this question the acquisition metaphor (to 'have') is dominant here, and the process of 'display' appears somewhat strange or at least non-normative (Sfard, 2006). Note also the 'before' and 'when' that signify distinct audit/evaluation purposes again, at the boundary between training and practising teaching (another boundary).

Ultimately it will be argued that the 'acquisition' of certain objects of knowledge (concepts, etc) in training practices, before teaching become tools in the subsequent practice of teaching. But second generation CHAT asks us to attend to contradictions arising from just this kind of linkage between the two. What is reified in one system tends to need a lot of work to become a useful tool in another. When assessment in training becomes a tool of audit, this links the training/teaching systems with a third system, and it becomes increasingly more difficult to structure it to the purpose of 'use' in teaching.

In part this becomes a matter of technology: can we devise assessment tools that bring the training practices into line with 'use', and satisfy the demands of the audit culture for some measure of knowledge that allows them to claim accountability is in place (credibly).

How can we audit/assess teacher knowledge: what tools/technologies do we have?

The need for the development of appropriate technologies is by now apparent: the demands of audit require a credible measure, but credibility and de-coupling demand a sense of authenticity in relation to the primary products of teaching. One very simple technology that our group has been working with in the field of formative assessment may be an apt case to discuss. It is one of a number conducted in which diagnostic assessment instruments were designed for students, but were adapted to assess their teacher's knowledge too. In addition to the case described here below see also (i) an account of the methodology used with student teachers in the context of ratio which showed they had quite good estimates of the difficulty of items but very little awareness of pupils' associated errors and misconceptions; (Misailidou and Williams, 2002). and (ii) an account of Primary and secondary children's teacher's knowledge about probability that nicely separated accounts of the effect of experience from subject matter knowledge, with the more experienced teachers better predicting

learners' errors, but the younger teachers showing better knowledge of the topic! (Afantiti-Lamprianou and Williams, 2003).

Let us now look at an example¹ in some detail: in Hadjidemetriou and Williams (2002) we describe how a diagnostic assessment tool, developed from items from the research literature (Bell et al., 1987a, 1987b; Bell & Janvier, 1981; Hart, 1981) was constructed (a) to elicit pupils' graphical conceptions and misconceptions, and (b) to function as a questionnaire for assessing (and measuring) teachers' perception of the difficulty of the items for their learners. The test instrument was given to a sample of pupils and their teachers to establish a link between these two groups in order to compare results. (Pupils' group interviews and teachers' semi-structured interviews helped us to validate responses and gain an insight into ways of thinking of learners and teachers).

The items of the diagnostic instrument were deliberately posed in such a way as to surface known relevant graphical conceptions. It developed from an analysis of the key literature in the field of children's thinking and involved misconceptions such as 'slope-height confusion' (Bell & Janvier, 1981; Clement, 1985; McDermott, 1987), the tendency towards linear, smooth and other 'prototypic' graphs that pass through the origin (Leinhardt et al, 1990), the 'graph as picture' misconception, pupils' tendency towards reversing the x and the y co-ordinates, misreading the scale (Williams & Ryan, 2000) and so on.

The pupils' test was 'scaled' using a Rasch methodology and the result was a hierarchy of responses, each level of which was described as a characteristic performance including errors which diagnose significant misconceptions (Hadjidemetriou & Williams, 2001b, 2002b). However, group interviews gave us the opportunity to validate the test responses, in particular that the interpretation of the errors found in the test are symptomatic of the misconceptions discussed in the literature. In general, we found such interpretations to be valid, with just one problematic case of a misconception concerning children's slope-height confusion (Hadjidemetriou & Williams, 2002).

Twelve experienced teachers also participated in the study. They were asked to answer all the items and:

- predict how difficult their children would find the items (on a five-point scale starting from Very Difficult, Difficult, Moderate, Easy, Very Easy)
- suggest likely errors and misconceptions the children would make and
- suggest methods/ideas they would use to help pupils overcome these difficulties.

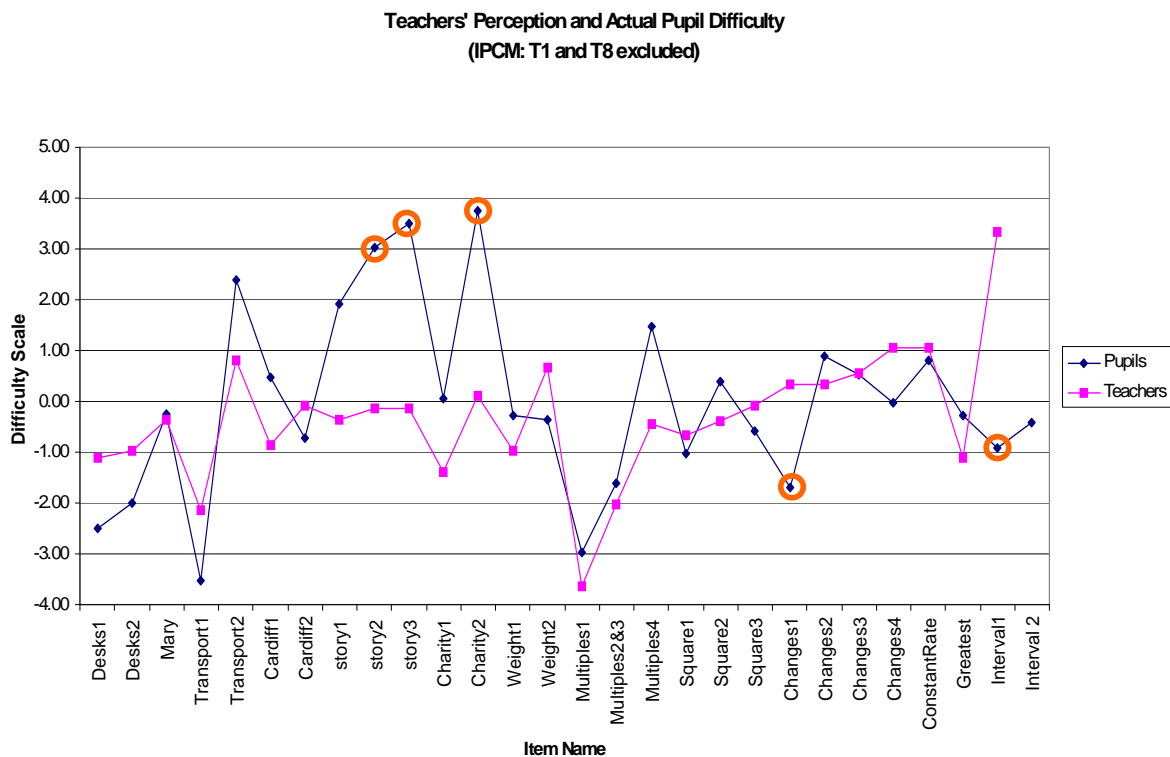
Teachers' knowledge was further explored through semi-structured interviews.

¹ NB the following account is an abbreviated account of that which appears in Williams et al. (AERA 2004) and draws from *Hadjidemetriou and Williams 2002-4*

From the teachers' rating scale data, using Rasch models, we scaled the teachers' perception of difficulty and contrasted it with the learners' difficulty hierarchy (see Figure 3). It was shown that some teachers over- or under-estimated the difficulties of some items. In Fig 3, the circled items are the items that the teachers 'most mis-estimated' in terms of their difficulty. Data from questionnaire and interviews suggested that these mis-estimations were due either to: (a) the teachers having the misconception the item was designed to elicit (content knowledge) or (b) the teachers incorrectly assuming that pupils required formal understanding of mathematical concepts to answer questions correctly, i.e. pedagogical content knowledge (Hadjidemetriou & Williams, 2001a).

The teachers' interviews on the other hand confirmed that the majority of them follow similar instructional sequences and that these are aligned with the prescribed National Curriculum. They also revealed that teachers' judgement of what is difficult is structured by this curriculum sequence: i.e. they sometimes incorrectly think that more 'advanced' in the curriculum implies more difficult.

Finally we were struck by these teachers' apparent lack of awareness in general of their children's conceptions and misconceptions. When asked what misconceptions



they might anticipate in their planning of teaching, few had much to say: yet when asked to predict errors in response to the test instrument, they were better able to predict what their pupils would do.

TEACHER MISCONCEPTION	1*	2*	3*	4*	5*	6*	7*	8*	9 I	10 I	11 I	12Q
<i>Slope height</i>					Q	Q	Q	Q				Q
<i>Linearity</i>											I	
<i>Y=X prototype</i>							Q	Q				
<i>Origin prototype</i>							Q	Q				
<i>Picture as graph</i>			Q	Q	Q	Q	Q	Q	I	I		
<i>Co-ordinates</i>			IQ	I		I	IQ	Q	I	I		
<i>Scale</i>	I	I	I	I	I	IQ	IQ	IQ			I	Q

‘Q’, ‘I’, and ‘QI’ indicate whether the misconception/error was mentioned by the teacher in the Questionnaire (Q), Interview (I) or both (IQ) while * indicates the teachers who were both interviewed and answered the questionnaire

Thus these teachers’ (who might generally be described as ‘leading’ teachers, in the sense that they were all experienced, promoted to leading positions, or active in education in their region) audited knowledge was highly sensitive to the methodology adopted to collect it (Hadjidemetriou & Williams, 2001a, 2002a). We concluded that their knowledge is distributed, and that well-researched tools might make all the difference in what they are able to articulate, with consequences for their planning of teaching perhaps, but also for the results of audit.

Shulman (1986) proposes that PCK appears in three different forms, propositional knowledge (e.g. knowledge of students’ errors and misconceptions), case knowledge (e.g. a vivid classroom experience of an error that a teacher was surprised by) and strategic knowledge (i.e. the art of acting in the moment, in particular despite potential information overload or ‘openness’, and perhaps lack of knowledge relevant to the situation).

Much knowledge is presented by teacher educators in the form of declarative statements or propositions possibly framed around a theory in a logical form. But these often lack richness of context and are therefore hard for practitioners to recall or use in practice. According to Shulman, these limitations make the propositional knowledge hard to apply. Case knowledge on the other hand, may bring these propositions to life and embed them in context:

Case knowledge is knowledge of the specific, well documented and richly described events. Whereas cases themselves are reports of events, the knowledge they represent is what makes them cases. The cases may be examples of specific instances of practice- detailed descriptions of how an instructional event occurred- complete with particulars of contexts, thought and feelings.’(Shulman, 1986, p.11)

By providing teachers with the appropriate tools that will ‘surface’ errors and misconceptions, we hoped to enrich this kind of well-organised but well-contextualised and usable knowledge. Thus such pedagogical tools might help mediate research knowledge, which might thereby be transformed aptly for teaching practice. ‘All’ that is then needed then is the strategic judgment to use the knowledge effectively in practice.

This link between ‘case knowledge’ and ‘propositional knowledge’ is in our view generally best conceptualised not just as a cognitive one (i.e. it is not only based on what teachers know and keep in their mind) but one which is socio-culturally

structured, i.e. mediated by well-researched tools in practice.

The following figure 4 illustrates the relationship proposed:

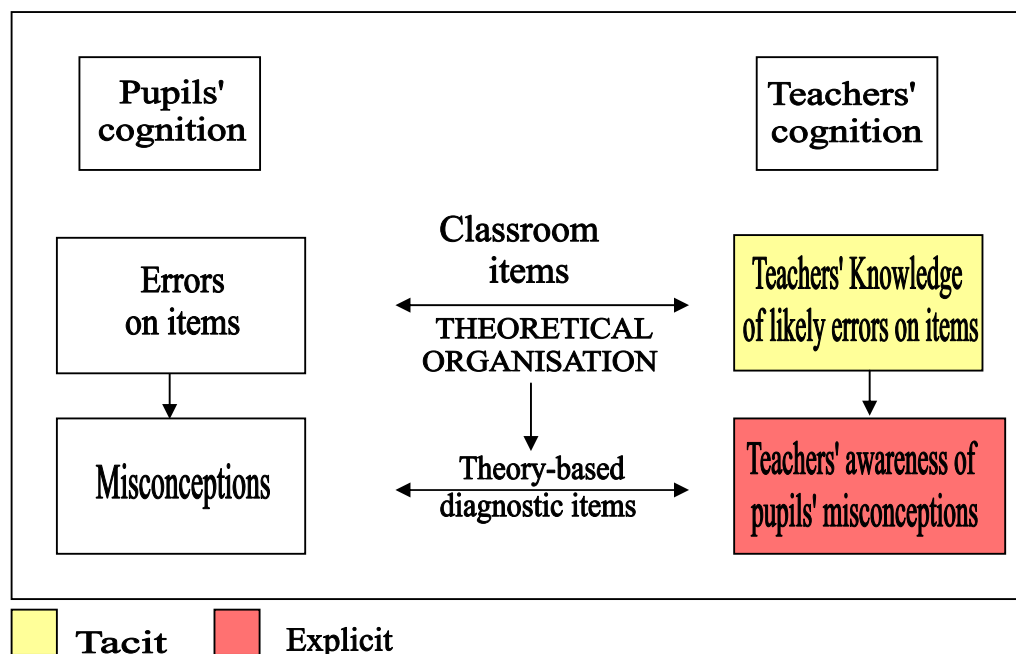


Figure 4: The emergence of knowledge about misconceptions

This suggests that teachers acquire (maybe through classroom experience) knowledge about their pupils' errors. This knowledge is tacit, based on the tasks and items used in the classroom. This also relates to teachers' propositional knowledge. However, if these propositions and pupils' errors and misconceptions are theoretically organised around tasks that aim to diagnose them then, firstly, deeper cognitive problems such as misconceptions come to the surface, and secondly teachers are made aware of them. We concluded that a well-designed diagnostic tool that includes items which will elicit errors that reveal theoretically-based errors (i.e. misconceptions), might help to transform teachers' tacit knowledge into explicit knowledge that could be used in teaching/planning.

In summary, our previous studies have shown that (i) the teachers we studied sometimes mis-judged their students' knowledge, and their judgments are influenced not only by their own mathematical knowledge, but also by their teaching experience and the intended curriculum; and (ii) their knowledge of their students can be strongly 'task-situated' and 'tool-mediated' rather than 'conceptual' and 'in the head'.

All this is suggestive of the obvious observation that audit and evaluation is tool-mediated, and that these tools shape the cognition in practice. But the triple objects of audit, training and teaching practices are at stake here: the tools we use are at the boundary of all three activity system, and need to reflect the exigencies of the three systems if they are to be stable.

Consequently, I suggest, audit/evaluation tools will be critical in shaping the results of audit and in their backwash, and need to be thought through in terms of their affordances for the colonising or decoupling practice. It seems to be important that

the tools we designed ‘bound’ ‘training’ and ‘teaching practice’ and the propositional and ‘case’ knowledge. It also seems to be important that they make measures, and hence offer tools for audit practices (summative). In this sense they seem to provide affordances for three systems/practices.

Towards a sociocultural perspective on pedagogy and teacher knowledge

I have appealed to social, cultural analyses of audit, evaluation and assessment practices in the foregoing and have focussed on how particular tools may mediate audit in significant ways; I now suggest some directions, and where the CHAT perspective might lead. It is increasingly obvious and widely recognised, I believe, that practice is mediated by tools, and that therefore audit is sensitive to the technologies of surveillance available. Less obvious or less well known – I feel – was our analysis of the social forces at work in audit and the consequences for understanding what is possible, and how and why productive or unproductive coupling, de-coupling or colonisation might be designed.

The CHAT perspective however may offer much more than this perspective on tool-mediation. First, CHAT recognises tool-mediation in ‘object-orientated’ activity as only one mediating factor among many that may be the source of significant contradictions and therefore dynamics. In addition we recognise the social division of labour, governed by social, cultural, historically formed ‘norms’ that position the subject inactivity. Furthermore, in particular, we recognise the inner contradictions in the ‘object’ of activity, and between activities and their bounding activities (Engestrom & Cole, 1993; Cole, 1996). Finally, CHAT recognises the possibility of ‘expansion’, for instance via the re-formulation of the ‘object’ of activity, or the collective ‘subject’. (Engestrom, 1987; 1991).

Where might these notions lead in the case of the audit of teacher knowledge? First, it is significant that ‘pedagogical knowledge’ is distributed in assessment tools, but this is only one ‘cultural crystallisation’ or reification of a more general social distribution of knowledge. If the ‘object’ of pedagogic knowledge then is the assembly of knowledge in an appropriate form in the learning-teaching environment then one must evaluate this in its social context. The result may be to question not why a teacher is unaware of the learners’ misconceptions, say, but why the scheme of work, the departments’/schools’ plans, the text book, and the assessment and professional development system as a whole are, ‘systemically’ unaware of the learners needs. In this view, a teachers’ knowledge can only be evaluated within the context of a pedagogic system, and the remediation of the system is at issue: the ‘blame’ for weaknesses becomes distributed. But so is the possible remedy, which demands a collective organisation of the ‘subject’, this raises the possibility of a community of teachers as collective subject (see Williams et al, in press). I think it makes sense to think increasingly of teacher knowledge in this way, as a collective property of a collectively cognising subject.

To credibly formulate the problem of auditing pedagogic knowledge at the level of

the individual, maybe just qualifying, teacher then, I argue logically requires the presumption of a 'normative' scheme of work, the departments'/schools' plans, the text book, and the assessment and professional development system, etc... But everyone knows that each institutional, pedagogical context is different. Appealing for a normative uniformity of affordances in school context for instance is not only unrealistic, it is equitable to the point of being revolutionary.

Then there is the question of the 'double bind' (Engestrom, 1987). The central contradiction of schooling, the principal source of alienation of learners, can be assumed to be that between exchange value (the learning of knowledge for accreditation, i.e. for advantage in the future distribution of resources, capitals etc) and use value (learning useful knowledge that enhances the capacity of the individual/social subject). The teacher experiences the same contradiction in relation to their pedagogic knowledge. The thrust of the argument for the formation of a collective subject is in finding allies that share an interest in escaping this double bind: in rewriting the rules... what might this mean in the case of 'teacher knowledge'? {I end this section with a question for pedagogical-dialogical reasons}.

Conclusion

I have argued that educational researchers need to understand the audit as a practice, and the contradictions inherent in it that might be politically exploitable. I have given an example from our own development work of how tools that audit knowledge-for-teaching provoke the realisation that knowledge is socio-culturally distributed, and a credible audit of knowledge-for-teaching invokes the development of contradictions between case and propositional knowledge that might be productive, and hence produce a de-coupling effect (a necessary corrective in these 'colonised times'). Finally I have flown a few kites as to where the CHAT perspective might lead- I hope this might provide room for discussion at the symposium.

References

- Adams, R. J. and Khoo, S. T. (1996). *Quest: The Interactive Test Analysis System*. Melbourne: Australian Council for Educational Research.
- Afantiti-Lamprianou, T. and Williams, J. (2003). 'A scale for assessing probabilistic thinking and the representativeness tendency.' *Research in Mathematics Education*, 5, pp ??
- Bell, A., Brekke, G. and Swan, M. (1987a). Diagnostic Teaching: 4 Graphical Interpretation. *Mathematics Teaching*, 119, 56-59.
- Bell, A., Brekke, G. and Swan, M. (1987b). Diagnostic Teaching: 5 Graphical Interpretation Teaching Styles and Their Effects. *Mathematics Teaching*, 120, 50-57.
- Bell, A. and Janvier, C. (1981). The Interpretation of Graphs Representing Situations. *For the Learning of Mathematics*, 2(1), 34-42.
- Bernstein, B. (1996). *Pedagogy, symbolic control and identity*. Taylor & Francis.
- Black, P. and Wiliam, D. (1998a). Assessment and Classroom Learning. *Assessment in Education*, 5(1), 7-74.

- Black, P. and Wiliam, D. (1998b). Inside the Black Box: Raising Standards Through Classroom Assessment. Retrieved at <http://pdkintl.org/kappan/kbla9810.htm>.
- Bourdieu, P. (1990). *The logic of practice*. Oxford: Polity Press.
- Bowen, G. M., Roth, W. -M. and McGinn, M. K. (1999). Interpretations of Graphs by University Biology Students and Practicing Scientists: Toward a Social Practice View of Scientific Representation Practices. *Journal of Research in Science Teaching*, 36, 1020-1043.
- Clement, J. (1985). Misconceptions in graphing. *Proceedings of the 9th Conference of the International Group for the Psychology of Mathematics Education*, 1, 369 - 375.
- Clement, J. (1985). Misconceptions in Graphing. *Proceedings of the 9th Conference of the International Group for the Psychology of Mathematics Education*, 1, 369 – 375.
- Cole, M. (1996). *Cultural Psychology: a once and future discipline*, Belknap Press of Harvard University Press, Cambridge, Mass.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsultit.
- Corbin, B., Mcnamara, O. & Williams, J. (2004) *BJES*
- Ebert, C. L. (1993). An Assessment of Prospective Secondary Teachers' Pedagogical Content Knowledge about Functions and Graphs. Paper presented at the annual meeting of the American Educational Research Association (Atlanta).
- Eisenberg, T. (1991). Functions and Associated Learning Difficulties. In D. Tall (Ed.), *Advanced Mathematical Thinking*. London: Kluwer Academic Publishers.
- Engeström, Y. (1987). *Learning by expanding: an activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsultit.
- Engeström, Y. (1991). Non scolae sed vitae discimus: Toward overcoming the encapsulation of school learning. *Learning and Instruction*, 1, 243–259.
- Hadjidemetriou C. and Williams J.S. (2001). 'Children's graphical conceptions: assessment of learning for teaching.' *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education*, 3, 89-96
- Hadjidemetriou C., and Williams, J.S. (2003). Using Rasch Models to Reveal Contours of Teachers' Knowledge, *Journal of Applied Measurement*.
- Hadjidemetriou, C., & Williams, J. S. (2002-a). Teachers' Pedagogical Content Knowledge: Graphs, from a Cognitivist to a Situated Perspective. *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education*, 3, 57-64.
- Hadjidemetriou, C., & Williams, J. S. (2002-b). Children's Graphical Conceptions. *Research in Mathematics Education*, 4, 69-87.
- Lave (1996) *Teaching, as learning, in practice*, Mind Culture and Activity
- Lave, J., and Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Leinhardt, G., Zaslavsky, O. & Stein, M. S. (1990). Functions, Graphs and Graphing: Tasks, Learning, and Teaching. *Review of Educational Research*, 1, 1-64.
- Leont'ev, A. N. (1978). *Activity, consciousness and personality*. Englewood Cliffs, NJ: Prentice Hall.

- Leont'ev, A. N. (1981). *Problems of the development of the mind*. Moscow: Progress.
- Misailidou, C. and Williams, J. (2002b). 'Ratio: Raising teachers' awareness of children's thinking.' *Proceedings of the 2nd International Conference on the teaching of Mathematics (at the undergraduate level), (ICTM2)*, <http://www.math.uoc.gr/~ictm2/Proceedings/pap143.pdf>
- Power, M. (1999). *The audit explosion: rituals of verification*. Oxford: OUP.
- Roth, W.-M. and Y. J. Lee (2007). "Vygotsky's Neglected Legacy: Cultural-Historical Activity Theory." *Review of Educational Research*.
- Roth, W.-M., S. Hwang, M.I. Mafra Goulart, Y.J. Lee. (2005). *Participation, Learning and Identity: Dialectic Perspectives*. Berlin, Lehmann Media.
- Ryan, J. & Williams, J. (2007). *Children's Mathematics 4-15*. Milton-Keynes, UK: Open University Press.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4-13.
- Sfard, A. (1991). On the Dual Nature of Mathematical Conceptions: Reflections on Processes and Objects as Different Sides of the Same Coin, *Educational Studies in Mathematics*, 22, 1-36.
- Sfard, A. and Prusak, A. (2005). Telling identities: in search of an analytic tool for investigating learning as a culturally shaped activity *Educational Researcher* 34(4): 14-22.
- Shulman, L. S. (1986). 'Those who understand: Knowledge growth in teaching.' *Educational Researcher*, 15 (2), 4-14
- Strathern, M. (Ed.). (2000). *Audit cultures: anthropological studies in accountability, ethics and the academy*. London: Taylor & Francis Ltd., Routledge.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.
- Wenger, E. (1998). *Communities of practice: learning, meaning and identity*. Cambridge, Cambridge University Press.
- Williams, J.S. (2005) *The foundation and spectacle of [the leaning tower of] PISA*; Melbourne: proceedings of PME-2005.
- Williams, J. S. & Ryan, J. T. (2000). 'National Testing and the improvement of Classroom Teaching: can they coexist?' *British Educational Research Journal*, 26(1), 49-73.
- Williams, J., Corbin, B., McNamara, O. (in press: 2007) Finding inquiry in discourses of audit and reform in primary schools. in *Sociocultural and Cultural-Historical Activity Theory perspectives on subjectivities and learning in schools and other educational contexts*. *International Journal for Educational research*, 46, 1-2. Special double issue.
- Williams, J., Black, L., Davis, P. (in press: 2007) Introduction to 'Sociocultural and Cultural-Historical Activity Theory perspectives on subjectivities and learning in schools and other educational contexts'. *International Journal for Educational research*, 46, 1-2. Special double issue.
- Wright, B.D. and Masters, G.N. (1982). *Rating Scale Analysis*. Chicago, MESA Press.
-