The effectiveness of formative assessment:
Does peer response improve the students writing?

Stephanie Dix,
School of Education,
The University of Waikato

The changes in New Zealand assessment policy have led to a greater emphasis on externally referenced assessment, and while standardised assessment has a valid purpose in terms of looking at trends in national and school achievement, it does not guarantee more effective learning

Introduction

It is recognised that "the intention of most educational systems is to help students, not only grow in knowledge and expertise, but also to become progressively independent of the teacher for lifelong learning" (Sadler, 1998, p.80). One factor that supports students' independence and control over their own learning is high quality formative assessment, especially when constructive feedback is provided. The value of formative assessment is well recognised, and as such the recent New Zealand assessment policy changes have raised concerns in some areas (Hill, 2000; Lee & Lee, 1998). The changes in New Zealand assessment policy have led to a greater emphasis on externally referenced assessment, and while standardised assessment has a valid purpose in terms of looking at trends in national and school achievement, it does not guarantee more effective learning, as indicated in the research on formative assessment (Black & Wiliam, 1998a; Cowie & Bell, 1999; Harlen, 1998; Sadler, 1998).

This article investigates the literature and focuses on peer response as a procedure that may help students improve the quality of their written messages.

Peer response played an important part in writing programmes during the 1980s and 1990s (Calkins, 1991; Elbow, 1998; Graves, 1983; Philips & Ward, 1992). From this researcher's work in schools there is little evidence of peer response in current writing programmes.

The Literacy Task Force Report (Ministry of Education [MoE], 1999) set clear expectations for year 4 writers. There are students who struggle to meet the criteria set down in this report (Dix & Ward, 2001).

Current pedagogy reflects the functional or genre based approach to writing as established in the New Zealand English Curriculum (MoE, 1994a). This means that students are often asked to write to criteria and are assessed against criteria. Adequacy of content, purpose, writer's voice and authenticity of the task are often ignored.

The National Monitoring Project Report (NEMP) on *Writing Assessment Results, 1998* (Flockton & Crooks, 1999) indicated that students at year 4 and year 8 made few meaning based changes to improve their writing. Most of the changes made were surface changes.

Peer response/peer assessment could encourage writers to re-evaluate and revise their writing.

Research on peer assessment and peer response

In both the research domains of assessment and written language there is little research data on either peer assessment or peer response. While there is a growing amount of research on the purposes and value of formative assessment, the focus has not been on the effectiveness of peer assessment in the classroom. Likewise, peer response in the written language literature is described as a procedure that can help writers clarify their writing. However, there appears to be a limited amount of research data regarding its effectiveness. What follows is an analysis of the literature, firstly from an assessment perspective and secondly from a literacy perspective.

Assessment perspectives

It is widely recognised that assessment plays an integral and essential role in identifying the learner's knowledge, skills and understandings in order to plan for further learning (Black & Wiliam, 1998b; MoE, 1993; MoE, 1994b; Sutton, 1995).

Stephanie Dix

For example, *The New Zealand Curriculum Framework* (MOE, 1993) states that:

> Assessment of individual students' progress is essentially diagnostic. Such assessment is integral to the learning and teaching programme. Its purpose is to improve teaching and learning by diagnosing learning strengths and weaknesses, measuring students' progress against the defined achievement objectives, and reviewing the effectiveness of teaching programmes (p. 24).

While in 1993 the New Zealand Ministry of Education recognised the multiple purposes of assessment, international emphasis has been on standards-based, externally referenced assessment (Johnston, Guice, Baker, Malone, & Michelson, 1995; Messenheimer & Packwood, 2002; Tierney, 1998); that is, assessment that measures educational outcomes in order to determine whether acceptable standards are met. This emphasis on outcomes based learning has concerned New Zealand educationalists because recent Ministry policy requirements have also demanded that a greater emphasis be placed on measuring students' progress against externally referenced assessment procedures (Hill, 2000; Lee & Lee, 1998; MoE, 1998; MoE, 2002).

Black and Wiliam's (1998a) in-depth literature search strongly indicates that it is formative assessment that makes the difference in raising achievement standards. Greatest gains in learning were evident when high quality formative assessment was implemented. Although formative assessment has been defined in a variety of ways, it is accepted that its common purpose is to improve and inform the student's learning, rather than just measure and record it (Cowie & Bell, 1996; Harlen, 1998). In *Assessment: Policy to Practice,* the New Zealand Ministry of Education (1994b) states that the purpose of formative assessment is

> to provide the student with feedback to enhance learning and to help the teacher understand students' learning. It helps build a picture of a student's progress, and informs decisions about the next steps in teaching and learning (p. 8).

Research on formative assessment identifies the importance of constructive feedback for the student learner with the purpose of improving learning. In peer assessment it is the students who are making decisions about other students' work. This differs considerably from peer tutoring, group assessment, and self–assessment purposes and procedures. While peer assessment can be implemented anonymously, using a range of "assessors to assess" and can be used to assess summatively, Race (2001) stated that:

> Peer assessment can also pay dividends when used in a purely formative way, where the real purpose is to allow students to gain feedback from each other, and any scoring or grading is just a means towards this feedback rationale. (p.4)

The MoE (1994b) further defines peer assessment as "assessment carried out by a student's peers following some agreed format and process (p. 49). Their statement in *Assessment: Policy to Practice* acknowledges the social skills that are inherent in effective peer assessment, but the document does not provide any in-depth rationale or justification for peer assessment. Much of the discussion centres on group collaborative work, which may or may not involve peer assessment feedback.

To investigate the rationale for peer assessment it is necessary to review the underlying principles of formative assessment. The first key principle is that it is essential that formative and peer assessment reflect the theoretical beliefs that underpin teaching and learning in the classroom (Black & Wiliam, 1998a; Harlen, 1998; Johnston et al., 1995). Harlen (1998) emphasises the need for assessment practice to be embedded in constructivist pedagogy. This essential principle is often overlooked as teachers attempt to implement formative and peer assessment procedures in their classrooms, and fail to acknowledge the central role of learners as those who "must be ultimately responsible for their learning as no-one else can do it for them" (Harlen, 1998, p. 3).

Constructivist learning theory requires learning to take place in meaningful contexts to ensure relevance and ownership of the learning (Harlen & James, 1997). The learner is expected to interact with the knowledge, to reflect on, debate and discuss the material to ensure that understanding, or 'deep learning' as opposed to surface learning, has taken place (Entwistle & Ramsden, 1983). The student in the constructivist classroom does not take a passive role (Biddulph & Carr, 1999). The learner is actively involved in constructing and reconstructing his or her own knowledge building on his or her prior understandings and making links to new experiences (Pearson & Spiro, 1982).

An issue concerning peer assessment and constructivist learning theory is that cognitive demands are placed on the learner, who must employ complex thinking skills to make judgements about the understanding or prior knowledge of his or her peers and to identify learning needs.

The second key principle that influences formative assessment is the effectiveness of 'constructive feedback.' Black and Wiliam (1998b) found that if feedback is continuous, immediate and specific to the task and the learner, it is more useful than comparative data or numbers that do not describe the learning. Harlen (1998) has described formative assessment as involving three interactive processes: gathering and collecting data; interpreting the data; and using it. The quality of these interactive processes is based on the interrelationships between the student learners, the teacher, and the knowledge that 'sits' behind the task (Black, Harrison, Lee, & Wiliam, 2001). The literature places a high priority on the teacher's subject knowledge, judgement making, and relationships between the teacher and learner (Cowie & Bell, 1999; Black & Wiliam, 1998a; Crooks, 1988; Gipps, 1994; Harlen & James, 1997; Sadler, 1998). These researchers have identified the teacher as having a key role in making assessment decisions.

This raises several issues for the effectiveness of peer assessment. First, as Sadler (1989) has maintained, "teachers' conceptions of quality are typically held, largely in unarticulated form, inside their heads as tacit knowledge" (p. 126). This is difficult to do for students who are still developing tacit knowledge. "A novice is, by definition, unable to invoke the implicit criteria for making refined judgements about quality. Knowledge of the criteria is caught through experience, not defined" (Sadler, 1989, p. 139).

Second, students have responsibility for providing feedback to their peers. This means they must fulfil the dual role of being the expert and the responder who gives constructive feedback, and the learner who recognises the value of the feedback and acts on it. In the role of expert and responder the student must provide constructive feedback to his or her peers. This requires evaluative and reflective thinking skills, and a working knowledge of the subject. Sadler (1989) has pointed out that students need to be able to compare objectively the actual levels of performance with desired standards. One way to do this is to follow an analytic approach and have the teacher provide criteria considered most relevant to the task. Another way to do this is to react to the work as a whole and make global judgements. As students are still developing a working knowledge of the subject Sadler (1989) perceives this approach as more difficult for the student.

The second role assumed by the student is that of the learner receiving constructive feedback. The learner must possess a concept of the standard required. This means the student must hold a concept similar to that held by the teacher. More importantly learners must act on the feedback to close the gap. They must not only desire to alter the gap but also have relevant skills and strategies to do so (Black & Wiliam 1998a; Harlen, 1998; Sadler, 1989, 1998).

The third key principle that influences formative and peer assessment concerns the learners themselves, their interactions with each other, the accessibility to perform the assessment task, and their desire to improve and be involved in the assessment procedure. Crooks (1988) maintaines that class evaluative activities have significant effects on students. Differing peer and social interactions and attitudes to learning also clearly influence learning acquisition (Wilkinson, Hattie, Parr, & Townsend, 2002). Stahl (1994) states that students working co-operatively on learning tasks

> tend to have higher academic test scores, higher self esteem, greater numbers of positive social skills, fewer stereotypes of individuals of other races or ethnic groups, and greater comprehension of the content and skills they are studying (p.1).

Crooks (1988) also discusses the importance of students having high 'self-efficacy'; that is, students' perceptions of their capability to perform certain tasks. This aspect appeares to have a strong influence on their effort and persistence with difficult tasks. The research literature indicates that cognitive learning is affected by the students' emotional and social factors. Further factors that warrant consideration include peer interactions, power relations, and the notion of friendship and help seeking (McCarthey & McMahon, 1992; Zajac & Hartup, 1997).

Written language perspectives

Independent writers are perceived as those writers who can communicate effectively using the written word and have the knowledge, skills and strategies to evaluate and develop their own writing (Carruthers, Phillips, Rathgen, & Scalen, 1991). Just as students need to know about their own specific skills and abilities, what they do well and where their writing needs further development, the writer also requires a response and evaluative assessment feedback that will support new learning. Carruthers et al. (1991), in summarising the range of writing assessment, stated that:

> Writing can be assessed in a variety of ways: formally, using stated criteria, or informally, based upon casual observation; during the process of writing or once a

piece has been handed in for marking; by the teacher, a peer or the writer him or herself- perhaps even by someone unknown to the writer; using impressionistic (holistic) or analytic methods, such as rubrics (p. 91).

The appropriateness of the procedure, however, depends on the purpose of the assessment. A grade, mark, or allocated level of achievement may be recorded as summative information for school records or when reporting to parents, nevertheless, a quantitative number offers little value for enhancing learning.

Constructive feedback is recognised as an effective strategy for improving the quality of writing (Calkins, 1991; Elbow, 1998; Graves, 1983; Hood, 1997: Phillips & Ward, 1992), and should be part of the teaching, learning and assessment cycle. Feedback, or responsive assessment as described by Carruthers et al. (1995), involves a sensitive appreciation of writing development. It requires careful listening on the part of the responder when the writer describes what has been included or excluded. This type of formative assessment became part of the "process writing approach" (Graves, 1983) implemented as writing conferences in the 1980s and early 1990s. It reflected the constructivist learning theory and sociocultural beliefs of the time, which acknowledged the ideals of writers constructing their own knowledge, learning at their own pace, learning through talk and interactions with others, and learner choice and ownership. The teacher provided scaffolding and facilitated learning for growth and independence.

The purpose of responsive assessment was to: provide feedback at any stage of the writing process; acknowledge the writers' views rather than the assessor's; leave content in the hands of the writer but teach revision skills; affirm what was being done but also provide a basis for change or improvement; and empower the writer by providing options and choices to change his or her writing. While there is research on teacher conferencing (Graves, 1983; Murray, 1982) and responding to children's writing, it must be noted that this research was written in the 1970s and 1980s. There is little specific data on the role and effects of peer response during writing time.

Elbow (1998) discusses two types of feedback as being important for writers, criterion-based feedback and reader-based feedback. To give criterion-based feedback the reader/responder must act like an expert, to stand off to one side and focus on the technique. The readers notice things they could otherwise miss as the process forces the criteria to be conscious and visible. Questions therefore focus on the quality of the content, the ideas, perceptions and point of view. Criteria-based feedback focuses on structure and organisation, the effectiveness of language, and inappropriate errors. It tells the writer how their writing measures up. Elbow (1998) offers the warning, however, that:

Conscious criteria can also be a screen between readers' words - a filter which keeps the reader from contacting and experiencing your words directly - leading them instead just to compare your words to a model, hold them up against a template, check off categories against a list (p. 250).

On the one hand criterion-based assessment is based on ideals or perfect models. Furthermore, Elbow (1998) maintains that readers often think they have to go looking for errors. This could be a concern for peer response.

Reader–based feedback, on the other hand, allows the reader to read for enlightenment or pleasure. It encourages the reader to become immersed and respond to ideas as a whole and the effect they have on them as a reader, to be aware of nuances that are difficult to categorise. Elbow (1998) regards reader-based feedback as being the most useful and efficient form of feedback, as it can lead to the fastest and most pervasive improvement in the text. "It is more apt to speak to the root causes of strength and weakness in your writing, not just the surface effects" (p.248). A major concern of reader-based feedback is however, that it can lead to narrow reading of the writing if the response is only an explanation of how the writing affects the individual reader.

These two forms of response have implications for peer assessment. Elbow (1998) has suggested that both types of responses have a role to play. Sadler (1989) by comparison, has considered student writers as unable to judge holistically; they require criteria to respond to. Criterion based assessment became the norm in the 1990s, emerging along with the genre approach to teaching writing.

Gere and Stevens' (1985) research investigated how oral language response in writing groups shaped revision. The young writers in the research study shared and responded to each other's text following a set pattern. Their comments were fed back, in turn, to the writer who had the option to elaborate, take on suggestions, or disregard them.

The research revealed three important findings. First, the type of responses that students gave were mostly content responses. The writers were challenged to clarify and provide more detail. Directing responses gave the author suggestions on what he or she could do to improve the writing. Second, the ability to act on the responses varied throughout the age groups. The fifth graders' revisions were simple and usually involved adding in. By contrast, the eighth graders debated language and worked for greater preciseness. The high school writers made more complex revision changes. Their deeper, more interactive, dialogue involved justification and elaboration, which often led to meaning-based or macro level changes.

The third finding identifieddifferences between peer and teacher responses. On the one hand teacher responses were highly generalised, lacked focus, except in the case of mechanics (spelling, punctuation), and gave few directions for rewriting. They were often vague, and it was difficult to determine whether the responses focused on content or form. Teacher responses showed little genuine reaction to the writing. Students, on the other hand, gave specific and focused feedback as well as explicit directions for rewriting. These differences were

described as providing 'different feed back functions.' Peer group response was immediate and addressed the writing during the writing process, when ideas were still being formed and clarified. Teacher responses, however, were written on the finished piece as summary statements with a formative intention of helping improve the writing. Gere and Stevens (1985) believe that:

> What group response is trying to form, then is actual text, one which communicates the meaning students find inherent in the text presented, a meaning which is often compounded of a variety of questions, comments, and criticisms of quite different interpreters who may each find a different meaning. What teacher response is trying to form is an ideal text, one which possesses certain abstract features of writing quite independently of any meaning (p.103).

The value of peer response as a formative assessment strategy is recognised in the following ways:

- Students who evaluate their own and others' writing learn to identify key aspects of quality writing
- There is transfer of learning
- Students are provided with a range of models of writing
- The readers who respond offer a variety of options; they generate other possibilities for the writer
- When peer response is ongoing the writer is provided with, and often takes greater risks
- Feedback should first focus on meaning, adequacy of information, and clarity of ideas.

This literature review presents the argument that if writers are going to respond to each other's writing, they need the skills to do so. Not only do they require the social skills for interacting positively and sensitively, but they also need to have the language to talk about writing and the knowledge to discuss the sub skills of composing and revising text. Teachers must continue to up-skill and extend their own knowledge. They do this through writing workshops, becoming writers themselves, and by sharing, discussing and assessing writing samples. Teachers must focus on the writing process as well as the final product. Furthermore, writing requires a reader's response throughout the process if meaning-based changes or higher order thinking revision is to take place. Sound teacher modelling and demonstrations are imperative before setting up peer conferences. Carruthers et al. (1991) "believe the teacher's main role is to open up the craft of writing, to strip it of its mystery" (p.94). They further believe that there should be dialogue around text and that the teacher should provide constructive ways to help writers respond to each other's text.

Summary

The assessment literature identifies shifts in assessment policy. There is a need to further investigate formative assessment procedures that will add to learners' knowledge and give them greater independence in preparing them for life long learning. When students are aware of the expected standards and are taught the skills and strategies to act on constructive feedback they will truly gain 'deep learning' – learning... with understanding. The literature identifies the value of criterion-based assessment to support focused judgements and feedback.

The literacy literature identifies peer response as an effective way of engaging dialogue around text. It provides opportunities for writers to be exposed to a range of writing models and specifically gain feedback that is focused on their own work. When presented with a range of suggestions, the writer can act and incorporate these possibilities to improve the writing. The literature identifies the value of reader feedback and making global judgements when responding to the strength and quality of the writing. More recent changes in teaching approaches to writing have led to the implementation of functional or genre based programmes. These not only present fewer opportunities for peer response, but also reflect the current emphasis on criterion-based assessment. This review emphasises the need for further investigation of assessment and peer response as viable teaching practices in classrooms today.

References

Biddulph, F., & Carr, K. (1999). Learning theories and curriculum. *Teachers and Curriculum, 3*, 31-35.

Black, P., Harrison, C., Lee, C., & Wiliam, D. (2001, April). *Theory and practice in the development of formative assessment.* Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in education: Principles, policy and practice, 5*, (1), 7- 74.

Black, P., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment.* London: Kings College.

Calkins, L. (1991). *Living between the lines*. Portsmouth: Heinemann.

Carruthers, A., Phillips, P., Rathgen, E. & Scanlen, P. (1991). *The word process*. Auckland: Longman Paul.

Cowie, B., & Bell, B. (1996, June). *Validity and formative assessment in the science classroom.* Invited keynote paper to the symposium on Validity Assessment, Dunedin.

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in education: Principles, policy and practice, 6*, (1), 101-116.

Crooks, T. (1988). The impact of classroom evaluation practices on students *Review of Educational Research, 58*, (4), 438-481.

Dix, S., & Ward, R. (2001) *Teachers and Curriculum, 5, 68-71.*

Elbow, P. (1998). *Writing with power: Techniques for mastering the writing process*. Oxford: Oxford University Press.

Entwistle, N., & Ramsden, P. (1983). *Understanding student learning*. London: Croom Helm.

Flockton, L., & Crooks, T. (1999). *Writing assessment results, 1998.* (National Education Monitoring Report No. 12). EASU. Dunedin: University of Otago.

Gere, A., & Stevens, R. (1985). The language of writing groups: How oral response shapes revision. In S. Warshauer Feedman (Ed.), *The acquisition of written language: Response and revision* (pp.85-105). New Jersey: Ablex.

Gipps, C.V. (1994). *Beyond testing: Towards a theory of educational assessment.*London: Falmer Press.

Graves, D. (1983). *Writing: Teachers and children at work.* London: Heinemann.

Harlen, W. (1998, December). *Classroom assessment: A dimension of purpose and procedures.* Paper presented at the annual conference of the New Zealand Association for Research in Education, Dunedin.

Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, policy and practice, 4,* (3), 365-379.

Hill, M. (2000). *Remapping the assessment landscape: Primary teachers reconstructing assessment in self managing schools.* Unpublished doctoral thesis, University of Waikato.

Hood, H. (1997). *Left to write too: Developing effective written language programmes for young learners.* Auckland: Berkley Curriculum Publishing.

Johnstone, P., Guice, S., Baker, K., Malone, J., & Michelson, N. (1995). Assessment of teaching and learning in 'literature- based' classrooms. *Teaching and Teacher Education, 11,* (4), 359-371.

Lee, H., & Lee, G. (1998, December). *Raising 'National Standards': Some Inconsistencies in the Green paper on Assessment for Success in Primary Schools.* Paper presented at the annual conference of the New Zealand Association for Research in Education, Dunedin.

McCarthey, S., & McMahon, S. (1992). From convention to invention: Three approaches to peer interactions during writing. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning.* Cambridge: Cambridge University Press.

Ministry of Education. (1993). *The New Zealand Curriculum Framework: Te Anga Marautanga o Aotearoa.* Wellington: Learning Media.

Ministry of Education (1994a). *English in the New Zealand Curriculum.* Wellington, New Zealand: Learning Media

Ministry of Education (1994b). *Assessment policy to practice.* Wellington: Learning Media

Ministry of Education. (1998, May). *Assessment for success in primary schools: Green paper.* Wellington: Learning Media

Ministry of Education. (2002). *Sharpening the focus: Kia hangai te titiro. A newsletter for boards of trustees, principals, and teachers on school management and administration,* 7, 1-4.

Messenheimer, T., & Packwood, A. (2002). Writing: The state of the state vs. thestate of the art in English and American schools. *Reading Literacy and Language, 36* (1), 11-15.

Murray, D. (1982*). Learning by teaching: Selected articles on learning and teaching.* Portsmouth: Heinemann.

New Zealand Literacy Taskforce. (1999). *Report of the Literacy Taskforce.* Wellington: Ministry of Education

Pearson, P., & Spiro, R. (1982). The new buzz word in reading is schema. *Instructor, 91 (* 9), 46-48.

Phillips, D., & Ward, R. (1992). Responding to writers and their writing: Recording and reporting. In R. Ward (Ed.), *Readings about writing: A resource book for supporting the developing writer* (pp. 37-53). Palmerston North: Dunmore Press.

Race, P. (2001). *A briefing on self, peer and group assessment.* York: Learning and Teaching Support Network (LTSN) Generic Centre.

Sadler, D. (1989). Formative assessment and the design for instructional systems. *Instructional Science, 18,* 119-144.

Sadler, D. (1998). Formative assessment: Revisiting the territory. *Assessment in Education, 5* (1), 77-84.

Stahl, R. (1994). *The essential elements of co-operative learning in the classroom.* Retrieved June, 13, 2002, from http://www.ed.gov/databases/ERIC- Digests

Sutton, R. (1995). *Assessment: A Framework for teachers.* London: Routledge.

Tierney, R. (1998). Literacy assessment reform: Shifting beliefs, principled possibilities, and emerging practices. *The Reading Teacher, 51,* 374-390.

Wilkinson, I., Hattie, J., Parr, J., & Townsend, M. (2000, June). *Influence of peer effects on learning outcomes: A review of the literature.* Auckland: Auckland Uniservices Ltd.

Zajac, R., & Hartup, W. (1997). Friends as co-workers: Research review and classroom implications. *The Elementary School Journal, 9*