

How do teachers make overall teacher judgments (OTJs) and how are they supported to make sound and accurate OTJs?

**Kerry Mitchell, The Education Group and Dr Jenny Poskitt, Massey University
Prepared for NZARE Conference, Auckland, December 2010**

Acknowledgment:

Thank you to the principals and teachers who willingly agreed to participate in this study and the Ministry of Education for funding the research.

Introduction

Within the educational international context of political concern about standards of education the New Zealand government has launched National Standards. The intention is to monitor student achievement and to ascertain whether standards are improving or declining across the school sector. To date, New Zealand has avoided the widely criticised national testing programmes introduced elsewhere, notably the 'No Child Left Behind' (NCLB) policy introduced in the USA in 2001. New Zealand has focused more on formative and cumulative teacher judgments with the intention of more valid and informative assessments of student progress. Central to this policy initiative is the notion of teacher judgment.

Teacher judgment

Sadler (1988, p.80) describes three elements that comprise teacher judgment of student work. The elements include teachers attending to the learning students produce, appraising this work against a reference framework, and making an explicit response such as feedback or judgment on the learner's work. Overall teacher judgment has an added complexity in the judgment process being applied to a range of data.

According to the Ministry of Education fact sheet (MOE, 2010) on overall teacher judgment,

"An OTJ draws on and applies the evidence gathered up to a particular point in time in order to make a judgment about a student's progress and achievement. Using a range of approaches allows the student to participate throughout the assessment process, building their assessment capability.... No single source of information can accurately summarise a student's achievement or progress. A range of approaches is necessary in order to compile a comprehensive picture of the areas of progress, areas requiring attention, and what a student's progress looks like."

'Triangulation' of information increases the dependability of the OTJ. See diagram 1 for more detail:

Diagram 1: OTJ in relation to National Standards



The issues identified by the MoE in the above article include:

- The need to confirm dependability of evidence and that the assessment is consistent with teacher experience
- Moderation both within and across schools
- Student participation
- Reporting to parents, families and whanau

National policy for OTJs implies that teachers understand what OTJs comprise and they are able to use the processes appropriately in classroom practice. However, the application of teacher judgment has been somewhat problematic in other countries. For example, a report in Queensland, Australia by Cumming, Wyatt-Smith, Elkins & Neville (2006, p.9) recommended guidance for teacher judgment by “descriptive elaborations of standards, citing the benchmark performance within such standards, should be developed to guide teacher assessment, with opportunities for teachers to engage through some process of moderation and sharing of exemplars.” These authors recommended further professional development with teachers on the interpretation and use of assessment data on which teachers formulate judgments.

The Queensland Studies Authority (QSA) funded several projects as part of a research programme to explore the capacity of teachers to make judgments against defined standards and whether their judgments could provide more valid and reliable information about student capabilities than standardised tests (Maxwell, 2002; Cumming et al, 2006). One literature review indicated that Queensland systems were underpinned by two assumptions, “that teachers as professionals are able to make appropriate judgments about students’ work, and moreover that teachers (and students) are best placed to make judgments about students’ work and to provide full information on student

performance in a range of contexts and through a range of assessment opportunities” (Cumming et al, 2006, p.16).

Interestingly, a research study by Meisels (2001) in the US investigated teachers’ capability to make consistent and therefore trustworthy judgments on student work. Meisels asked teachers to collect and evaluate samples of student work in structured portfolios, to use checklists and prepare summary reports three times a year. The researchers compared this Work Sampling System (WSS) with student results on the Woodcock-Johnson (WJ) battery for literacy and math measures and found correlations between the WSS and WJ ranging from 0.5-0.75. Such correlations indicate sufficient degree of correlation to trust and support teacher judgments but insufficient degree of correlation to suggest that WSS and WJ measure the same dimensions. Meisels (2001) concluded that teacher judgments from student work samples are inadequate for high stakes decision-making or comparison of students, but helpful for informing ongoing teaching and learning programmes, ongoing monitoring of student progress and for informing parents.

Many authors (e.g. Sadler, 1988; Cumming et al, (2006) emphasise the value of formative assessment for ongoing learning, in contrast to summative assessments based on standardised tests that reduce validity, narrow the curriculum and can have a negative impact on student learning. A system that values teacher judgments supports formative assessment practices. Cumming et al (2006, p.16) cite Kellis and Silvernail (2002) who argue that teacher judgment of student progress has multiple benefits: “provides immediate feedback, secondly teachers are able to use more recent information on student learning to make instructional decisions rather than waiting months for test scores; and thirdly, owing to the unobtrusive nature of teacher judgment techniques such as observation, teachers are able to made deeper judgments of student learning that go beyond fact based, short answer responses.”

However, other authors argue that such benefits only derive from teacher judgments that are accurate and consistent. Indeed, Gardiner, Tom, & Wilson (2000) argue that the key issue in the role of teacher judgments is consistency. In order for the wider educational community to be confident about teacher judgments teachers need to be consistent so that “judgments about student learning are not dependent on the individual teacher, student, location or time, and are based on a shared understanding of syllabus and standards of learning. “ (NSW Department of Education and Training (2008, p.1) To be assured of teacher judgments users of the judgments need to be confident that such judgments are consistent with the judgments of other teachers in a school and teachers in other schools.

Increasing confidence about OTJs is purportedly developed through extensive processes of moderation (e.g. Maxwell, 2002). An extensive discussion of moderation is beyond the scope of this paper. What is pertinent here is to indicate that research literature demonstrates that the application of OTJs can be problematic unless teachers are clear about *what constitutes an overall teacher judgment*, that they have *common understandings of ‘standards’*, that such *understandings are supported by clear criteria and exemplars* of student work demonstrating achievement of standards, and that *moderation processes are used to ensure appropriate and consistent judgments* are made (Gardiner, Tom & Wilson, 2000). Given that New Zealand is only beginning to use OTJs in

the primary school sector, this research investigation set out to explore teacher understandings about OTJ and how they were applying OTJs in the classroom and school.

Context and methodology of the research study

Although some studies have been conducted in countries such as the US and Australia about teacher judgments, little was known about New Zealand teachers' understanding, and application, of OTJ or any related professional learning needs. Because little was known about teachers' understandings an exploratory (intrinsic) case study research design was selected (Punch, 2009). Exploratory case studies enable researchers to develop a better understanding of a case (teacher understanding and use of OTJs), in natural settings (schools), within the complexity of school life. The phenomenon of OTJs was explored within the bounded context of a sample of ten primary schools in the 2010 Assess to Learn professional development contracts (for information about AToL refer for example to Mitchell & Poskitt, 2009).

These schools were invited to participate because they were already part of a larger evaluation study of the Assess to Learn (AtoL) contract, and hence willing to engage with external evaluation and research. Furthermore, their participation in formative assessment professional learning meant they were likely to have some understanding of classroom and school-wide assessment processes that could be drawn on for making overall teacher judgments. If these teachers struggled with the concept and application of OTJs then this would have significant implications for nation-wide professional development programmes.

Typically case study research design uses multiple sources of evidence. Within the broader evaluation of AtoL a range of data were collected: researcher, teacher and facilitator observations, student achievement data, national questionnaires of teachers and principals. In this smaller exploratory case study semi-structured interviews were the primary source of data, supplemented by document analysis. Semi-structured interviews enable researchers to prepare interview questions in advance (Punch, 2009). Advanced planning also enables interviewees to be at ease, knowing the questions to be asked, and to gather relevant information, think or reflect prior to the interview. Structured questions also enhance the likelihood of collating and analysing responses from a range of participants. Incorporating some flexibility is possible with semi-structured interviews which allow for spontaneity beyond the pre-prepared questions and further exploration of ideas in depth with participants.

Interviews were conducted with school principals, lead teachers and classroom teachers. An interview protocol was developed and used for all interviews; with a slightly different version for principals. During the interviews participants were encouraged to refer to samples of teacher or school documentation related to understandings of OTJ, benchmark or reference points against which judgments were made and processes used in the school.

The interviews were recorded and analysed. The first interviews took place mid-year, June 2010 and the same people were followed up again in November 2010.

Schools from four regions were invited to participate, from which ten schools chose to be involved: three in Auckland, three in Christchurch, three in Manawatu and one in Wellington. This represented five (out of eight) AtoL professional development providers.

Findings from the initial interviews were reported to the Ministry of Education (MoE) and to the AtoL providers so that assessment facilitators could respond to the identified needs during the second half of 2010. Findings from the second interviews were also reported to the MoE and providers.

Data analysis

Interviews were initially transcribed and then collated according to type of interviewee and interview question. For example, all principal interviews were labelled and collated into a principal response file; then responses to each question analysed question by question. Coding sought to summarise the data by identifying themes and patterns. The first phase established descriptive codes and themes (finding and describing relationships amongst data), which later developed into second-level, inferential pattern codes (giving meaning or interpretation to the data relationships). Punch (2009) refers to pattern codes as being more abstract and interpretive than descriptive codes. Clear links were made between descriptive and pattern codes to ensure coding could be checked and that there was inter-coder reliability between the two researchers.

Analysis of the findings reveals several themes related to teacher understanding and application of OTJs in their classrooms and schools. The reader is cautioned however, that this is an exploratory study of only ten schools so the findings, whilst potentially useful to inform future professional development provision, have limited generalisability.

Findings

Principals (10) and teachers (30) were interviewed to explore understandings about making overall teacher judgments and processes used in the school to support teachers. The data are presented here in two broad categories: teachers and principals/lead teachers.

Teacher findings (June data)

Understanding of overall teacher judgement (OTJ)

Schools and teachers varied in the way they conceptualised overall teacher judgments (OTJ). A few teachers (n=5) viewed OTJ as a “gut feeling” from all their professional interactions with students, while most teachers viewed OTJ as a professional judgment based on considered analysis of a combined range of assessment data. Some (n=7) teachers added to this mix the student’s self or peer assessment and the student view on their overall performance. Others considered that OTJ was identifying ‘best fit’ between criteria and performance and triangulating data. Finally, a few teachers (n=3) compared the combination of assessment data to a framework or reference point (such as writing matrices, numeracy project framework, or NS).

Two interview extracts illustrate the range of understandings about OTJs:

My initial thinking is it is a gut feeling and reaction. I observe at mat time, listen and see their [children's] response to my questions. I sometimes gather information at the end of a lesson to support my gut feeling. (PS, T1, 7/6/2010).

OTJs are formed by combining a range of data from informal and formal assessment, incorporating the child's views as well as the teacher's. I draw on conversations with the child, something in writing (like their use of self and peer assessment), and refer to my modeling books. I incorporate the kid's judgments and evidence from a range of writing like narrative and report writing and highlight when we have evidence that reached something has been achieved in relation to the [writing] matrix. I use a range of data, at least three sources because otherwise the information could be affected by influences on the day like playground interactions. (LS, T4; 16/6/2010).

Teacher disposition towards either anxiety or confidence about making OTJs appeared to reflect the professional leadership perspective of the school, the extent of professional discussion and professional development the school had engaged in and the nature of their student community (i.e. the proportion of students likely to meet NS). The schools where principals were supportive of teachers making OTJs and where systems were put in place to strengthen decision making (such as moderation processes) contained teachers with greater confidence and skill in drawing on a range of data to inform their judgments.

Use of evidence in making OTJs

Teachers varied in the extent to which they referred to evidence or data in supporting their OTJs. A few teachers (n=4) relied extensively on their years of teaching experience (they “knew where a child was and what to do based on their experience”), whilst most teachers mentioned referring to a range of tools such as Running Records, classroom observations while teaching groups of children, formal tests, including aSTtle, PATs, exemplars and modelling books. Conversations with students ‘We check with kids.... can they explain it?’ were also mentioned as evidence supporting OTJs.

Three interview extracts illustrate use of limited, and a range of information:

Our professional judgment and knowledge that we call on after years of experience [teaching this level]; accumulated over the years, combined with observations of children and assessments and talking with each other (BS, T1, 8/6/2010)

Mostly visual and working in small groups; lots of discussion (PS, T1, 7/6/2010)

I use a range of information like standardised tests, PATs, diagnostic surveys and knowledge tests in maths, modeling books, kids work, matrix in writing, feedback and feedforward giving kids and whether kids acting on it – look to see if they have acted on – write in pencil and edit in black pen and see whether change has occurred since I made the comment (written or oral). Self assessments they do and where I make comment alongside child's. Sometimes I discuss with teacher aide progress students making and discuss with them before making judgment. (LS, T4, 13/6/2010).

What factors influence teachers in making judgments?

At one end of the continuum, individual circumstances of the child were mentioned by many teachers as influencing their overall judgments, while at the other end of the continuum a minority of teachers accepted only what they saw in front of them at the time of the assessments. Individual circumstances for students included 'level of stress;' 'sensing when kids are tired' 'student health' 'what is happening in peer group and family'. A few interviewees mentioned that parent concerns and comments about a lack of progress, standardised test results that differed from teacher expectations or surprising results caused them to review their data and OTJs.

One interviewee appeared to rely solely on parent information,

Often chat with parent and find out how reading going, to help me out. If parent says the child is struggling then I know the reading material is a bit hard (BS, T2, 7/6/2010).

Another teacher indicated that she responded to children on a day-to-day basis in making assessment decisions:

Depends on the day – e.g. if child has been with one particular parent that weekend, whether they have slept and been fed properly, knowing if child in good or poor space – might delay doing Running Record until they are more rested; using standardized data and what should be doing and what can be doing. Use different assessments because some assessments benefit or limit others, and what I know from working with children on daily basis. (LS, T2, 13/6/2010).

Other interviewees drew on a range of information and involved several personnel in the OTJ decision-making process to minimize the effects of one-off results:

Our good classroom practice, anecdotal notes, modelling books, traditional stand tests, conversations with pupils, setting next step, feed-forward and feedback; involve parents if got problems for ongoing learning rather than making judgements; e.g. basic facts or fluent reading; moderating with kids, writing discussions between teachers. Starting student led three way conferences and explain where at and what next step. (HS, T3, 7/6/2010)

What do teachers do when they are uncertain about a judgment?

Most teachers discussed the judgment with colleagues either informally or formally. Moderation processes were often a part of the formal process. Rounding down to a lower level when in doubt was mentioned as a process by a small number of teachers.

Re-do with more preparation; or if stuck I go to more experienced teacher or leader or another teacher. Discussions with the children. (PS, T1, 7/6/2010).

Ask other colleagues' views on student work and 'where to' level. I approach experienced staff or those with deep knowledge of curriculum and NS and find out more information

about particular characteristics that are indicative of a stage or level. I now have increased familiarization of standards with our matrix. (HS, T1, 7/6/2010).

Ask other people – work with people within the team; moderating took samples of different level and worked with people in team and compared samples of work we thought were in a particular level. Curriculum leaders help, RTLit, Numeracy adviser, exemplars, advisors, moderation, seek the views of other schools if really stuck. (LS, T2, 13/6/2010).

What tools are used in making an OTJ?

Almost all teachers talked of using students' work and student observations alongside tests to complement decisions.

Focused classroom observations, student workbooks modeling books, Running Records, learning conversations, standardized tests e.g. PATs, Gloss. NS focused us on making a difference and making it higher and lifting our game; some kids formerly okay and now thought perhaps borderline. Makes for focused teaching and knowing what to do differently to move kids on and kids are aware now too in AtoL. Fortunate in implementing new curriculum and AtoL, reporting to parents and NS – all aligned; outstanding facilitation of facilitator has been integral to staff understanding and taking on board. (HS, T3, 8/6/2010). Having more than one piece of evidence – a range of assessments to get best fit. Not sure – need one piece of standardised assessment, but I place those in background especially if done some time ago – in process but not the dominating factor. I may make up a test according to the criteria I have and use that as one piece in the puzzle. (LS, T4, 13/6/2010).

How do teachers ensure judgments are soundly based?

Moderation and collegial support were the methods used by most teachers. 'Gut feel' was mentioned by a minority but in all cases used in addition to other tools which provide back up evidence of the 'gut feel'.

Gut instinct; talking with other teachers and getting second opinion. (PS, T1, 7/6/2010).

Starting to develop own school matrices and hope to use that for moderation. Matrices give us confidence – recorded in writing then got support to back up judgments if asked. Matrices are really valuable for less experienced teachers. We also refer to criteria, tests used, combined with professional experience from what we know where kids need to be. We look at 'where at' and 'where need to get to'. We focus on where they started from; realize kids all different. (BS, T2, 8/6/2010).

Having more than one piece of evidence – a range of assessments to get best fit. It is important to involve child in the process as well. PD side is teachers working together in teams as a school and all getting on same wavelength. We discussed in staff meetings our diagnostic survey for numeracy - how to think about it and administer in similar ways. This

type of PD is relevant and useful. It is part of our job (the PD) to ensure we all understand the OTJ and all on same path. (LS, T4, 16/6/2010).

PD around matrices and alignment with NS; level and alignment on tki. So reassuring – you know in your gut intuition and then validate with some data. (HS, T1, 8/6/2010).

What professional development assistance or resources would be helpful?

Some teachers requested guidance in the tools that should be used while others were concerned that OTJ methods could become too prescriptive and restrictive, limiting the recognition of teacher professionalism. There was also a difference of views on the time required to make OTJs. On one hand some teachers indicated that moderation is time consuming while a minority stated that time is not an issue as OTJs are made on a daily basis.

Need PD on moderation – tools do not seem to align, like asTTle with National Curriculum Exemplars just leave me confused. (PS, T1, 7/6/2010).

Perhaps more time in classroom to assist with reading/writing – and recording of assessment information would be useful. (BS, T2, 8/6/2010)

Really need a common reference point, like the maps on TKI that show the various tools teachers can use in reading to come to a judgement of where children are at in a particular curriculum level. If a range of scores was given for each tool and where they fit for each curriculum level and each National Standard that would make teachers' judgments easier to form and then more consistently applied throughout New Zealand. (HS, T3, 8/6/2010).

PD on how to moderate and time to do so! (WS, T2, 9/6/2010).

How are teachers reporting to parents at mid-year?

There was a wide variety of responses to midyear reporting in 2010. Responses included the provision of student led conferences for parents, with reference to progress towards NS at the end of the year. Some teachers were awaiting information from the NS workshops scheduled during the year and planning to make changes to the end of year report, but not the mid-year report in 2010.

I look at all the evidence and make an overall teacher judgment based on the evidence. (ACS, T1, 10/6/2010)

Many teachers reported their intent to show progress over the year and for many reporting on end of year achievement was still a 'work in progress'.

What is challenging about making OTJs?

Responses ranged from 'no challenges' to a variety of challenges including 'knowing what is enough', ensuring clarity about differences between specific writing genre are understood and identified, 'time', contrast/contradiction with NCEA's criteria ranging from 'not achieved' to 'achieved with

excellence', getting parents to understand assessments and challenges of consistent messages from different teachers. One principal (not reported here) considered that teachers needed to know how to give difficult messages to parents when their child's performance was below what parents understood it could/should be.

Summary of teacher findings

The brief literature review in this paper indicated that unless four factors are met, teacher judgments can be problematic. The four factors were: *what constitutes an overall teacher judgment*, having *common understandings of 'standards'*, that such *understandings are supported by clear criteria and exemplars* of student work demonstrating achievement of standards, and that *moderation processes are used to ensure appropriate and consistent judgments*. In June 2010 teachers varied in their understandings of OTJ, how to make and use them. Whilst the majority of interviewed teachers understood the need to gather and interpret a range of information gathered over time, and the variability of student performance from day to day, there was not a common understanding about the 'standard' nor about how to make judgments. Only some schools referred to clear criteria and exemplars and few were aware of the need for moderation or comparing judgments against a benchmark (such as school matrices, TKI progress maps) and any reference point varied from school to school. This variation is cause for concern because it means the basis on which teachers and schools are making judgments is not consistent. In conclusion, in June 2010 teacher judgments across the ten schools were problematic because there was not common understanding of what constituted OTJs, understandings were rarely supported by clear criteria and exemplars, and few schools had robust moderation processes in place.

Principal and Lead Teacher findings

Principals and/or lead teachers were asked how schools moderate overall teacher judgments and responded to questions as follows.

1. What is your understanding of overall teacher judgment?

Principals/lead teachers participating in this exploratory study expressed a range of understandings about overall teacher judgments (OTJs). A few principals (5) referred to making judgments from evidence, others to essentially what teachers do continuously in their role, and others made reference to National Standards.

Always done teacher judgments, now it is in relation to specified National Standards. (HS, 8/6/2010).

OTJ - a way to recognize what teachers do. In the absence of national test it is a more acceptable way for teachers to intuitively make assessments about children's progress. Collecting of all teachers see and do and judge and combination of heart and what do. I am

concerned about how NS will work, how it will align and how it might work in relation to standard normed tests. (LS, P, 16/6/2010).

Instead of relying on one piece of information we use a range of things such as teacher observation, conversations with student, PATs, STAR, e-asTTle (previously asTTle V.4. Involving the child in developing own assessment capability. Student led learning conferences. (MS; P; 9/6/10)

OTJS involve teachers making a judgment call around a number of assessments and evaluations, based on knowledge of student as a learner. Involves other colleagues, depends on experience of teachings. It is drawing a line in the sand using various tools and input of colleagues. Background (e.g. ESOL) is also considered. (BI; P, 16/6/10)

2 What process does your school use to support teachers making soundly based OTJs?

A range of views were expressed in regard to school processes. Themes varied from OTJs not being new, teachers struggling (need for clarity, curriculum knowledge, time, agreed processes), implementing professional learning processes to support OTJs (such as in-depth staff discussions, systematic observations, development of exemplars, development of learning culture amongst staff, allocation of buddy classes for critical friend processes when examining data), developing teacher skills in formative assessment (through AtoL professional development) and involving parents. Interview extracts illustrate these themes below.

We are very aware of the need for consistency. Classic example is the consistent approach taken with Running Records (training each year on the way we expect it done in this school, and how to interpret data – important information for new teachers and good reminders for other teachers), then we can be assured that everyone is administering and interpreting results consistently. (LS; 16/6/2010)

We used the MoE rubric and realized our knowledge was low. Teachers need acceleration and are trying hard but don't have the skills to do it. AtoL is best way to do it. (TP; P; 8/6/10)

We needed AtoL before we could be ready for OTJ – teachers have been struggling. They get plenty of support but they need curriculum content knowledge. (TPS, P. June 2010)

New Principal acknowledges fact that children were underperforming although teachers thought they were at the level (AS, P, 5/6/2010)

Staff have been resistant. Last year I introduced a new staff position to support teachers implement AtoL and literacy in classroom. This person keeps the pressure on staff with no let up between facilitator visits [so we will ready to make sound OTJs]. (TPS, P. June 2010)

Formal testing, records of classroom observations, moderation – looking at levels between class and overall school; Quality Learning Circles in syndicates and professional conversations about children; professional development from AtoL. (BSP; 8/6/2010)

We are involving parents in the process through our student led conferences (BIS, June 2010)

What criteria/ framework are used to ensure appropriateness of OTJs?

Principals varied in their understanding of the need for a consistent framework and how to implement an associated process in their school.

On learning curve for that one; beginning discussions with other schools in local cluster. Realize need to fill some gaps in teacher understanding. (BSP; 8/6/2010)

Maths: I can, PATs. English: exemplar matrices, TKI, STAR, PATs, RR, PM Benchmark; ARBs. (HS; 8/6/2010)

School has a set of general criteria teachers can work off. Also range of assessment tools available. (PS; 9/6/2010)

Consistency promoted through PD prior to administration of assessment tools (ensuring consistent application and interpretation). Would appreciate more tools for writing at junior level and reading in senior level. (LS; 16/6/2010)

Time needed to understand what NS are. What information needs to be fed in? What is evidence? What needs to be recorded? Needs to be manageable. (AS; P; 9/6/10)

How does your school know that OTJs are consistent¹ and appropriate?

Principals varied in their understanding of the need for moderation and how to implement robust processes.

Part of learning curve – next step would be moderation within and then across schools. We will know through interpretation of tests and assessment information – it will be interesting to see range of interpretation across schools. (HS; 8/6/2010)

Consistency comes from professional development, discussion and providing the right tools for teachers to use. Team and curriculum leaders scrutinize student evidence and assist teachers with judgments they unsure about. (LS; 16/6/2010)

Through moderation processes, listening to more and less experienced teachers sharing ideas. (BSP; 8/6/2010)

Moderation and involvement of experienced teachers. Teachers have experience and also gaining confidence with experience. AtoL is used school wide and all are learning together. The facilitator also models the moderation process in writing. (BI; P, 16/6/10)

¹ Consistent means coming to the same judgement when presented with the same or similar evidence.

What future Professional Development might your school need?

Principals indicated professional development needs for their schools related to teacher content knowledge, preparation for moderation, and developing common understandings of criteria for curriculum levels and particular standards.

Continuing in AtoL next year; need whole staff involved and thinking in similar ways about OTJs, Quality Learning Circles fostering professional discussions. (BSP; 8/6/2010)

Moderation – AtoL facilitator starting to lead us with that (HS; 8/6/2010)

Sustain current focus and depth in literacy numeracy; deepen teacher curriculum knowledge in other areas e.g. Science, Social Studies; greater access to assessment tools in those areas. (LS; 16/6/2010)

Huge amount of PD needed: Reading/Writing/Maths. Need to build in moderation. Ensure thorough content knowledge. Where coming from and where going to? Need evidence. (WB; AP; 3/6/10)

Recent clarity about learning progression. Level 3 needs to know what level 1 looks like. (WB; AP; 3/6/10)

How might you communicate this information [achievement of national standards] to parents?

Principals reported a range of strategies for communicating information about National Standards and children's achievement in relation to National Standards. Strategies varied from written information in newsletters, student portfolios and written reports, through to dedicated face-to-face parent information sessions.

Examples from interview extracts:

Through newsletter – drip feeding through newsletters; Portfolios highlight where child working at reading and writing at age band and that fits well with NS and next learning steps. (BSP; 8/6/2010)

Use relevant phrases on reports and align judgements with standardized tests and TKI interpretive chart (indicating where assessment tools sit in relation to NS) (HS; 8/6/2010)

Entry level meetings with parents, discussion of 6 year net results, 3 way conferencing, student portfolios, parent education sessions at school, written reports, newsletters, website especially blogs. (LS; 16/6/2010)

2 x NS meetings for parents so far this year with 1 divided into cultural groups. Chinese parents want tests. Indian parents ask about moderation (fairness). Middle Eastern

parents want children kept back if not achieving. African parents are worried because they saw similar standards introduced but not working. Maori parents very positive. (WB; AP; 3/6/10)

Now using 3 way student led conferences instead of parent meetings. Parents have an understanding of students leading their learning. This is new for us. Feedback is oral, presented positively and not formally assessed. End of year report language will include 'moving towards' rather than 'well below' NS. It is crucial to show progress over the year. Results will be triangulated to ensure assessment is accurate. If asTTle test results are lower than teachers' OTJ, both would be reported with qualification to inform parents. (BI; P; 16/6/10)

Most principals believed it was important to involve students in the reporting process, “empowering kids to share their learning with their parents”, “the more we involve the students the greater the ownership and deeper the learning”, “parents astonished at depth of student knowledge and sophisticated communication about their learning at 3 way conferences; blogs”.

What concerns might you have about communicating OTJs to parents?

Principals and lead teachers interviewed for this exploratory study expressed concerns about a range of social issues, (demands made on teachers to fix up social issues beyond jurisdiction of schools), communication (balance needed between reporting in plain language and education jargon), teacher professional knowledge (variable levels of teacher knowledge about curriculum and understanding of standards), not having common understandings of standards, parent perceptions (children’s assessed achievement levels differing from primary to intermediate level), the realities of student learning (student plateaus in learning at particular stages, learning in different ways and at different rates), and the speed at which NS were being implemented. Some schools believed they had sound assessment systems and expressed few concerns.

Summary of June Principal/Lead Teacher Interviews

Although principals generally had a greater understanding than teachers of the intent of OTJs, there was insufficient evidence of teachers within their schools sharing a common understanding of how to reach consistent OTJs, nor consistent processes in establishing explicit criteria, exemplars or a reference point against which to make OTJs. Moderation processes were well established in a few schools, but sporadic or absent in other schools.

End of Year interviews

In November 2010, researchers returned to the same schools for a second time, to investigate progress made in understanding and implementing OTJs.

End of Year interviews with Teachers

Given the length of this paper, only key distinctions are reported here in relation to progress by November. Interview data have been analyzed in relation to the four key factors (derived from the literature review) that are necessary for sound overall teacher judgments to be made.

What do teachers understand about OTJs?

A few teachers reported minimal difference in their understanding of OTJs. Indeed three teachers stated their perception now of it being a “snapshot at this stage of the year, rather than progress over time.” In contrast, two teachers reported now realizing the importance of considering class work as well as test results, the need for reflection and moderation, and triangulating data from multiple sources. Another group of teachers reported changes in now comparing beginning and end of year data to show progress, as well as now needing to distinguish students performing below, at and above standard. The final group of teachers indicated that they now ensure they have at least three pieces of evidence which they triangulate and on which they base OTJs.

In summary, a small group of teachers are now able to articulate appropriate understanding of OTJ. Of concern is that even within an assessment professional development contract, there are teachers who do not fully understand the need for valid, reliable and usable data on which to base OTJs. A tentative conclusion (given the small sample size and limited generalizability), is that it will take time for the teaching professional to develop shared understandings of OTJs and to embed practices that are valid, reliable and usable.

How common are teacher understandings of ‘standards’?

Frequently stated by teachers interviewed in November, was the need for sufficient information or evidence on which to make a judgment and the need to work with colleagues to check understanding. Fewer teachers mentioned the need to refer to standards (or a reference point) to match the achievement with teacher judgment. Sadler (1988) argues that comparison with a reference framework is one of three requirements for making effective teacher judgments.

A concerning trend was reliance on, or deference to, a standardized test result when it differed from informal assessment results, particularly for less experienced or less confident teachers.

Our professional knowledge over the year has grown so [my team leader] said I mark them harder because my standards may have changed as my knowledge increases... If I have done my OTJ I'd put in the writing level they got in the test or would I actually.. no you'd put in your OTJ.. and change the level that way.. well I am not sure really. (TPS, T2, Nov 2010)

Many teachers were not sure how to ‘triangulate’ or synthesize data, particularly when results varied according to the tool used.

In summary, teacher understandings of ‘standards’ are not common. There is mixed interpretation of standards, uncertainty and ambiguity as to which assessment results to use when variability

arose. Professional development is needed on aggregating and interpreting assessment information, as well as clearer guidance on appropriate interpretation of standards.

To what extent are teacher understandings supported by use of clear criteria and exemplars?

Some teachers discussed the need for more exemplars and deeper discussion of what constitutes achievement of particular standards, “we need to tease out what we mean by particular indicators in our matrices”, “working through differences in our marking”, “even out teacher judgments”. Only one group of teachers reported comparing their analyses and judgments against standards and exemplars.

In summary, acknowledging the small sample size, a tentative conclusion is the need for professional development on clear criteria, use of exemplars, standards and reference points against which to make consistent teacher judgments. Whilst teachers are beginning to compare their judgments with judgments of colleagues, few teachers appear to make judgments against curriculum levels, or detailed analysis of the National Standards.

To what extent are moderation processes used to ensure appropriate and consistent OTJs?

Two of the ten participating schools of this exploratory study had well established and successfully functioning moderation processes in place. Five schools were beginning to use moderation, and three schools were yet to begin systematic moderation.

It [moderation] is going to be a difficult process ... It's not easy but we could meet and talk about it and see how it goes. (TPS, T2, Nov 2010)

Moderation hasn't really happened yet. We moderate for writing but not in reading and math as it's all been a bit of a rush (AS, T1, Nov 2010)

We have moderation teams- both year levels-swapping with a mixture of teachers. (BS, T1, Nov 2010)

Our team moderates. Moderation occurs firstly in pairs [of teachers] and we bring samples you are unsure of to a meeting and discuss together as a team... We highlight children where we're a bit unsure- these are the ones we make sure we have enough evidence... we talk to children, not just look at the assessment results and we refer to the standards... (MS, T2, Nov 2010).

In summary, two of the ten participating schools used systematic moderation processes to ensure appropriate and consistent OTJs were formulated. Eight of the schools were either beginning to moderate or were yet to start any consistent moderation. On the basis of these ten schools it is likely that many schools need time and professional support to establish effective moderation processes.

End of year interviews with Principals and Lead Teachers

What do principals understand about OTJs?

Analysis of principal interviews revealed some confusion and ambiguity as to whether OTJs were to represent progress over the year or end of year summative data. The MOE definition of OTJs clearly states, "... up to a particular point in time". (MOE, 2010, OTJ factsheet)

Several principals indicated that their understanding had deepened during the year, as these interview extracts portray:

*We have had a lot of PD around what makes an effective OTJ and the process to make one. It is really important to have something written up for the process so it becomes part of the assessment schedule even though it takes more time than we anticipated. (MS, P, Nov 2010)
Grown heaps - teacher were making judgments but they weren't well informed. They have become less reliant on test tools and more aware of triangulating judgments now. (WBS, P, Nov 2010)*

When I first came here teachers were reporting to parents against nothing it was just your child's doing well. The first thing we had to do was report against standardised testing so parents had to get some honest data about where their children were at in terms of literacy and numeracy and this time we are giving them that but we are also giving them the OTJ. We learnt how to moderate and we've practiced as a whole staff. I put a newsletter out yesterday saying that this is what you're going to get but don't you know the government hasn't given us anywhere near enough time to perfect this and we're learning and this is a work in action. (TPS, P, Nov 2010).

Whilst principals indicated growth in understanding and confidence, there was still confusion amongst some principals about aggregation of data and how to reach an OTJ on the basis of both standardized test results and cumulative informal assessments.

In summary, by November most principals in this exploratory study had deepened their understanding and confidence in teachers making OTJs, but some of the principals continued to express uncertainty about the composition and effective processes to determine OTJs.

How common are principals' understandings about 'standards'?

Principal (and teacher) understandings about 'standards' and OTJs were not commonly shared. There was still considerable variation on when and how to interpret and apply 'standards', as signaled in the following range of interview extracts.

Still unsure about the difference between a snap shot and the big picture, using observation to know where a child is at and the combination of all the information that you have. (WBS, P, Nov 2010).

Teachers have got to get their heads around national standards and a whole lot of new learning and so you lose the focus on your formative assessment. It's about getting all the relevant evidence together and then making an informed decision. In the comment I sent home to parents yesterday I said that standardised testing is one part of the assessment. But it's really confusing for parents. (TPS, P, Nov 2010)

Conscious of the time needed to be put aside for teachers to have a chance to learn about it and try it. Focus on teaching as inquiry- really changed teachers practice in making OTJs. (MS, P, Nov 2010)

All schools have different interpretations against the national standard- "mucky" understanding of NS. (AS, P, Nov 2010)

Given that consistent and defensible judgments require shared (common) understandings and interpretations of OTJs and 'standards' (Sadler, 1988; Wyatt-Smith, Klenowski & Gunn, 2010) this small sampling of principals is cause for concern. It appears that more time, information and professional learning are needed in order for principals and teachers to reach more common understandings of OTJs and 'standards'.

How are principal and teacher understandings supported by clear criteria and exemplars?

Practice varied across the ten schools involved in the exploratory study. Two schools had well developed criteria, exemplars and systems for supporting staff learning and judgments. However, this practice was not widespread, with six schools beginning to develop awareness and practices, while two of the schools were yet to focus on criteria.

We talked about the importance of using a range of tools, having a common understanding across the staff, referring to rubrics and using your knowledge of the child and the everyday work that they do (MS, P, Nov 2010).

All students are being assessed against national standards and I'm saying remember to have all those documents [National Standards] open in front of you. When you are making the comments to the parents then that's when you may indicate that we have reported against national standards but in fact that's not necessary because your child has only been in the country for da da da so we do not expect your child to be working at national standards at this stage. (TPS, P, Nov 2010)

'National standards' is not a precise science it's a journey that takes twist and turns. Trial and error this year, exploring answers to issues, levelling understanding and clarity. (AS, P, Nov 2010)

Schools may need further time and professional learning support to develop awareness of the need for explicit criteria and exemplars to support teachers in making consistent judgments about student

achievement of national standards. Until this occurs judgments will be neither defensible nor consistent.

To what extent are moderation processes used to ensure appropriate and consistent judgments?

Even amongst these schools engaged in professional development contracts related to assessment (Assess to Learn), variability in the use and quality of moderation processes existed. Two of the schools were using moderation processes effectively to ensure that appropriate and consistent teacher judgments were being made. Most other schools were beginning to use moderation (usually in writing), but were yet to embed their practices. Two schools were yet to embark on moderation processes.

I can't tell. I would be lying if I said that they are consistent and appropriate this year. This is a work in progress that will continue into future years. It is good having some beginning teachers so that we can grow them into this process. Some long serving teachers not up to the mark. (AS, P, Nov 2010).

Moderating was a smooth process for this school. We are not going to moderate with the whole class every time, might just choose a few and everyone have a go. (MS, P, Nov 2010).

We don't, and I would be lying if I said we do. Because we haven't had enough time I mean it's just ridiculous. It's the best we can do in the short time frame we've had. I mean moderation doesn't have to be overnight. These national standards are so complicated because they're not standards they're aspirational goals they are flowery you know they're not specific enough you can't do a checklist so that takes time for teachers to learn that. (TPS, P, Nov 2010).

By November two of the ten schools were using moderation processes effectively to ensure appropriate and consistent OTJs were being made about student achievement in relation to NS. However, most other schools in this study would benefit from additional time and expert support to develop robust moderation processes.

Conclusion and Implications

Deep or meaningful change takes time, generally considerably longer than anticipated (Timperley, Wilson, Barrar & Fung; 2007) so it is not surprising in the first year of implementation of OTJs and NS that across ten schools, teachers and principals vary in their understandings and practice. Although two of the ten schools contained principals and teachers who understood what constituted an OTJ and within their staff had some degree of commonality in interpreting standards, they were still refining criteria and exemplars through their moderation processes. However, despite having a focus on assessment professional development, the other eight schools in the study were not in agreement about how to make OTJs, how to interpret NS nor did they have useful processes to moderate judgments.

While generalisation from this small study is limited and these results need to be treated with caution, there is nevertheless sufficient confusion and uncertainty amongst teachers in New Zealand to be concerned about the appropriateness and consistency of OTJs in 2010.

More New Zealand wide professional development attention is needed on developing common understandings of the NS and to have universally used reference frameworks supported by explicit criteria and exemplars. These aspects need to be developed within a context of continuous professional learning and moderation processes if New Zealand is to achieve appropriate, consistent and defensible OTJs.

References

- Cumming, J., Wyatt-Smith, C., Elkins, J., & Neville, M. (2006). *Teacher judgment: building an evidentiary base for quality literacy and numeracy education*. Report. Griffith University. Retrieved on 30/11/2010 from:
www.qsa.qld.edu.au/downloads/publications/research_qsa_teacher_judgment.pdf.
- Gardiner, J., Tom, C., & Wilson, K. (2000). *Consistency of teacher judgment: research report*. Brisbane: Office of Queensland School Curriculum Council.
www.qsa.edu.au/downloads/publications/research_qsc_teacher_judgment.pdf. Retrieved 4/12/2010.
- Kellis, M. & Silvernail, D. (2002). *Considering the place of teacher judgment in Maine's local assessment systems*. Maine Centre for Education Policy, Applied Research and Evaluation, University of Southern Maine.
- Maxwell, G.S. (2002). Moderation of teacher judgements in student assessment. Discussion paper on assessment and reporting. School of Education: The University of Queensland.
- Meisels, (2001). Trusting teacher judgment. Retrieved from
www.fairtest.org/trusting_teacher_judgment. 4/12/2010.
- Ministry of Education. (2010). Factsheet: Making overall teacher judgments (OTJ).
www.tki.org/factsheet/overall_teacher_judgment. Retrieved 4/12/2010.
- NSW Department of Education and Training (2008).
www.curriculumsupport.education.nsw.gov.au/consistent_teacher/consistency.htm. Retrieved 4/12/2010.
- Sadler, D.R. (1988). Formative assessment: revisiting the territory. *Assessment in Education*, 5 (1), 77-85.
- Punch, K. (2009). *Introduction to research methods in education*. London: SAGE Pubs.
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development. Best Evidence Synthesis Iteration*. Wellington, New Zealand: Ministry of Education.
- Wyatt-Smith, C., Klenowski, V., Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*. 17:1, p.59-75.