ELSEVIER

Feedback to writing, assessment for teaching and learning and student progress

Judy M. Parr *, Helen S. Timperley [1]

Faculty of Education, The University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142, New Zealand

Available online 11 June 2010

Abstract

Traditionally, feedback to writing is written on drafts or given orally in roving or more formal conferences and is considered a significant part of instruction. This paper locates written response within an assessment for learning framework in the writing classroom. Within this framework, quality of response was defined in terms of providing information about: (i) where students were positioned relative to the performance desired; (ii) about key features of the desired performance, and (iii) what was needed to achieve the desired performance. A study of teachers (maximum $n = 59$) in six schools provided data regarding their ability to give quality formative written feedback to a piece of writing. This quality score related significantly to gains in their students' achievement on a nationally standardised measure of writing ($r = .685$, $p < .01$). The ability to provide written response that serves assessment for learning functions appears to be a powerful component of teacher practice that develops student writing. The paper suggests that considerable teacher pedagogical content knowledge is required to provide such feedback.
© 2010 Elsevier Ltd. All rights reserved.

Keywords: Formative writing assessment; Feedback; Written response; Writing achievement; Teacher knowledge

1. Introduction

Written response to writing, while considered a common form of writing instruction, has not been a central theoretical concern for research (Phelps, 2000). Although such feedback is intended to improve student learning in writing, ironically, most studies of written feedback have been conducted outside of a pedagogical context (Fife & O'Neill, 2001) or, indeed, of any

* Corresponding author. Tel.: +64 9 623 8899x88998.
  E-mail addresses: jm.parr@auckland.ac.nz (J.M. Parr), h.timperley@auckland.ac.nz (H.S. Timperley).
[1] Tel.: +64 9 623 8899x87401.

theoretical or communicative frame (Huot, 2002). Research has largely focussed on describing characteristics of the response and student interpretation of, and attitudes to, such response. With some exceptions, there is a lack of work that considers the interactive and contextual nature of response or work that considers response in relation to writing outcomes. This paper locates written response as an instructional act of feedback within the theory of formative assessment (often now known as assessment for learning). The study defines the dimensions of response likely to make it effective in relation to this concept of assessment and feedback for learning, and tests the idea that a teacher's ability to provide quality, written feedback for learning is an important component of teacher practice in writing by examining its relationship to student progress in writing.

1.1. Research on written response

Providing written feedback to writers is presented in the literature as a problematic practice. Comments on students' drafts are seen, in general, as not effective in improving writing (Hyland, 2000; Muncie, 2000). Generally, it seems that the nature of feedback influences impact. Certain sorts of feedback, like that focussing on personal qualities, can impede learning by shifting focus from instructional to social goals (Kluger & DeNisi, 1998), while outcomes-focussed feedback (e.g. a grade) seldom provides sufficient information to advance learning. The nature of feedback can also encourage surface versus deep learning. Feedback that focuses on the correctness of content in a domain generally contains insufficient information to affect the development of knowledge construction, whereas feedback directed at deeper learning may trigger forms of cognitive processing such as searching for relationships or developing knowledge to elaborate information (Balzer, Doherty, & O'Connor, 1989).

Studies of the nature of written response to writing have shown that college teachers' comments tend to focus on low level, technical concerns, rather than on meaning-making (e.g. Connors & Lunsford, 1993; Sommers, 1982). It has also been reported that school teachers similarly give excessive consideration to surface features, particularly with regard to revision (Hargreaves & McCallum, 1998). Teachers have been portrayed as unable to articulate deeper feature, rhetorical concerns (Schwartz, 1984). Although several of the studies documenting the nature of teachers' responses have been criticised as having methodological weaknesses (Ferris, 1997), findings regarding surface level feedback have been replicated more recently (Stern & Solomon, 2006).

However, this research has largely neglected the influence of context (Huot, 2002), treating the texts that teacher–responders create as if they stand alone, ignoring the perspective that the meaning of text will be constructed differently depending on the 'discourses' brought to bear on the text by the reader (Murphy, 2000). In particular, theoretical and cultural orientations affect interpretations (Ball, 1997); teachers also respond to extra-textual features within the context of the classroom (Wyatt-Smith, Castleton, Freebody, & Cooksey, 2003). Research adopting a more contextualised view (for example, by Weigle, Bildt, & Valsecchi, 2003) suggests that the characteristics of the writing task and of respondents, in particular the conventions and emphasis of a discourse community, play a part in influencing the criteria used to evaluate (and, presumably, to respond to) writing. Similarly, there is some evidence from research into the writing of primary students that it is amount and type of feedback that predicts the quality of final drafts (Matsumura, Patthey-Chavez, Valdes, & Garnier, 2002).

An additional consideration in evaluating the impact of feedback concerns the potential for action suggested by the feedback. This potential is compromised if students, as reported, have

difficulty in processing teacher written response (Zellermayer, 1989), or are confused, misinterpret feedback or cannot decipher comments (e.g. Nelson, 1995; Richardson, 2000; Sperling & Freedman, 1987). Further, the comments may not invite student response. Huot, for example, argues that response should be more transformative, the comments "open-ended, forcing students back into the text" (2002, p. 132). Feedback that is corrective rather than designed to foster development is unlikely to engage students. If students are not critically involved, the likelihood of their acting on feedback, or of the evaluative process becoming internalised and having effect beyond the current piece, is lessened (Muncie, 2000). Not only do students often find written feedback unhelpful regarding what to do but it is not seen as useful as a catalyst for discussion (Maclellan, 2001). Arguably, if messages are complex or difficult to decipher, students need opportunities to construct, through discussion, an understanding of them (Higgins, Hartley, & Skelton, 2001). The non-developmental and non-dialogic nature of much of the reported practice may contribute to student reaction to feedback and the reported lack of impact.

1.2. Locating written response within assessment for learning

Assessment for learning is a pedagogical context designed to promote learning and student engagement in their learning (Black & Wiliam, 2009). Feedback is a key element. Assessment for learning is designed to provide information about student performance that can be used to support learning and to modify teaching (Black & Wiliam, 1998; Shepard, 2005; Torrance & Pryor, 1998). Using assessment information in this way can improve the quality of teaching and learning outcomes (Black, Harrison, Lee, Marshall, & Wiliam, 2003; Black & Wiliam, 1998). While early formative assessment discourse focused on the role of teachers in gathering information and using it to inform their teaching, more recently there has been a re-conceptualisation. Formative assessment has been reframed as a social, collaborative activity, aligned more with learning (Black, McCormick, James, & Pedder, 2006; Gardner, 2006). The emphasis has shifted to the teacher and the students, working in partnership (Hawe, Dixon, & Watson, 2008) to enhance student learning.

Realising the benefits of assessment for learning requires that teachers help their writers to understand what the goals of learning are and provide opportunities for them to have feedback on progress towards such goals. The learners' understanding of the quality performance aimed for, what success in a task looks like, and what they might do to achieve it is directly related to the instruction and feedback received (Black & Wiliam, 1998). Effective feedback not only helps learners to evaluate where they are but provides them with an indication of where to proceed next and how best to accomplish this forward movement (Hattie & Timperley, 2007). Writers need response in the form of feedback not only for monitoring their progress and moving forward but also as a means of discovering their readers' needs (Zellermayer, 1989).

In this paper, written response is located within the pedagogical context of assessment for learning. For feedback to contribute to development in writing it needs to have particular characteristics. These characteristics include that the feedback given allows a writer to see where s/he is currently positioned relative to the quality performance desired; that the feedback indicate what the key features of a quality performance are, and what is needed to bridge the gap between current and desired performance. Feedback should also allow a writer to see how s/he might go about achieving the desired performance. We utilise these concepts regarding the nature of feedback that have been generated by the literature and theory related to assessment for learning (e.g. Ramaprasad, 1983; Sadler, 1989, 1998; Wiliam & Thompson, 2007) to define the characteristics of quality written response.

1.3. Feedback for learning and teacher knowledge

Implementing assessment for learning in classrooms is challenging (Hall & Burke, 2003; Torrance & Pryor, 2001). It requires a culture change in the classroom culture and expansive learning on the part of the teacher (Webb & Jones, 2009). Research literature suggests that the practices associated with assessment for learning, in general, make a far greater demand on a teacher's expertise than where judgements are simply about a student's ability in relation to others (Marshall & Drummond, 2006). Numerous researchers contend that knowledge of the subject, particularly from the point of view of teaching it to others (the pedagogical content knowledge discussed by Shulman, 1986, 1987), is important in order to respond appropriately to students while they are engaged in learning. In formative assessment terms, teachers need to notice a 'gap' (Sadler, 1989) in student performance between the current and the desired; they need to recognise contradictions in student learning (Ball & Bass, 2000; Cowie & Bell, 1999; Newton & Newton, 2001). They then need to work out how to move forward (Black & Wiliam, 1998) in order to effect improvement. Without a sound grasp of subject matter, from the point of view of teaching others, teachers are not able to formulate effective comments or questions. They are less able to anticipate where students may have conceptual difficulties or to develop support for learners to take the next learning steps, to scaffold that learning (Jones & Moreland, 2005; Shepard, 2005; Twiselton, 2000).

In writing, a problematic issue concerns the body of subject matter or content knowledge that a teacher needs to know, to transform in order to teach it. As Phelps and Schilling (2004) point out with respect to reading, it is not a discipline; there is no single group of scholars who identify what there is to be known about it. In terms of a disciplinary base, Snow, Griffin, & Burns (2008), writing about knowledge to support the teaching of reading, argue it is about language structure, its systems and sub-systems and how it is used but that teachers need usable knowledge, embedded in practice. Studies have shown that teachers tend to lack basic knowledge of how language works (Moats, 2000; Moats & Lyon, 1996; Wong-Fillimore & Snow, 2002).

In writing, like reading, there is no group of scholars who define what there is to be known about it. Indeed, the field is decidedly ecumenical in terms of epistemology. And, writing, as a construct, is complex. As a social and cultural act, it is problematic to specify what 'develops' or progresses and what it develops towards (Marshall, 2004), under what conditions. There are different Discourses (Gee, 1990) that operate when writing in different contexts and discipline areas. With respect to subject matter knowledge needed to teach writing, there is a dearth of research that suggests what such might be and how it relates to student learning outcomes. There is a body of work that relates professional learning opportunities with respect to writing to practice and, subsequently, to features of student writing (e.g. research on the National Writing Project, cited in Borko, 2004).

Clearly, teachers of writing need more knowledge than an average, competent adult writer. They need to know, at a conscious level, how texts work to achieve their communicative, rhetorical purposes, including knowledge of the features of text most commonly employed to support writing for a particular purpose. This involves a detailed knowledge of language and of text structures, which might be considered subject matter knowledge. But, to teach this to others also involves the ability to articulate and make accessible to developing writers what is implicit and often at a level below conscious thought. Teachers have to be able to unpack what writers are doing as they engage in the writing process, including the strategies more expert writers use in the complex activity of writing.

This explicit knowledge of language and how texts work and are formed in contexts, together with the knowledge of processes and strategies, needs to be married with knowledge of the developmental trajectory that may operate in learning to write and of the approaches, activities and resources most efficacious to employ with students. Most importantly, in the context of the present study, teachers need to utilise this package of knowledge and apply it in the light of their interpretation of evidence of writing achievement to provide feedback and support to move learning forward. Both lack of knowledge of the content of the subject, writing, and the difficulty in specifying a quality performance may partly explain the nature of unintended outcomes or lack of outcomes reported in the majority of the literature on teacher response to student writing.

1.4. Research aims

In this paper we have two, inter-related aims. First, we locate written response within a pedagogical, theoretical framework, that of assessment for learning. Within assessment for learning we define the hallmarks of quality feedback as those that support evaluating performance in terms of goals, ascertaining the nature of the gap between current and desired performance and identifying what is needed to move learning forward. Second, we present evidence to show that a teacher's ability to give this quality written feedback to writing that would support formative assessment practices is related to his/her students' progress over a school year in terms of writing achievement.

2. Method

2.1. Context

The present study was part of a larger research endeavour accompanying and informing a national professional development project in literacy (schools focussed on either reading or writing) for schools with students in years 1–8 of schooling. The professional development was school-based and job-embedded. An expert facilitator[2] worked with teachers and leaders in several schools over a two year period. Learning was needs based. Facilitators and their schools together gathered evidence at the beginning of the initial year about both student achievement and teacher knowledge and practice. Both the needs of students in terms of progressing their literacy learning and what teachers needed to know in order to address those needs were considered in concert in designing the professional learning. Evidence of teacher knowledge and practice was gathered through questionnaires and classroom observations.

In the professional development focusing on writing, facilitators emphasised building content knowledge of teachers about writing and how to teach it. This was achieved in a number of ways. Facilitators and literacy leaders supported teachers to interpret and apply the detailed criteria in the scoring rubrics contained in the diagnostic writing measure employed (Assessment Tools for Teaching and Learning: Writing (asTTle Writing), Ministry of Education, 2001, 2004). Utilising the criteria to score scripts required building teacher knowledge of language structure and features. Likewise, facilitators supported systematic moderation procedures for scoring of writing. The literature suggests that teachers are able to make reliable judgements of written work through

---

[2] Although the term literacy coach is common, we use the term facilitator as the role of these persons is to facilitate change at all levels of a school in terms of raising student achievement.

the process of construct referencing. Understanding of a construct is refined through experience and particularly through processes like moderation where collegial discussion is involved. The shared meaning that develops among those interpreting the evidence in the form of the writing becomes part of professional or guild knowledge (Sadler, 1987; Wiliam, 1998). Teachers were supported, often in professional learning circles, to interpret the information about their own students' writing from the diagnostic measure and to plan appropriate instructional moves to address the demonstrated needs (note that giving written feedback to writing was not a specific focus of the professional development but practices relevant to it may have arisen in the course of the professional learning circle discussions or classroom observations). Teachers were also observed in their classroom and were engaged in learning conversations (Robinson & Lai, 2006) to analyse their practice. As part of classroom observation, facilitators or school literacy leaders enquired of students their extent of understanding of learning goals; of criteria for successful performance and if they knew what they personally should be working on. The results of this inquiry became part of the evidence brought to the learning conversation and teachers were encouraged to utilise similar inquiry in their practice (Timperley & Parr, 2009).

2.2. Participants

Participants in the study were from six primary schools that focussed on writing. We sought and gained appropriate ethical approval for this study. As this was a school-wide professional development project, principals, senior management and classroom teachers participated. At Time 1 there were 59 educators who gave informed consent to participate and provided responses to the teacher knowledge scenario that contained a writing sample for feedback. The experience of these educators ranged from .2 to 36 years with a mean of 14.5 years. Data for those who completed the scenario at two points in time were available for 49 participants. The loss is attributable to the fact some participants left their school or were absent. However, the sample size for considering teacher feedback knowledge in relation to student progress is much smaller ($N = 30$) as some teachers in management positions had no teaching responsibilities and teachers of years 1–3 were not included as there is no suitable assessment tool for writing that has national normative data for these years. Student achievement data in writing for two points in time were obtained for a total of 375 students.

2.3. Procedure

Data from both teachers and students were collected at the beginning and then end of the school year in the first year of the project (in February and in November). The beginning of year data formed part of a needs analysis to determine the learning needs of students and of teachers in relation to that of the students. The teacher data relevant to this study included feedback responses to a writing sample presented within a scenario describing classroom teaching practice in writing. Responses were completed in a staff meeting.

2.4. Measures

There were two measures employed in this study, one to gauge the quality of teacher written response to writing and one to measure student achievement and progress in terms of that achievement.

2.5. The measure of teacher feedback quality

The quality of teacher feedback to writing was measured from their response to a student writing sample. The sample was presented, together with the hypothetical classroom writing lesson associated with the piece of writing. The lesson was described in a brief scenario (similar, parallel forms of this were used at the two points in time). We developed the scenarios in consultation with two literacy experts, both of whom were leading practitioners of in-service and/or pre-service teacher development in literacy, and trialed them on several practising teachers. Each scenario described the context (year/grade level of students etc), the teacher's aims for the lesson and then some significant aspects of teacher practice. Respondents had to consider specific aspects of the practice (like the lesson aim, a teacher action like modelling, or the nature of the activities developed). They were asked to rate them and give reasons for the rating. These aspects of practice provided some pedagogical context for the sample of draft writing from a student in the class that participants were responding to (Appendix 1 contains an example of a scenario and the associated writing sample. Note that the student sample was a genuine piece of writing but the lesson description was constructed).

In response to the student sample associated with the lesson, teachers were asked to imagine themselves as the classroom teacher and give feedback to this student to help improve the piece. First they were asked to write a brief general response to the student. This was to allow teachers to respond more personally or globally if they wished. They wrote responses like "*Wow Keri. You are so lucky to get to go to Sydney and Brisbane again. When I went 3 years ago I loved to go shopping and I loved the hot weather*" or "*Bet you had a great time visiting all these interesting people.*" Then they were asked for two key aspects of feedback they would give to the student that would be most helpful in improving his/her writing. They were provided with several lines to write in for each piece of feedback and it was these comments that were analysed.

2.6. The diagnostic assessment tool for student writing

Achievement data were obtained from a criterion referenced measure of writing (Assessment Tools for Teaching and Learning (asTTle), referenced to the New Zealand curriculum: Ministry of Education and University of Auckland, 2004) that has associated national normative data (Years 4–8 of schooling). A curriculum map outlines the higher-order dimensions of the curriculum area by which learning targets or outcomes or achievement objectives can be categorised. These dimensions represent major teaching and learning areas. For writing, there are six major communicative purposes for writing identified and over 60 different tasks encapsulating these communicative purposes. Seven dimensions of writing are included (four deep features – audience, structure, content and language resources, and three surface features – grammar, spelling and punctuation). For each of the six writing purposes, for each dimension and curriculum level of achievement, detailed criterion statements comprise the scoring rubric (as an illustration, Appendix 2 contains the scoring rubric for the deep features of the writing purpose "to recount").

The design of the writing assessment acknowledged that texts differ not only according to social purpose but also according to cultural practice and situational context and the interactions between them. The theoretical framing drew on the work of those who view writing broadly as serving social purposes (e.g. Chapman, 1999; Knapp & Watkins, 1994, 2005; Kress, 1993; Wray & Lewis, 1999). Writing was conceptualised as serving major functions or purposes (after Knapp & Watkins, 1994) that encapsulate what the text is doing. Knapp and Watkins (2005) call these a core set of generic processes. These processes are: to inform or entertain through narrating or

"storying" (imaginative narrative, personal interpretive/expressive); to inform or entertain through recount; to report and describe by classifying and organising information; to instruct or lay out a procedure; to argue or persuade, and to explain.

An analytic rubric for each of the major purposes that inform, or processes that form, texts provides descriptions of features and text structures (similar to what Knapp & Watkins, 2005, call the codes of writing) *commonly* associated with the generic social purpose. This originates from the ideas of Kress (1993), namely, that given the relative stability of social structures, forms of text produced (i.e. genres) in and by specific social institutions (like within schooling) will have some stability. The descriptions of features and text structures drew on functional linguistics, a tradition of linguistic research positioned somewhat differently in its theoretical view (e.g. Martin, Christie, & Rothery, 1987): functional linguists view meaning as determined by the language system and structures of texts. In describing linguistic features they tend to view text as product (e.g. Derewianka, 1990; Wing Jan, 1991). Text, as Knapp and Watkins (2005) note, can be seen, therefore, from two perspectives: "a thing in itself that can be recorded, analyzed and discussed, and also a process that is the outcome of a socially produced occasion" (p. 13).

In assessing writing, all dimensions of analysis of the text are seen as interdependent in terms of judging the effectiveness of the piece of writing but, for purposes of diagnostic assessment, the dimensions are considered and scored separately. Within each dimension, the criteria at each of the curriculum levels articulate a developmental progression. This allows tracking over time and helps teachers to work out what the next level of development would look like for any given writer. Teachers are able to see where each writer's strengths and weaknesses lie and to use the developmental characteristics provided by the progress indicators to plan teaching and learning activities and provide feedback and feed-forward to students. The asTTle tool is a diagnostic tool with associated normative data. It places the assessment in the hands of the teacher so the particular purpose selected to test student writing achievement depends on the teacher's classroom programme needs. The test software generates the appropriate prompt (from the bank of prompts which also have appropriate associated stimulus material) for the writing task, according to the purpose selected. Normally, a teacher or school would choose the time to test but for research purposes this was standardised.

3. Data analysis

3.1. Teacher feedback quality

The two specific feedback comments to the student writing sample were categorised and scored. We anticipated that a comment would likely be multi-dimensional. For example, while identifying a focus and explaining why it was problematic, it would also contain a suggestion as to a course of action.

There were five dimensions of comment scored. These derive from concepts in formative assessment, considered to contribute to a quality piece of feedback.

(1) The first concerned the evaluative (Tunstall & Gipps, 1996 call this descriptive) quality of the comment, namely, that it gave an indication of the extent to which the writer had met important learning objectives (in terms of the lesson aims or the writing purpose) or had achieved a quality performance. Such a comment should allow the writer to know how successful s/he was being in terms of certain understood aims or criteria.

(2) The second dimension on which the comment was coded was whether it was specifically linked to the writing function or the lesson aim as opposed to being a generic statement.
(3) Thirdly, the comment was coded for its accuracy in terms of the evidence in the writing sample, the features commonly associated with the purpose for writing, and the available details of the rhetorical situation.
(4) Fourth, the comment was considered in terms of whether there was an adequate indication to the writer of what action to take to improve the piece.
(5) Finally, the level of text to which the comment referred was coded using the seven dimensions in the asTTle scoring rubric (four deep features and three surface feature levels).

With respect to the first dimension, if the comment were explicitly evaluative and/or contained an indication of the extent to which aims for or requirements of a quality performance had been met or if it identified what the "gap" was (what was required to make progress to reach the desired performance), it scored two points. The rationale for this was that such feedback was more likely to provide opportunity for learning to occur. Examples of such comments were: "*You have set the scene well. You have included details of Who, What, Where, When, Why and How . . ..*" which tells the writer how well s/he has met the aim of an orientation. The following comment is an example of one that indicates clearly to the writer that criteria for a recount have not been met and the major reason why not. "*Keri have a look at my recount – see how it is talking about what has already happened. A recount is about something – an event or experience – that has happened. Yours is not all about something that has happened. Can you change yours so you are writing about what has happened instead of what you are going to do?*" Where the feedback comment contained only an implicit evaluation concerning the extent to which the learning aims had been met or was implicit regarding what a quality performance might contain, or where it was a non-specific personal or positive response, these scored only one point. This decision was made on the basis that there is no evidence that the latter, positive feedback, is related to improved performance (Hattie & Timperley, 2007) and an implicit or indirect comment has considerable potential for misunderstanding. Examples of positive feedback included more personal responses or non-specific positive responses of which an example was "*Sounds like you had a fantastic time! I like the way you used an interesting start.*" The latter comment is regarded as non-specific in that it is not clear what makes the start interesting or, indeed, whether the teacher's liking of the opening means it was an effective one! Implicit comments are where the message is not stated but is, rather, implied. For example, many comments contained the implication that something needed to be attended to and these were often framed in a passive voice, like the comment "*recounts are in the past tense.*" The implication is that the current piece is not written in this way. Sometimes what is needed to improve the piece is buried in an apparently personal desire or need. The comment "*I would like to know what you thought were the three (or two) best things that happened last time you went*" is presumably meant to suggest, in a round-about way, to the writer that a recount highlights a few salient things rather than being a "dawn to dusk" saga.

The writer scored a point for meeting each of the criteria for the three other dimensions, that is, for being specific to the writing function (e.g. to recount) or the lesson aim for the class as described in the scenario; for being accurate in terms of the function and aim, and whether the comment contained an adequate indication of particular action or actions that the writer should take. An example of part of a comment that specified actions that could be taken was, "*. . . I would suggest that you read this account out loud to a partner. This would help you to see where you need to put in punctuation to break it up so the reader can understand it better.*" Comments that were confusing in that they gave mixed messages or were partly inaccurate did not score. An

example of such a comment would be "*It sounds like planning your trip was such fun but we need more detail.*" While the plea for more detail is an implicit (and accurate) comment about a demonstrable need of the piece to provide amplification, describing the planning for a trip is not entirely consistent with features of a recount (in terms of the task). Finally, the level of text at which comments were directed was coded from deeper features of text (rhetorical/audience, structure, content and use of language resources) through to surface features (grammar, spelling and punctuation). A reference to deep features in the feedback scored one point. Each piece of feedback could, therefore, potentially score six points (total of 12 for the two feedback comments).

Reliability of coding of responses was obtained by a second rater coding all comments for a 20 percent sample of questionnaires. The percentage agreement between raters was then calculated (the number of agreements divided by the number of agreements + the number of disagreements × 100/1). An acceptable level was obtained, namely, 81 percent agreement.

3.2. Assessment of student writing

Writing samples from an asTTle writing task are scored using the assessment tool's rubric for the tested writing purpose (see Appendix 2). The marker determines which curriculum level description best fit each of the seven dimensions of writing, then decides whether on each dimension, the score should be augmented (Penny, Johnson, & Gordon, 2000) as being slightly below, or above the assigned level without being in the next level above or below. Each dimension of a script is given a score ranging from less than Level 2 Basic to greater than Level 4 Advanced. Within the software accompanying the test, the total score for each script is found by averaging the seven dimension scores. Curriculum-level scores are transformed into a numeric standardised score that represents performance against that of the norming population used by asTTle to calibrate the assessment tasks. Rasch modelling allows scores to be located on a common continuum (Hattie et al., 2003) with a mean of 500 and a standard deviation of 100. Scaled scores typically range from 150 to 1250. No score is below 100 and it is possible to score above 1250. As a common scale is used, gain can be calculated from total scores at Time 1 and Time 2 and viewed in relation to published normative expected gains for each year level.

With respect to reliability of scoring, the scoring of the normative sample for the assessment tool demonstrated that consistency and reliability of scoring writing was possible with light training. In that instance, there was more than 67% adjacent grade consensus, an average Cronbach alpha of greater than .80, and a dependability index (φ) greater than .80 for each purpose scored by one of seven separate panels (Brown, Glasswell, & Harland, 2004; Brown, Irving, & Sussex, 2004).

In this study, a member of the research team, highly experienced in scoring writing, scored student scripts. For five of the six schools participating this meant all students. The sixth school was larger and we scored all of the scripts at Time 1 and about half at Time 2 (hence there were only 196 students with researcher scored scripts on both occasions from that school). However, schools also scored their own scripts separately and engaged in moderating processes with their facilitator as part of their professional learning.

Gains in total writing score were used as the measure of progress as the assessment tool employs a common scale. For each individual teacher's class an average gain in total raw score was calculated. Then, individual teacher feedback scores, gained from their responses to the writing sample at Time 2, were considered in relation to the average student gain for their class.

Time 2 was selected as teachers would have had the opportunity to utilise the learning from the professional development with their writers.

4. Results

The major empirical question concerning the relationship between the ability of a teacher to provide quality assessment for learning feedback and the progress of their students in writing is considered first. The score each teacher obtained at the end of a year of professional learning, indicative of the quality of their feedback in terms of key components necessary to engaging in assessment for learning in writing, is viewed in relation to the progress in writing of students in their class. Then the details of the feedback to the writing sample are briefly examined for patterns of responding across time.

4.1. Teacher feedback quality in relation to student progress

For the purposes of this study, a measure of progress (gains in raw score) was used as this more accurately reflects the actions of the teacher than a straight achievement measure. Students made significant gains in writing. In the six research schools, the average effect size gain over two years calculated using mean score (Cohen's *d*) was 1.19 (for the project nationally it was 1.28) a large gain. The average expected effect size gain nationally, calculated from published cross-sectional data, is .2 per year or .4 over two years (calculated from published data in asTTle Version 4, Manual 1.1, chapter 3, 2005).

There was a significant relationship demonstrated between the quality score for teacher feedback at Time 2 and the average extent of gain in writing scores for students in their class ($r = .685$, $p < .01$). This is a large coefficient suggesting a relationship that has implications in practical terms. The more able a teacher was to provide the type of quality feedback to writing required in the context of assessment for learning, the greater the progress of his or her students in writing.

4.2. Feedback quality

The giving of written feedback to a student in an assessment for learning context involves knowing what a quality performance looks like (in the particular context); evaluating the extent to which the writing produced meets this; diagnosing the gap or problematic aspects of text given the quality performance or learning aimed for, and then articulating an appropriate solution for the student. The expectation was that feedback comments at Time 2 (end of one year on the project) would demonstrate the qualities to a greater extent.

The quality of feedback on writing was considered from a snapshot at each point in time that included all participants who responded. The basic unit of analysis was the comment and, potentially, each of the comments could contain more than one dimension of quality feedback. This analysis of comments does show some changes. For example, in terms of the evaluative quality dimension (that gave students an indication of where they were positioned or how well they had met learning goals or what was required to meet desired performance) there was a decline in the proportion of comments falling into what was the largest sub-category at the beginning, implicit responses, comments that contained a message that implied that the writing was lacking. The responses containing an implied message fell from about 73% at Time 1 to 50% at Time 2 (see Table 1). Correspondingly, comments containing an explicit evaluation and/or quality performance

Table 1
Percentage of comments: evaluative quality of teacher feedback at Time 1 and Time 2.

Evaluative nature	Time 1	Time 2
Implicit evaluative message	73	50
Explicit evaluation/indication of desired quality performance	22	47

criteria increased at Time 2. Those comments that contained an evaluative component that was explicit grew from 22% to 47%.

The percentage of feedback comments specifically linked to the writing purpose or lesson aim rather than being a generic comment also increased. At Time 1 around 48% of responses contained an identifiable link to purpose of lesson or of writing; this grew to 60% at Time 2. Accuracy levels remained similar as few comments were, at either time point, coded as inaccurate either in terms of the features of recount or of the writing sample or because the comment contained a mixed or confusing message. Similarly, comments containing an indication of action to take remained at a similar level. The nature of the actions suggested, like their relative power to affect writing quality, was not judged.

Most noticeably, at Time 2, teachers were giving feedback about the deeper features of text. The proportion of comments in the various categories that described levels of text changed and this is shown in Fig. 1. In particular, the percentage of comments directed to rhetorical (2) and structural (3) dimensions of writing approximately doubled (from 11% to 23% and 16% to 31%, respectively) from Time 1 to Time 2. The change in proportions was statistically significant, that is the movement in proportions from lower or surface levels of text to higher or deep features of text was significantly different from the movement from higher/deep to lower/surface levels ($p = .004$, McNemar test).
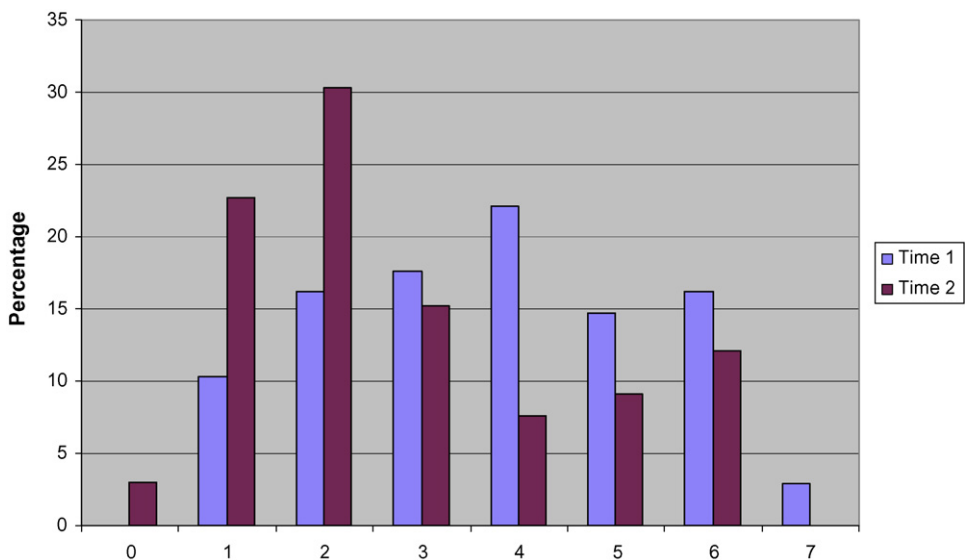

Fig. 1. Percentage of comments at different levels of text Time 1 and Time 2. *Note*: 1 = rhetorical, 2 = content/ideas 3 = structural, 4 = language resources 5 = syntax/grammar, 6 = mechanics 7 = other.

5. Discussion

The finding of a strong relationship in this study between teacher ability to give quality assessment for learning feedback and student progress suggests that this ability may be an indicator of teacher knowledge applied to practice. In the response to student writing produced as part of the lesson described in the scenario, teachers had, in providing feedback to the student, to retrieve knowledge from memory. This included content knowledge of how written language works to achieve communicative purposes, particularly in relation to recounting, but also their guild knowledge of what achievement in this particular instance (this writing purpose, this lesson and its aims, this year level of student) should look like. Then, this bundle of knowledge had to be applied to the student writing to diagnose the gap between produced and ideal; then further extended to work how best to respond to the student in ways that would allow him or her both to perceive the gap and also see a way to address it. This transforming of subject or content knowledge into the content of instruction, by adding the closely entwined pedagogical dimension is a challenging task for teachers. The notion that transforming is difficult seems to have support from work in mathematics. In a study looking at teachers' knowledge of fractions, the first part of the questions that Ward, Thomas, & Tagg, 2007 presented to teachers required them to use content knowledge to recognise whether a student answer was correct or not (the questions encapsulated common misconceptions), while the second part asked for what key understanding was needed to get the right answer or asked teachers to explain the reasoning behind the student's answer. Teachers were better on the content knowledge than on the latter, involving explaining the understanding or reasoning.

Writing, however, is far less straightforward than mathematics. The issue with providing assessment for learning feedback is related to working out how to move forward (Black & Wiliam, 1998) along a progression (Harlen, 1998) after noticing a 'gap' (Sadler, 1989). Whether this progression in writing is towards a defined goal or a broad horizon is a dilemma discussed by Marshall (2004). The goal model of progression is like a skills or knowledge model which assumes that what is necessary to be good at writing is "known, quantifiable and reducible to a systematic teaching programme" (Marshall, 2004, p. 102) while the horizon model suggests less specified outcomes and multiple pathways.

It seems that in order to understand progression in writing, teachers may rely only partly on a knowledge model which is then married with value judgements that are formed through the extensive process of making those judgements. Sadler (1989) wrote of the guild knowledge of teachers. The work of Wiliam (1998), building on that of Sadler, suggests that teachers are able to make reliable judgements of written work through the process of construct referencing. Understanding of a construct is refined through experience and processes like moderation where collegial discussion is involved. It is likely that the use of the diagnostic assessment tool rubrics served to help teachers understand more with respect to content knowledge about language as well as about expected progression (Parr, Glasswell, & Aikman, 2007). The often extensive process of moderation undertaken in these six schools is likely to have contributed to building shared meaning of quality performance in writing. Others have supported the idea that examining samples of student work collectively, against specific criteria, can foster teacher learning within a professional community (Little, Gearhart, Curry, & Kafka, 2003).

This study demonstrates the strong relationship between a particular instance of what could be viewed as requiring pedagogical content knowledge, in this case the ability to give quality feedback for learning, and student progress in writing. Others have positioned formative assessment and feedback within a model of self-regulated learning (e.g. Nicol & Macfarlane-Dick, 2006), postulating principles that support and develop self-regulation in students. Arguably, the

qualities of feedback provided within an assessment for learning framework, together with the its pedagogical intent provide both models for students and the type of information for them to engage in practices that support self-assessment and help them move towards self-regulation of their learning in writing.

Appendix 1. Writing scenario

The following is a brief description of a teacher's instruction while modelling writing for the whole class (a Year 4–5 class) where the intended task for the children was to write a recount of what they did on a recent trip or holiday. The purpose was to write to tell other people about what happened on the trip or holiday and the aim, according to the teacher, was to produce writing with appropriate structure that would be of interest to a reader (in this case, their peers).

Miss A modelled the writing process by composing aloud and on the whiteboard a description of a beach she had been to on holiday not long ago. She asked the children to picture a place they had been to recently that really stood out in their minds. Several children contributed by telling about their trips to fun parks and to beaches. The teacher then talked about and emphasised the need for specific vocabulary to convey a visual picture to the reader.

She got the children to draw a picture of the place they visited and label features to help them with vocabulary they might use. Then they swapped pictures with a partner and asked each other questions about the drawing. Then the children wrote their recount.

Rate the effectiveness of the following aspects of the lesson using the 1–6 scale. Give reasons for your rating.

highly ineffective	moderately ineffective	slightly ineffective	slightly effective	moderately effective	highly effective
1	2	3	4	5	6

The aim for the lesson (paragraph 1)   Rating:
Reasons for rating:

Introducing the writing task (paragraph 2)   Rating:
Reasons for rating:

Arranging the paired activities to support the writing (paragraph 3)   Rating:
Reasons for rating:

Here is a **draft** piece subsequently produced by one of the "target" Year 4 students from the class introduced above (the typing preserves the original spelling and punctuation but does not allow you to see the crossings out and insertions the student made).

When you read the piece, imagine you are Miss A. Write a brief response or comment to the student. Then note what you consider would be the two key aspects of feedback you would give to this student that will help to improve his/her writing.

A.1. Keri's writing

Hi I am at home planning my next trip to synedy and Brisbaned for Christmas and New years day I am going to stay in syned for a week and I am going to stay in Brisburnd for a week with

my mum's flat mate. When I went over to synedy I met rua his a dog of chris's. Chris is one of marys flat mate now last time when I went there I had to count his money. There is bazz she's another one of marys flat mate and nan she has 3 children one is around 14 years old the seoncod is 2 years old and the youngs child is nine mothes old. Then I went to Brisbured. When I got to synedy I am going to go to all this fantsey parks and wet and wild and I am going to stay in a hotal.

Overall response/ comment:

Feedback 1:

Feedback 2:

Appendix 2. Progress indicators for recount deep futures

Dimension	Level 2 Proficient	Level 3 Proficient	Level 4 Proficient	Level 5 Proficient	Level 6 Proficient
Audience Awareness and Purpose	● The writer recounts to tell of a past experience or happening.	● Language use and writing style appropriate to recounting a past happening.	● The text attempts to entertain/Inform.	● Writer explains events, situations, etc. as appropriate for the audience.	● Writer's awareness of audience is seen in ending of text.
	● Recognises that he/she is writing for an audience other than the self, but may be limited by assumption of shared knowledge.	● Audience has moat information needed to make sense of the experience recounted.	● Beginning of text attracts attention and provides adequate context for recount.	● May experiment with point of view or voice (e.g. use a "hook"/quote to engage reader's interest).	● Speaks directly/indirectly to audience making strong connections and drawing audience into recount.
		● Some inference may be needed.	● The text provides sufficient description of setting, situation, etc. for audience's understanding.	● Writer's developing awareness of audience is seen in reflective comments on the action.	Adjusts content details, and style to effectively bridge gap between writer and audience.
		● Evidence of attempts to capture the audience's interest.	● First person recount speaks to audience.		
Content/Ideas	● Writer recounts events.	● Writing includes evidence of foregrounding of significant content.	● Necessary elements of recount are present.	● Orientation is comprehensive, yet succinct.	● Writer's sense of looking back and analysing event is integrated fully through the recount.
	● Writing begins with an orientation (background information) using some of the elements of recount (when, where, who, what, and why).	● Evidence of attempts to add detail to, comment on, or evaluate selected points of interest.	● Focus on and development of some specific events and interest areas.	● Clear focus on and development of specific events and interest areas but not always edited for relevance.	● Recount is a enriched with interpretive comments, evaluation and observation, and sustained appropriate selection of details.
	● May be some evidence of selection of events for inclusion or of comment on events.	● There may be an attempt to conclude.	● Content selected for narrative rather than thematic impact.	● Thematic understanding/or interpretation of recount is attempted.	● Conclusion is linked thematically to recount content.
	● May include content not relevant.		● Simple appropriate conclusion.	● Some evidence of interpretive reflection on events in recount, possibly through sharing thoughts and feelings with the audience.	

References

Ball, A. (1997). Expanding the dialogue on culture as a critical component when assessing writing. *Assessing Writing*, *4*, 169–202.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In: J. Boaler (Ed.), *Multiple perspectives on teaching and learning* (pp. 83–104). Westport, CT: Ablex Publishing.

Balzer, W. K., Doherty, M. E., & O'Connor, R., Jr. (1989). Effects of cognitive feed-back on performance. *Psychological Bulletin*, *106* (3), 410–433.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham: Open University Press.

Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn: Assessment for learning: A theoretical enquiry. *Research Papers in Education*, *21*, 119–132.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, *5*, 7–73.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*, 5–31.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33* (8), 3–15.

Brown, G. T. L., Glasswell, K., & Harland, D. (2004). Accuracy in scoring of writing: Studies of reliability and validity using a New Zealand writing assessment system. *Assessing Writing*, *9*, 105–121.

Brown, G. T. L., Irving, S. E., & Sussex, K. (2004, August). *Scoring of a nationally representative sample of student writing at level 5–6 of the New Zealand English curriculum: Use and refinement of the asTTle progress indicators & tasks*. asTTle Tech. Rep. #48, University of Auckland/Ministry of Education.

Chapman, M. (1999). Situated, social, active: Rewriting genre in the elementary classroom. *Written Communication*, *16*, 469–490.

Connors, R. J., & Lunsford, A. A. (1993). Teachers' rhetorical comments on student papers. *College Composition and Communication*, *44*, 200–223.

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education*, *6*, 101–116.

Derewianka, B. (1990). *Exploring how texts work*. Sydney, Australia: Primary Teachers Association.

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly*, *31*, 315–339.

Fife, J. M., & O'Neill, P. (2001). Moving beyond the written comment: Narrowing the gap between response practice and research. *College Composition and Communication*, *53*, 300–321.

Gardner, J. (2006). Assessment and learning: An introduction. In: J. Gardner (Ed.), *Assessment and learning*. London: Sage Publications.

Gee, J. (1990). *An introduction to discourse analysis*. New York: Routledge.

Hall, K., & Burke, W. (2003). *Making formative assessment work: Effective practice in the primary classroom*. Buckingham: Open University Press.

Hargreaves, E., & McCallum, B. (1998). *Written feedback to children from teachers* ESRC Project 'Teaching, Assessment and Feedback Strategies'. Project Paper No. 7. London: University of London, Institute of Education.

Harlen, W. (1998). Classroom assessment: A dimension of purposes and procedures. Paper presented at the *Annual conference of the New Zealand association for research in education*.

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

Hattie, J. A. C., Brown, G. T. L., Keegan, P. J., Irving, S. E., MacKay, A. J., Pipi, A. H., et al. (2003, December). *Assessment tools for teaching and learning (asTTle) manual: Version 3 2004*. Wellington: Learning Media.

Hawe, E., Dixon, H., & Watson, E. (2008). Oral feedback in the context of written language. *Australian Journal of Language and Literacy*, *31*, 43–58.

Higgins, R., Hartley, P., & Skelton, A. (2001). Getting the message across: The problem of communicating assessment feedback. *Teaching in Higher Education*, *6*, 269–274.

Huot, B. (2002). *Re-articulating writing*. Logan, UT: Utah State University Press.

Hyland, F. (2000). ESL writers and feedback: Giving more autonomy to students. *Language Teaching Research*, *4*, 33–54.

Jones, A., & Moreland, J. (2005). The importance of pedagogical content knowledge in assessment for learning practices: A case-study of a whole-school approach. *Curriculum Journal*, *2* (16), 193–206.

Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current Directions in Psychological Science*, *7* (3), 67–72.

Knapp, P., & Watkins, M. (1994). *Context, text, grammar: Teaching the genres and grammar of school writing in infant and primary classrooms*. Sydney, Australia: Text Productions.

Knapp, P., & Watkins, M. (2005). *Genre, text, grammar: Technologies for teaching and assessing writing*. Sydney, Australia: University of NSW Press.

Kress, G. (1993). Genre as social process. In: B. Cope & M. Kalantzis (Eds.), *The powers of literacy: A genre approach to teaching writing* (pp. 22–37). Pittsburg, PA: University of Pittsburg Press.

Little, J., Gearhart, M., Curry, M., & Kafka, J. (2003, November). Looking at student work for teacher learning, teacher community, and school reform. *Phi Delta Kappan*, 185–192.

Maclellan, E. (2001). Assessment for learning: The different perceptions of tutors and students. *Assessment and Evaluation in Higher Education*, *26*, 307–318.

Marshall, B. (2004). Goals or horizons – the conundrum of progression in English: Or a possible way of understanding formative assessment in English. *The Curriculum Journal*, *15*, 101–113.

Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, *21*, 133–149.

Martin, J. R., Christie, F., & Rothery, J. (1987). Social processes in education. In: I. Reid (Ed.), *The place of genre in learning: Current debates*. Geelong, Australia: Deakin University Press.

Matsumura, L., Patthey-Chavez, G., Valdes, R., & Garnier, H. (2002). Teacher feedback, writing assignment quality, and third-grade students' revision in lower- and higher-achieving urban schools. *The Elementary School Journal*, *103*, 3–25.

Moats, L. (2000). *Speech to print: Language essentials for teachers*. Baltimore: Paul H. Brooks.

Moats, L., & Lyon, R. (1996). Wanted: Teachers with knowledge of language. *Topics of Language Disorders*, *16*, 73–86.

Muncie, J. (2000). Using written feedback in EFL composition classes. *ELT Journal*, *54* (1), 47–53.

Murphy, S. (2000). A sociocultural perspective on teacher response: Is there a student in the room? *Assessing Writing*, *7*, 79–90.

Nelson, J. (1995). Reading classrooms as text: Exploring student writers' interpretive practices. *College Composition and Communication*, *46*, 411–429.

Newton, D. P., & Newton, L. D. (2001). Subject content knowledge and teacher talk in the primary classroom. *European Journal of Teacher Education*, *24* (3), 369–379.

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

Parr, J. M., Glasswell, K., & Aikman, M. (2007). Supporting teacher learning and informed practice in writing through assessment tools for teaching and learning. *Asia-Pacific Journal of Teacher Education*, *35*, 69–87.

Penny, J., Johnson, R. L., & Gordon, B. (2000). The effect of rating augmentation on inter-rater reliability: An empirical study of a holistic rubric. *Assessing Writing*, *7* (2), 143–164.

Phelps, L. W. (2000). Cyrano's nose: Variations on the theme of response. *Assessing Writing*, *7*, 91–110.

Phelps, G., & Schilling, S. (2004). Developing measures of content knowledge for teaching reading. *Elementary School Journal*, *105*, 31–49.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, *28*, 4–13.

Richardson, S. (2000). Students' conditioned response to teachers' response: Portfolio proponents, take note!. *Assessing Writing*, *10*, 80–99.

Robinson, V. M. J., & Lai, M. (2006). *Practitioner research for education: A guide to improving classrooms and schools*. Thousand Oaks, CA: Corwin Press.

Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education*, *13*, 191–209.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.

Sadler, R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, *5*, 77–84.

Schwartz, M. (1984). Response to writing: A college-wide perspective. *College English*, *46*, 55–62.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*, 4–14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*, 1–22.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership.*, *63*, 66–70.

Snow, C., Griffin, P., & Burns, S. (Eds.). (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world*. Jossey-Bass: San Francisco.

Sommers, N. (1982). Responding to student writing. *College Composition and Communication*, *33*, 148–156.

Sperling, M., & Freedman, S. (1987). A good girl writes like a good girl: Written response and clues to the teaching/learning process. *Written Communication*, *4*, 343–369.

Stern, L. A., & Solomon, A. (2006). Effective faculty feedback: The road less travelled. *Assessing Writing*, *11*, 22–41.

Timperley, H. S., & Parr, J. M. (2009). What is this lesson about? Instructional processes and student understandings in the writing classroom. *The Curriculum Journal*, *20*, 43–60.

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Buckingham: Open University Press.

Torrance, H., & Pryor, J. (2001). Developing formative assessment: Using action research to explore and modify theory. *British Educational Research Journal*, *27*, 615–631.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal*, *22*, 389–404.

Twiselton, S. (2000). Seeing the wood for the trees: The National Literacy Strategy and Initial Teacher Education; pedagogical content knowledge and the structure of subjects. *Cambridge Journal of Education*, *30* (3), 391–403.

Ward, J., Thomas, G., & Tagg, A. (2007). What do teachers know about fractions? Paper presented at *National numeracy conference*.

Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice*, *16*, 165–184.

Weigle, S. C., Bildt, H., & Valsecchi, M. (2003). Effects of task and rater background on the evaluation of ESL student writing: A pilot study. *TESOL Quarterly*, *37*, 345–354.

Wiliam, D. (1998). The validity of teachers' assessments. Paper presented to the *22nd annual conference of the international group for the psychology of mathematics education*.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In: C. A. Dyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah, NJ: Erlbaum.

Wing Jan, L. (1991). *Write ways: Modelling writing forms*. Melbourne, Australia: Oxford University Press.

Wong-Fillimore, L., & Snow, K. (2002). What teachers need to know about language. In: C. T. Adger, C. E. Snow, & D. Christian (Eds.), *What teachers need to know about language* (pp. 7–54). McHenry, IL: Delta Systems Co.

Wray, D., & Lewis, M. (1999). *From learning to teaching: Towards a model of teaching literacy*. Unpublished Report, England (ERIC Document Service No. ED 432 777).

Wyatt-Smith, C., Castleton, G., Freebody, P., & Cooksey, R. (2003). Teachers' qualitative judgements: A matter of context and salience, Part 1. *Australian Journal of Language and Literacy*, *26*, 11–32.

Zellermayer, M. (1989). The study of teachers' written feedback to students' writing: Changes in theoretical considerations and the expansion of research contexts. *Instructional Science*, *18*, 145–165.

Judy Parr is an Associate Professor of Education at The University of Auckland. Her research interests concern teaching and learning, focusing on optimising literacy development. Her areas of specialisation encompass classroom practice, assessment and the nature of knowledge for teachers of writing, and she has published articles on these and related issues appearing in numerous international journals.

Helen Timperley is a Professor of Education at The University of Auckland. Her primary research interests focus on ways to promote professional learning of school leaders and teachers that positively impact on student outcomes. Her papers appear in international journals including *Journal of Curriculum Studies* and *Review of Research in Education*.