"My teacher's judgement matters more than mine": Comparing teacher and student

perspectives on self-assessment practices in the classroom

Lois R. Harris

The University of Auckland

Gavin T. L. Brown

The Hong Kong Institute of Education

Introduction

Self-assessment is considered a formative practice (Black & Wiliam, 1998;

Clarke, 2005), thought to help students engage (Woodward & Munns, 2003) , develop

self-regulation and metacognition (Andrade, 1999; Boud, 2005; Hattie & Timperley,

2007), and better understand criteria used to evaluate their work (Andrade, 2000;

Andrade & Valtcheva, 2009). In the Assessment *for* Learning framework (Black &

Wiliam, 1998), self-assessment asks pupils to evaluate their own work in relation to a

specific learning intention, goal, or set of criteria. Some advocates would consider any

kind of reflection about the quality of work to be an example of self-assessment (e.g.,

Munns & Woodward, 2006; Woodward & Munns, 2003).

While this complex social practice has significant promise (Black & Wiliam,

1998; Cowie, 2005), studies have noted problems with implementation (e.g., Cowie,

2005, 2009; Ross, 2006). Ross (2006) found teachers were concerned students would

over or undervalue their work; pupils feared peers would cheat, with some seeing

assessment not as their job, but their teacher's. In Cowie's (2005, 2009) work students

described complex issues of disclosure (both to other students and their teacher), as well

as problems relating to control and accuracy. Several New Zealand studies have also

found students to be skeptical about the utility of self-assessment practices (Brown,

Irving, Peterson & Hirschfeld, 2008; Harris, Harnett, & Brown; Peterson & Irving, 2008).

Students seem to perceive, rightly or wrongly, that only teacher-controlled assessment

matters (Bourke, 1996; Brown et al., 2009).

In line with theories of planned or reasoned behavior (Azjen, 2005), the reasons,

intentions, and beliefs educational stakeholders have about assessment are thought to

shape their actions within school environments. Hence, how teachers and students understand self-assessment is likely to affect how they enact such practices. This study sought to examine:

- How and why do teachers use self-assessment in the classroom?

- How do students understand self-assessment?

- What similarities and differences exist between students' and teachers' understandings?

New Zealand policy context

The New Zealand Ministry of Education has an explicit commitment to the use of Assessment for Learning (Ministry of Education, 2001) through the provision of resources (e.g., Assessment Tools for Teaching and Learning asTTle) and professional development (e.g., Assess to Learn Program, AtoL). Unlike the United States and the United Kingdom, New Zealand has long resisted accountability measures like state and national testing (Crooks 2002). Schools in New Zealand are self-governing and choose what assessment tools they will use; school quality is evaluated by the Educational Review Office which makes its judgements using a range of measures (e.g., school visits and observations, achievement data). There is no standardised, compulsory assessment in New Zealand until students are in Years 11, 12, and 13, when students complete national school qualifications. Hence, for primary (Years 1-6), intermediate (Years 7-8), and early high school (Years 9-10) student assessment is completely at the discretion of the school. This lack of high-stakes testing has led to conditions that would seem favorable for self-assessment and there appears to be less tension between improvement and accountability goals than in other contexts abroad (Harris & Brown, 2009).

Method

This study utilized data collected during Study 2 of the Measuring Teachers'
Assessment Practices (MTAP) project which examined how teacher conceptions of
assessment relate to their assessment practices, students' conceptions of assessment, and
pupils' academic outcomes. Three teachers were selected for this multi-method study
after participating in MTAP Study 1 where they completed the Teacher Conceptions of
Assessment (TCoA) inventory (Brown, 2006) and an hour long semi-structured interview
about their conceptions of assessment (Harris & Brown, 2009; Brown & Harris, 2009).
From the 26 possible participants from Study 1, three were selected to maximize the
diversity of the sample, taking factors like school sector, geographic location, level of
experience, and views on assessment into account. This paper examines data about self-
assessment gathered from each teacher through a week of classroom observations, a
teacher follow-up interview, and focus groups with students from the teacher's class.

The three teachers were female and worked in co-educational, public schools, but
were from diverse educational levels and came from differing geographic areas of
Auckland. Isabel taught a Year 6/7 class at a mid-socioeconomic (i.e., decile 6) primary
school in a semi-rural area in southern Auckland. She was an early career teacher in her
third year of teaching. Danielle taught a Year 7 class in a low socio-economic (i.e., decile
3) intermediate school in southern Auckland. Danielle had spent the majority of her
teaching career working in her native country, India. She had only recently completed
further study in New Zealand and this was her third year teaching in New Zealand
schools. Sylvia taught a year 10 advanced English class at a mid-socioeconomic (i.e.,

decile 6) high school in west Auckland. She was a Senior Teacher and Literacy Leader in her school and was also responsible for other Years 11 and 12 English classes.

Each classroom was visited for one calendar week. All maths and/or English lessons occurring during this week were observed and video recorded; detailed field notes were taken with particular attention paid to assessment activities. The classroom teacher helped select students with mixed achievement levels for focus group interviews; each group lasted 40-60 minutes. In all, 6 focus groups were analyzed (*n*=40), two from each class. A follow-up teacher interview, lasting approximately 1½ hours, was conducted on the Friday after all other data had been collected. All interview and focus group recordings were transcribed verbatim; self-assessment events within the video recordings were also transcribed and triangulated with field notes and artifacts collected during observations.

Each teacher and her students were considered a unique case (Yin, 2009). Categorical analysis (Coffey & Atkinson, 1996) was used to identify and code data relating to self-assessment. Within case analysis took place initially (i.e., particular interview, observation, or focus group transcript), then these data were systematically triangulated with other data from within the same case to examine how the practice was used and understood by the teacher and her students. Finally, cross-case synthesis (Yin, 2009) was conducted to examine similarities and differences across cases.

Results

Within this study, teachers described using student self-assessment to gain information about how students thought they were progressing and frequently used these data to make instructional decisions. Students in all classes expressed skepticism about

the validity of self-assessment, although a few saw it as having potential benefits. In this

section, individual cases will be examined in turn; cases will be compared and contrasted

in the subsequent discussion section.

Isabel

All three teachers were observed using differing forms of self-assessment. Isabel

described self-assessment in her Year 6/7 classroom, saying:

> I get them to mark it themselves and I'll get them to self-assess on how well they
> did, something in accordance with what the whole learning intention was. And I'll
> prompt them like 'how well did you do this' or 'did you not do so well at this' and
> that'll feed their comments and then I'll just initial it and then say "I agree' or 'I
> think you're being too hard on yourself.'

This practice generally occurred when criteria were objective (i.e. for spelling, grammar,

handwriting) and after answers or examples had been provided. Typical student

comments were, "I think I did alright with this activity" and "I did well, but I think I can

do better next time;" these generally evaluated task difficulty or personal effort. Only

occasionally did comments discuss accuracy (i.e. "I did well with sequencing even

though I got 5 wrong they were only the tricky ones and I tried which is good"). Isabel

explained self-assessment was:

> Just to see how well they think they are [doing]…. it pulls them down quite a bit
> because they're like "I have to be honest here" and it works really well. It's
> [assessment's] a two way thing. It's not just me doing it all. It also cuts my
> marking down a heap of a lot.

Here, she talked about using self-assessment for her own evaluative purposes, although

acknowledging it also let assessment be "a two way thing" and forces kids to be honest.

Additionally she notes a pragmatic purpose for use of this practice as it 'cuts marking

down.'

Isabel's students agreed that self-assessment was mainly so adults could understand their perceptions of progress, saying:

> Howard: Cause the teacher wants to know how well you do, how well you think you do.
> Frances: And also when your book gets sent home at the end of the year, your parents, it's good for your parents because they know how you feel about that.

One pointed out, "If she says you guys have to get a bit more detailed for your comments, sometimes I just make stuff up," demonstrating that pleasing the teacher was a main purpose.

Isabel's students were clear that they did not see this practice as being assessment; to them, assessment was tests and other teacher evaluated, formal processes. Some students struggled to see the point of self-assessment; as one noted "I don't really care what I think about what I've done. I just think about what Mrs. Cooper thinks and what my friends think." While students were happy to take part in self-assessment activities, they didn't appear to see their own self-regulation as a major goal.

Danielle

The second teacher, Danielle, commonly used a practice called Traffic Lights, developed out of the King's-Medway-Oxfordshire Formative Assessment Project in England (Organization for Economic Co-operation and Development & Centre for Educational Research and Innovation, 2005), to help her Year 7 students assess their understandings. At the end of small group maths lessons, students would rate their understanding: green meant they felt ready to move on, orange indicated they needed more practice, while red showed complete lack of understanding. Danielle explained:

And when they come to me and show it to me, it's always registered in my head. And I know where they are. Today Jim was orange. He was not green. So I asked him to stay down on the mat.

Information from these self-assessments helped Danielle make instructional decisions; she didn't explicitly discuss any student purposes.

Danielle's students were aware that she was their audience. As one student noted:

I think it's good that she tells us to do traffic lights. It's for her to know where we're at and if we need some help and things like that.

However, when asked if they frequently marked down red, they responded no and explained:

Chelsea: Because if you put red, you feel like she might tell us off or something.
Natasha: She'll tell us to listen more or something, like we weren't listening.
Michael: She always turns it against us if we're not learning.

While they said that with previous teachers they would have been more honest, in this class, they feared the consequences of admitting that they did not understand something.

Additionally, some students were afraid of what peers might think. One student explained:

Sally: Yesterday, we were learning to simplify fractions and I couldn't actually understand what she [the teacher] was saying so I don't actually know how to simplify fractions.
Researcher: That's okay. Have you told Mrs. Johnson that? Did you put red in your traffic light?
Sally: I did orange.
Researcher: And why did you do orange?
Sally: Um, Jason was looking at my work and if I did red, he'd think I was dumb, so I just did orange.

Here, due to concerns about peer perceptions, the student put down what she knew was an inaccurate self-assessment in order to 'save face.'

Danielle confirmed that her students seldom put 'red' in their traffic lights and explained:

> …if I get a green, I'm telling myself my lesson was planned well. If I get a red,
> that's telling me 'Danielle, you're going off track. Hold it, caution, that's a
> caution, that's a hazard there."

She said that when students put down red, "I feel I'm not making maths reachable for them. Approachable, enjoyable, relevant." In this particular case, she described student self-assessment responses as a measure of her lesson's successfulness that reflected on her as a teacher.

Like Isabel's students, they didn't consider self-assessment to be 'assessment' per se as it was informal and didn't result in a mark or score. While several students noted positives about self-assessment (i.e., it could help students identify mistakes, improve on virtues like honesty, and create an avenue of communication with their teacher), in the main, Danielle's pupils appeared concerned to admit mistakes, with some describing events where they purposely wrote down inaccurate evaluations of their progress. This case highlighted that fundamental to self-assessment is a classroom environment where students feel comfortable with teachers and peers enough to respond honestly about their progress. Additionally, it showed that students react as individuals to self-assessment and perceive the possibilities and threats differently.

Sylvia

The third teacher, Sylvia, said she did not do much self-assessment with her year 10 English students, articulating concerns that students would over or undervalue their work, instead preferring peer assessment and other formative strategies. She noted, "I think self-assessment works best when you have no doubt about the criteria, um, and I don't know that that's that easy in English because there's always a little bit of doubt in

the criteria." For self-assessment, she used a KWL chart to get students to identify what

they already knew and what they needed to learn about their current topic, film review,

"hoping that they were engaging mentally with what they did and didn't know." She

explained, "… I had material that I could have given them but I wanted to select that

material based on what they wanted or needed."

> None of her students considered this KWL activity as assessment, explaining:
>
> Allison: It's just to find out what we know and we don't know
> Kelly: She doesn't grade us on it
> Allison: Yeah, it's not like, we're not getting a mark on it
> Joe: Just a question answer thing
> Interviewer: Question answer?
> Dillon: Just so she knows what she has to teach us
> Prama: It like a survey type thing

This activity was seen as information gathering for the teacher; it was not considered

assessment as it did not result in a mark. Sylvia's students didn't appear to value self-

assessment saying:

> William: Because it doesn't matter. The teacher can do it.
> Nancy: Cause you mark yourself easy
> Cathy: Cause also for that one, you might think you are better than what you
>   actually are. [You might] Not think you're doing anything bad, but really
>   the teacher looks at it different.
> Hugh: You might think it's good as, but the teacher.
> Cathy: The teacher might have a different opinion to you.
> Interviewer: So do you think your judgement of your work matters or only your
>   teacher's?
> Nancy: I think my teacher's judgement matters more than mine
> Hugh: You base your judgement on what the teacher thinks, so it's pretty much
>   the teacher's judgement that really counts.

Her students appeared to view assessment as a teacher controlled domain where their

judgements didn't really matter. This view may be accurate in this context as ultimately

only teacher judgements do count towards their grades. Additionally, students raised

concerns over the accuracy of student judgements. Like the other two cases, Sylvia's

students didn't described self-assessment as increasing their ability to be more self-regulated learners.

Discussion

While this study only examined three cases of PASA use within the classroom, limiting the generalizability of these results, there are patterns which emerged across the cases which can give pause to those advocating these kinds of formative practices. All teachers were trying to implement self-assessment and use the data in formative ways (i.e., to improve teaching and learning). However, neither they nor their pupils described self-assessment as being about students developing their abilities as reliable assessors of their own work. It is questionable if the practices Sylvia and Danielle described as self-assessment should be considered 'assessment' at all because students lacked explicit criteria to use when making judgments; Isabel's use of self-assessment sometimes included criteria, but still lacked precision. Perhaps students can rightly refer to the observed practices as 'surveys' since teacher information gathering appeared to be a primary purpose of these data. However, the accuracy of such surveys must be questioned as these data are far from anonymous and in some classrooms, students perceive that 'incorrect' responses could lead to negative personal consequences.

Additionally, students across cases consistently reported threats to self-assessment. First, there were issues of disclosure both with teachers and peers who may inadvertently view these assessments. Clearly a high level of trust is necessary for honest self-assessment to occur. Second, there were issues of accuracy reported, with students genuinely concerned about their abilities to assess correctly. Students need explicit

instruction in how to assess their work and must develop confidence in their skills if they are to be comfortable in an assessor role. Apathy was another serious threat, with students across cases saying they didn't care what they think of their work, preferring teacher or peer judgments. This apathy may be rooted in a lack of understanding of how self-assessment skills will personally benefit their learning. It is important to note, however, that these threats were not perceived equally among students within and between cases; hence it is clear that student perceptions and experiences of self-assessment are highly contextualize on a personal level and implementation may need to be individually tailored to minimize these threats.

These three cases appear to be examples of what Shepard (2006, p 641) describes as "adopting a formative "technique" without a corresponding philosophical shift [which] is likely to undermine efforts by leaving in place traditional attitudes." While all three teachers were trying to implementing formative practices, they used self-assessment primarily for teacher purposes. Additionally, their students retained a mindset in which assessment is viewed as tests and grades. The self-assessment practices observed in this study might be better classified as collaborative pedagogical practices, rather than assessment. This would be consistent with Black and Wiliam (2006) who hinted that Assessment for Learning is more than a theory of assessment, but rather an incipient theory of pedagogy. The student judgment that these self-assessment activities are not actually assessment may be accurate. Hence, more work is required to clarify to teachers and teacher educators what is necessary to make self-assessment an 'assessment' activity. Additionally, teachers must make this clear to their students so pupils understand how self-assessment can help them become more self-regulated learners.

Also, as highlighted in Danielle's case, particular self-assessment practices may not be viable in all classrooms. Good classroom relationships between teacher, students, and their peers appear paramount. Without a classroom philosophy that views mistakes as an opportunity for learning and that encourages honest reflection, self-assessment may become a dishonest exercise where students write what they think others want to hear. Hence, teachers must think strategically about when and how to incorporate self-assessment into their classrooms. There are some classes where, due to teacher-student dynamics, self-assessment may not be a viable option. In other scenarios, it may be best utilized later in the year when positive relationships have had time to be built.

Within these data it appears students still see their teacher as the preferable source of feedback (something teachers and schools reinforce through grading practices), making them struggle to understand and value the purpose of self-assessment. As Nicol and Macfarlane-Dick (2006, p. 200) noted, "if formative assessment is exclusively in the hands of teachers, then it is difficult to see how students can become empowered and develop the self-regulation skills needed." This study is not invaliding self-assessment, but critiques how it is often enacted in schools. Clearly, students and teachers need training not only in the techniques of formative assessment, but also in the theories that underpin them if implementation is going to result in students becoming more self-regulated learners.

References

Ajzen, I. (2005). *Attitudes, personality and behavior* (2nd ed.). New York: Open University Press.

Andrade, H. G. (1999). *Student self-assessment: At the intersection of metacognition and authentic assessment.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5), 1-7.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice, 28*(1), 12-19.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81-100). London: Sage.

Boud, D. (1995). *Enhancing learning through self-assessment*. London: Kogan Page.

Bourke, R. (1996, December). *But is it right? Students' conceptions of assessment.* Paper presented at the Annual Conference of the New Zealand Association for Research in Education, Nelson, NZ.

Brown, G. T. L. & Harris, L. R. (2009). Unintended consequences of using tests to improve learning: How improvement-oriented resources heighten conceptions of assessment as school accountability. *Journal of Multidisciplinary Evaluation, 6*(12), 68-91.

Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. *Psychological Reports, 99*, 166-170.

Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. F. (2009). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning & Instruction, 19*(2), 97-111.

Clarke, S. (2005). *Formative assessment in the secondary classroom*. London: Hodder Murray.

Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies*. London: Sage.

Cowie, B. (2005). Pupil commentary on assessment for learning. *Curriculum Journal, 16*(2), 137-151.

Cowie, B. (2009). My teacher and my friends help me learn: Student perspectives and experiences of classroom assessment. In D. M. McInerney, G. T. L. Brown & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about Assessment for Learning* (pp. 85-105). Charlotte, NC: Information Age Publishing.

Harris, L. R., & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy and Practice*, *16*(3), 365-381

Harris, L. R., Harnett, J. A., & Brown, G. T. L. (2009). "Drawing" out student conceptions: Using pupils' pictures to examine their understandings of assessment. In D. M. McInerney, G. T. L. Brown & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about improving school outcomes* (Vol. IX, pp. 53-83). Charlotte, NC: Information Age Publishing.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Munns, G., & Woodward, H. (2006). Student engagement and student self-assessment: the REAL framework. *Assessment in Education: Principles, Policy & Practice, 13*(2), 193-213.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Peterson, E. R., & Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*, 238-250.

Organization for Economic Co-operation and Development, & Centre for Educational Research and Innovation. (2005). *Formative assessment : improving learning in secondary classrooms*. Paris: OECD.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation, 11*(10), 1-13.

Shepard, L. A. (2006). Classroom assessment. In R. L. Brennan (Ed.), *Educational Measurement* (pp. 623-646). Westport, CT: Praeger.

Woodward, H., & Munns, G. (2003). *Insiders' voices: Self-assessment and student engagement.* Paper presented at the New Zealand Association for Research in Education and Australian Association for Research in Education Joint Conference, Auckland.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Los Angeles: Sage Publications.