Educ Asse Eval Acc (2014) 26:107–133
DOI 10.1007/s11092-013-9187-5

Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses

Lois R. Harris · Gavin T. L. Brown · Jennifer A. Harnett

Received: 21 December 2012 / Accepted: 30 December 2013 /
Published online: 1 February 2014

Abstract While feedback is a key factor for improving student learning, little is known about how students understand and experience feedback within the classroom. This study analysed 193 New Zealand primary and secondary students' survey responses alongside drawings of their understandings and experiences of feedback to examine how they experience, understand, and respond to feedback. It found that despite New Zealand's strong commitment to student-centred Assessment for Learning practices, the majority of students still drew, selected, and endorsed teacher-led feedback practices, with pictures dominated by written comments or grades. However, they generally depicted and described this feedback as positive and constructive, suggesting that negative emotional responses to evaluative comments and grades may be lessened if students perceive such feedback will help them improve.

Keywords Feedback · Student perspectives · Classroom assessment · Formative assessment

L. R. Harris (✉)
School of Education, Central Queensland University, CQUniversity Australia Rockhampton Campus, Building 33, Bruce Highway, Rockhampton 4702, Australia
e-mail: l.harris@cqu.edu.au

G. T. L. Brown · J. A. Harnett
School of Learning, Development, and Professional Practice, Faculty of Education, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

G. T. L. Brown
e-mail: gt.brown@auckland.ac.nz

J. A. Harnett
e-mail: j.harnett@auckland.ac.nz

1 Introduction

Feedback is a key facet of classroom assessment because learning occurs through the use of formative feedback (Hattie 2009; Hattie and Timperley 2007; Tunstall and Gipps 1996a). This paper draws on the definition of feedback by Hattie and Timperley (2007, p. 81); that is, "information provided by an agent (e.g. teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding". This definition reflects key perspectives adopted by New Zealand primary and secondary schools as part of their adherence to the global Assessment for Learning (AfL) movement, which encourages educators to move the focus of assessment away from end-of-course (i.e. summative) testing or examinations to in-course (i.e. formative) improvement-oriented interactions between learners and instructors (Black and Wiliam 1998). While teachers may continue to be a primary source of feedback to students, this definition reflects the importance of peer- and self-assessment practices (e.g. Andrade 2010; Strijbos and Sluijsmans 2010); in these, students are seen as legitimate providers of feedback (e.g. Black and Wiliam 1998; Black et al. 2003).

While feedback is universally endorsed as beneficial to learning, what counts as 'good feedback' is contested, with contrasting perspectives about the optimal source, form, content, frequency, and timing of feedback (Shute 2008). Additionally, student viewpoints about their roles in feedback situations are not well understood. While it is well established that students react to, respond to, and potentially utilise feedback (e.g. Hyland 2003; Peterson and Irving 2008), the thinking that shapes how students understand feedback remains unclear. This paper aims to contribute to knowledge about how students understand and experience feedback within the classroom. It examines the research questions:

- What feedback practices do students report experiencing in New Zealand upper primary and lower secondary classrooms?
- What roles or purposes do students identify for these feedback practices?
- What emotions do students associate with these feedback practices?

2 Theorising feedback

There are numerous typologies and theorisations around what constitutes quality feedback (e.g. Butler and Winne 1995; Kluger and DeNisi 1996; Hattie and Timperley 2007; Nicol and MacFarlane-Dick 2006; Shute 2008; Tunstall and Gipps 1996a). Butler and Winne (1995) focused on the student's role within the feedback process. According to their model of self-regulation, the student drives the learning process and must start by interpreting the task's properties and by setting goals for themselves. As they work through the task, drawing on their own tactics and strategies, students must monitor their learning, utilising their own and external feedback about their progress. Here, feedback information from external sources (e.g. teachers, peers, and computerised feedback systems) combines with the learner's own perspectives to influence his or her knowledge and beliefs about the task, greater learning domain, learning processes and products, and his or her personal performance. They noted that

problems can arise when this external feedback is inconsistent (with itself or with the student's own feedback to him/herself), vague, or unrelated to the task. Additionally, this model also takes student emotions into account, noting the importance of motivation on a self-regulated learning process. The strength of this model is that it acknowledges the agency of students within feedback situations; for feedback to be useful, students must need it, receive it in a timely manner, and be able and willing to use it (Shute 2008).

While feedback is generally depicted as positive, several reviews have noted the neutral or negative effects it can have on learners and their motivation, once again foregrounding the importance of student agency in the feedback process. The meta-analysis of Kluger and DeNisi (1996) was centred around the premise that feedback interventions actually change the locus of a learner's attention relating to task learning, task motivation, and/or meta-task processes. They found that feedback interventions actually reduced performance in one third of the studies they examined. Three variables were found to reduce feedback's effects: praise, interventions that threatened self-esteem, and oral feedback. Concerns about the role of praise within feedback were also noted in the review by Hattie and Timperley (2007). Hattie and Timperley identified four types of feedback: task (i.e. whether work was correct or incorrect), process (i.e. comments about the processes or strategies underpinning the task), self-regulation (i.e. reminders to students about strategies they can use to improve their own work), and self (i.e. non-specific praise and comments about effort). While self-regulation feedback was considered to lead to greater student engagement, effort, and self-efficacy, making it the most powerful type (Butler and Winne 1995; Hattie and Timperley 2007), self feedback was problematic as praise may be used to mask low performance and seldom contains useful information about students' actual performance that could improve their learning (Shute 2008).

While teachers may often be the providers of feedback, which leads researchers to focus on their actions and the actual content of their comments, it is important not to lose sight that it is students and the uses they make of the feedback that will ultimately determine its effects on learning. Hence, it is important to gain a better understanding of student perspectives of the feedback they generate and receive.

3 Student perspectives on feedback

To date, most research on student understandings of feedback has been conducted in tertiary settings (e.g. Brown 2007; Carless 2006; Higgins et al. 2002; Nesbit and Burton 2006; Poulos and Mahony 2008; Weaver 2006), leading to a gap in the literature related to student perspectives within primary and secondary schooling, the focus of this paper.

Research suggests that even young students are able to identify and interpret feedback interactions. The study by Tunstall and Gipps (1996b) on UK Year 1 and 2 students found that almost all were able to describe descriptive and evaluative feedback they had received about their work, with the majority of this feedback being spoken comments from the teacher. Most students identified that standards were in place and articulated that they were asked to do work again when it did not reach these levels, showing the presence of formative feedback. While most students described positive feedback experiences, for the small minority who had received very negative

evaluations of work (e.g. teachers ripping up work), it was clear that this kind of negative feedback could damage student self-esteem. The students also indicated that negative reactions from peers were at times worse than negative teacher reactions.

Some research has examined what school students want from feedback and which processes they find especially helpful for improving their learning. Tjeerdsma (1997) found that the 6th grade students and their physical education teacher who participated in her study were relatively aligned, citing improving performance and reinforcing correct behaviour as the two most important functions of feedback. Some students also articulated that motivating and encouraging them was an important role of feedback, with most feeling that receiving feedback was a generally positive experience. The study by Lee (2007) with Hong Kong second-language learners found that the students described the error correction feedback received from their teachers as helpful, but admitted low recall of this feedback and its limited effects on their actual writing. They preferred receiving multiple modes of feedback (e.g. spoken, written, in-class discussions), believing these were preferable for retention, but described having few opportunities to revise and improve work after receiving feedback. They reported being sensitive to criticism, acknowledging that negative comments might decrease their motivation in the subject. While generally positive about opportunities for peer feedback, some had concerns about their ability to accurately correct others' work and described having few opportunities to use peer-assessment processes.

In the study by Pajares and Graham (1998) of US middle school teachers' and students' understandings of feedback on a poetry assignment, students overwhelmingly wanted honest and accurate responses to their work, saying they saw through teacher efforts to mask unsuccessful performance with praise. They knew teachers cared about them when they gave them honest and accurate feedback, not when they received empty praise and encouragement. While there is some data around what students report wanting from feedback, student emotional responses to feedback received have been largely ignored. It would be assumed that feedback would engender fewer negative emotional responses than those reported in the literature around assessment (e.g. Goetz et al. 2007; Harris et al. 2009; Wheelock et al. 2000b), however, as feedback does often contain evaluation, negative emotional response are possible.

Within New Zealand, the study by Peterson and Irving (2008) of Year 9 and 10 students found pupils reported that teachers usually generated positive and encouraging feedback, but the majority of students preferred feedback to be honest and constructive, noting feedback often told them what areas they needed to improve, but not explicitly how to make these changes. Students admitted to ignoring feedback, especially when it was given in conjunction with a grade, seeing the grade as clearer and more honest feedback than the corresponding narrative comments. However, within the focus groups, there were some minority opinions voiced, with some students preferring good grades and positive comments, regardless of the quality of the work. Although students reflected within the focus groups about what they wanted and needed out of feedback, there was little evidence they were using feedback data for goal setting or self-regulation. The self-report *Student Conceptions of Feedback* version III (SCoF-III; Irving and Peterson 2007) instrument, developed from these focus groups (Peterson and Irving 2008), captures student conceptions of feedback purposes. A New Zealand survey of 626 New Zealand Year 9 and 10 students using the SCoF-III found seven inter-correlated factors (i.e. *feedback from peers*, *feedback from parents*, *feedback*

motivates, *teacher feedback is trusted*, *feedback communicates standards, feedback is irrelevant*, and *feedback is enjoyable*), with good fit (χ^2=1,367.87, *df*=488, χ^2/df= 2.80, p=0.09; CFI=0.91, RMSEA=0.054, gamma hat=0.97, Irving et al. 2008). In their study, the SCoF instrument was used in tandem with a feedback practices checklist. Irving et al. (2008) found that, with cluster analysis and confirmatory factor analysis, responses aggregated into three types (i.e. Teacher Feedback, Symbolic Feedback, and Feedback from Others). Teacher feedback had a strong regression to mathematics achievement (β=0.40), while the other two types had small negative regressions to achievement (β=−0.13 and −0.20, respectively). While well-fitting, two items within Symbolic Feedback were around teacher suggestions and information for improvement, which seem logically more consistent with Teacher Feedback.

While most studies have focused more explicitly on student reactions to and perceptions of teacher feedback, some have also examined their understandings of the student-led feedback situations (e.g. self-assessment, peer-assessment) promoted within an Assessment for Learning framework. Research suggests that the validity and effectiveness of feedback from students is dependent on inter-personal relationships and psychological issues related to self-disclosure and trust (Black et al. 2003; Harris and Brown 2013; Cowie 2009; Peterson and Irving 2008; Topping 2013; van Gennip et al. 2010) as students take up the complex role of assessor (Topping 2010). Students may be more likely to perceive peer feedback as effective if they are able to question and interact with peers when receiving the feedback (Black et al. 2003) or provide justifications for their evaluations (Gielen et al. 2010). Studies have indicated that some students and teachers question the validity and reliability of the feedback received through these practices (e.g. Harris et al. 2009; Harris and Brown 2013; Peterson and Irving 2008), as especially among younger students, student feedback cannot be expected to have the accuracy of expert feedback (Gielen et al. 2010). van Gennip et al. (2010) found that while students' trust in feedback from peers grew after partaking in peer-assessment, it was much more difficult to get students to gain confidence in their own skills as feedback providers, despite feeling psychologically safer than when receiving teacher feedback.

The limited research on student perspectives of feedback within primary and secondary school settings suggests that while most students want honest, direct, and constructive feedback from teachers, individuals have differing perspectives, with some seeking positive reinforcement from feedback. Students seem to like receiving multiple forms of feedback and may prefer teacher feedback to that generated by themselves or peers due to concerns about accuracy. However, not enough research has been conducted to paint a consistent and compelling picture of student perspectives of feedback. There has been insufficient examination of student understanding of the feedback they experience and the emotions associated with it.

4 New Zealand context

This study explored the conceptions of feedback of a large sample of New Zealand students. New Zealand has adopted an Assessment for Learning framework as its national strategy and actively promotes formative uses of assessment through professional development and assessment resources with "a deliberate focus on the use of professional teacher

judgment underpinned by Assessment for Learning principles rather than a narrow testing regime" (New Zealand Ministry of Education 2010, p. 5). The national assessment policy prior to Year 11 (students nominally 15 years old) emphasises voluntary, school-based assessment for the purposes of improving instruction and raising achievement relative to the learning outcomes and objectives specified in the national curriculum (Crooks 2010). The New Zealand curriculum is child centred, non-prescriptive, holistic, and integrated (Fraser 2001), with learning outcomes and objectives specified across multiple levels. While teachers provide formal feedback to students and members of the school community through written reports, they are also expected to give formative feedback to students in the interactive process of teaching and learning. This study was carried out before the introduction in 2010 of National Standards (mandating the reporting of student achievement against benchmark standards to both parents and the Ministry of Education) and so it is expected that within this study, students' experiences will be driven largely by formative feedback practices.

However, New Zealand's formative approach to assessment is not anti-testing. There has been considerable investment into the creation and implementation of a toolbox of standardised testing resources that schools use as part of their diagnostic analysis of student learning needs. The toolbox includes an electronic formative testing system, Assessment Tools for Teaching and Learning (asTTle), that allows teachers to create on-demand, norm referenced, standardised tests, which match the particular curriculum objectives they are teaching and provide diagnostic analyses of performance (Hattie and Brown 2008). The Ministry-funded toolbox also includes standardised tests [e.g. Progressive Achievement Tests (PAT), Supplementary Tests of Achievement in Reading, Observation Survey of Early Literacy Achievement] related to multiple curricular areas (Brown et al. 2008). Additionally, New Zealand teachers frequently administer tests, quizzes, and performance assessments they have created to monitor learning progressions against achievement objectives (Croft et al. 2000). Nonetheless, in accordance with a strong child-centred pedagogy and formative assessment policy, New Zealand teachers also make extensive use of in-the-moment interactions with children to guide instruction and reporting. Hence, it would be expected that students would receive feedback based on a diverse range of assessment techniques, making the New Zealand context an ideal environment to discover whether student conceptualisations of feedback are aligned with the formative ideals espoused in policy and practice.

Within New Zealand, both primary and secondary school teachers generally subscribe to the conception that assessment is for improved teaching and learning (Brown 2011) and strongly endorse the notion that feedback should improve learning rather than enhance students well-being (Brown et al. 2012). There is evidence that this feedback is usually focused on the task and processes underpinning it, not general praise and comments about behaviour and effort (Harris et al. 2013). However, there are distinctions between primary and secondary schooling which may affect how feedback is practiced.

First is the organisation of teaching and learning. Within primary (Years 1–6) and intermediate schools (Years 7 and 8), students usually remain with the same generalist teacher and peers for their school day. In secondary schools (Years 9–13), teachers are subject specialists and students normally have different teachers for each subject. Secondary students are also unlikely to share all classes with the same peers due to

timetabling, especially after Year 10. Hence, primary teachers work intensively and flexibly with one class of students across subject areas, while secondary teachers work with a much higher number of students. It may be that these structural differences encourage teachers to provide feedback in slightly different ways.

Second, while assessment in New Zealand is generally low-stakes, from Year 11, students work towards the National Certificate of Educational Attainment (NCEA), a high-stakes national certification system where students earn credits by demonstrating proficiency on curriculum standards through moderated internal and external assessments. Thus, while the NCEA system is high-stakes for individuals, since it determines educational and employment opportunities, it is also indirectly high-stakes for schools in that NCEA results are published by school name, impacting on how schools are perceived. Nonetheless, the NCEA system is not intended as a school evaluation system as is often seen in national testing systems. Secondary school assessment is often focused on preparing for or implementing NCEA (Crooks 2010), with considerable washback from Year 11 onto Years 9 and 10 practices (Mizutani 2009). The NCEA system has clear standards of performance related to learning objectives, criteria, and targets, with exemplars of performance made available to students and teachers. As the system is not dependent on the rank ordering of students, there is a strong element of formative, capacity, and competence building feedback, suggesting primary and secondary students may have similar views, despite the potential influence of high-stakes testing on feedback practices as students progress through secondary school.

This study seeks to add to the research on student perspectives of feedback through the combined use of a self-report inventory (SCoF-III), feedback practice checklist, and a relatively new technique in which students draw a picture of feedback as they experience it. These techniques aimed to gather data not only about the kinds of feedback students report actually receiving, but also about their emotional and learning responses to this feedback within a policy context where Assessment for Learning is prioritised.

5 Method

5.1 Participants

This study recruited 193 (female=110, male=81, not given=2) students from 13 teachers' classes in 11 Auckland region schools who were already participated in a multi-year investigation into teacher assessment practices. Of these students, 88 were in lower secondary school (i.e. Years 9–10), with the remaining 105 in upper primary or intermediate (i.e. Years 5–8) classrooms. In compliance with University of Auckland Human Participants Ethics Committee regulations (ref: 2008/521), all participating teachers, students, and their parents gave written consent after reading a supplied Participant Information Sheet. Students' ages ranged from 9 to 15 (M = 12.07, SD = 1.80). Student self-reported ethnicities were predominantly New Zealand European (n= 100, 52 %), followed by Asian (n=39), Pacific Islander (n=21), and indigenous Māori (n=6). The balance (n=27) were not given. The proportion for the New Zealand European and Māori were both less than the expected proportions in the New Zealand population. Thus, the sample, while reasonably large, is not fully

representative of the ethnic makeup of the nation and consequently analyses according to student ethnicity have been avoided.

5.2 Instruments

All data were obtained from an individually completed self-report survey instrument consisting of three sections.

Section 1—Feedback practices. Students were first asked to indicate from a list of 15 different practices those they had experienced this year. This technique is an adaptation of the previous SCoF-III study with secondary students (Irving et al. 2008), which found that, with cluster analysis and confirmatory factor analysis, responses aggregated into three types (i.e. Teacher Feedback, Symbolic Feedback, and Feedback from Others).

Section 2—Student conceptions of feedback. The previously developed, fixed-response, SCoF-III questionnaire about the purposes of feedback was given to the students as the second part of the survey. While this instrument was designed for students in lower secondary school, its Flesch-Kincaid reading grade level (as calculated by MS Word) was assessed as 4.7, making it suitable for children in Years 5 and above. Additionally, to make sure students were not having problems understanding the prompts, members of research team assisted in its administration in two classrooms where it was hypothesised that students may experience difficulties (i.e. a low SES Year 5 class and a Year 9 class with a high number of special needs students); in both instances, students were able to independently complete the instrument after receiving the standard instructions. The response system for the 42 items was a six-point, positively-packed agreement scale (Lam and Klockars 1982) with two negative options (*strongly disagree* and *mostly disagree*) and four positive options (*slightly agree*, *moderately agree*, *mostly agree*, and *strongly agree*). This scale was used because it was believed that most students would agree to some degree with all of the purposes of feedback (Brown 2004).

Section 3—Draw a picture. After completing the two fixed-format questionnaire items, students were asked to create a free-response drawing about their experiences of feedback. Free-response drawings can be an effective way of understanding the complexity and subtlety of student thinking and emotions about phenomena (Clarebout et al. 2007; Diem-Wille 2001). Student drawings have already proven useful in examining the complexity of student thinking and emotions about assessment (Carless and Lam 2012; Haney et al. 2004; Harris et al. 2009; McKillop 2006; Picker and Berry 2000; Wheelock et al. 2000a, b). This instrument was added to the inventory to provide possible triangulation of survey results, establishing whether students were drawing similar practices to those that they were selecting in the checklist and evaluating in the questionnaire. Additionally, the draw-a-picture instrument allowed further dimensions of the student feedback experience to be examined that are not explored thoroughly in the survey items: students' emotional response and potential responses to feedback. It was expected based on previous studies (e.g. Harris et al. 2009) that student pictures would contain imagery representing emotions (e.g. positive/negative symbols, facial expressions) which could shed light on their feelings about feedback. The

drawings also provided students with the opportunity to show their role in the feedback situation and what they might potentially do as a result of receiving feedback.

With the instrument, students were randomly assigned one of four prompts, each of which was worded slightly differently to test student sensitivity to the prompt's wording (i.e. (1) Draw a picture of feedback. This picture can be about what you think it is and how it makes you feel. If you are not sure what feedback is, look at other pages to get some ideas. (2) Draw a picture of how you get feedback about your learning to help you know how you are doing and what you need to work on next. (3) Draw a picture of how you find from your teacher, your peers, or yourself whether you are learning and how you figure out what you need to work on next. (4) Draw a picture showing how you figure out what you are doing well and what you need to work on next with your learning.). While these prompts were not pilot tested, they were worded in a similar way to the successful prompt from a previous study about student conceptions of assessment (Harris et al. 2009). Chi-square tests for differences in frequency of categories according to prompts indicated there were no statistically significant differences; hence, this paper reports aggregated results.

5.3 Procedures

While two classes had the research team administer the questionnaire with support from their normal classroom teacher to confirm the readability of the instruments, the balance were overseen by the students' normal classroom teacher during a standard lesson and returned using postage-paid envelopes. While the instruments contained explicit written instructions that students could read independently, teachers were encouraged to go through these instructions orally with their students and make sure any questions were answered prior to commencement of the instrument. Only one student out of 193 did not complete the draw-a-picture instrument and 10 students did not complete Section 1 Feedback Practices list. Nearly three quarters of students answered all 42 items of the SCoF-III and expectation maximisation procedures were used to impute missing responses (maximum missing per person was three responses for just four students and maximum missing per variable was 19 cases for one variable). Hence, the grand majority of students were able to follow these procedures successfully.

5.4 Analysis

Fixed-format questionnaire items Confirmatory factor analysis was used to test the pre-existing 33 item, seven-factor solution for the SCoF-III. Because the fit did not meet conventional standards, exploratory factor analysis, following procedures outlined in Courtney (2013) was used to find the most likely factor structure. In accordance with exploratory factor conventions, items were accepted as belonging to their intended factor only when their loading was >0.30 and if cross-loadings, as indicated by modification indices, were low (Bandalos and Finney 2010). Subsequently, confirmatory factor analysis was used to determine the fit of exploratory result. Likewise, the Feedback Practices instrument did not fit well with the original model; exploratory

factor analysis was used to identify the dimensionality and confirmatory factor analysis used to determine the quality of fit.

Maximum likelihood confirmatory factor analysis of the variance-covariance Pearson correlation matrices, using AMOS software (IBM 2011), was used to test the confirmatory factor models. In line with suggested practice (Cheung and Rensvold 2002; Fan and Sivo 2007; Marsh et al. 2004), models with statistically non-significant χ^2 per *df*, GAMMA hat of >0.90, root mean square errors of approximation (RMSEA) of <0.08, and standardised root mean residuals (SRMR) close to 0.06 were considered sufficiently close to the data to not be rejected. Correlations between the Conceptions of Feedback and Feedback Practices factors and the drawing content themes were also found to explore relationships across methods.

5.5 Drawings

Content analysis (Bell 2001, p. 13) was used to analyse the drawn pictures of feedback and is "an empirical (observational) and objective procedure for quantifying recorded 'audio-visual' (including verbal) representation using reliable, explicitly defined categories". The visual elements (e.g. symbols and images) of the drawings themselves can be analysed for meaning; for example, McKillop (2006) inferred that students had a negative attitude towards assessment by referring to painful, oppressive imagery in drawings (e.g. a student crying with pain, a darkened room with a chair in the corner, a student cringing in the corner and looking very small, a student with a gun in his mouth).

To classify the drawings, initially, the first author and third authors independently viewed each picture as a whole, iteratively identifying major categories of imagery. Both authors then worked together to synthesise and group these codings, identifying four major categories (i.e. form of feedback, content of feedback, emotional impact of feedback, and student's response to feedback) with subcategories beneath each. The first and third authors independently coded all images based on mutually agreed upon subcategory definitions, using the student captions to better understand the intent of the drawings. When a picture did not contain imagery representing a particular subcategory, it was scored as a zero, indicating missing data, rather than suggesting the student had purposely excluded the construct. Inter-rater agreement was calculated using the kappa coefficient and Pearson's correlation (*r*). Values of kappa greater than 0.80 indicated that agreement between raters was independent of chance processes (Cohen 1960; see Table 1).

Table 1 Inter-rater reliability statistics for classifying drawings by four major categories

	Agreement statistic		
Category	Consensus (%)	Pearson (*r*)	Kappa (κ)
Form of feedback	97	0.89	0.83
Content of feedback	94	0.82	0.85
Emotional impact	96	0.86	0.87
Student response to feedback	96	0.86	0.88

As sample size was small for some subcategories, in order to better test for the statistical significance of the findings, some were combined to form conceptually meaningful aggregated variables (see Table 2). In this analysis, form of feedback was examined through the subcategories Teacher Feedback, Student-Led Feedback, Spoken Feedback, and Written Feedback. The content of feedback was examined through the new subcategories Task Feedback and Self Feedback, while Emotional Impact was simplified into Positive Emotions and Negative Emotions. When examining student responses to feedback, Student Accepts Feedback and Student Rejects Feedback were retained from the original categories, while Acts on Feedback and Shares Feedback were created by combining existing subcategories. Once the new subcategories were formed, χ^2 tests were run to examine whether there were statistically significant distributions in these categories according to school sector, with small cell sizes precluding analyses by teacher or ethnicity. Cramer's V (V) tests were used to measure the degree of association between two nominal or categorical variables, while adjusting for the number of cells. Values below 0.40 are considered weak to moderate, with values above 0.60 being strong (Rea and Parker 1992).

6 Results

6.1 Conceptions of assessment

Exploratory factor analysis suggested three inter-correlated factors which after trimming poor fitting items was found to have acceptable fit (χ^2=617.96, *df*=347, χ^2/df=1.78, p=0.18; CFI=0.87; gamma hat=0.91; RMSEA=0.064, 90 % CI=0.056–0.072, SRMR=0.076). The three factors were Comments for Improvement (13 items), Interpersonal Feedback (7 items), and Negative Feedback (8 items) (Appendix 1). Four items about trusting the teacher's feedback had negative loadings on the Negative Factor; hence, these items were reverse scored before mean scores were created (Table 3). Very strong endorsement was given to Comments for Improvement, moderate positive endorsement was provided to Interpersonal Feedback, while weak disagreement with Negative Feedback was seen. Interestingly, no items related to self-assessment loaded onto the Interpersonal Factor; however, this was likely because the wording of these three items positioned self-evaluation as superior to alternative sources of feedback (e.g. Comments about my work are not necessary because I already know how well I am doing), rather than simply acknowledging utility or value in this practice. Improvement was moderately correlated with Interpersonal Feedback (Table 3), indicating that adult sources of improvement comments were seen as relatively independent of student interpersonal sources. The Negative Feedback factor was strongly and inversely correlated to Improvement, with no statistically significant correlation to Interpersonal Feedback, suggesting interpersonal sources of feedback are randomly associated with negative aspects of feedback.

Feedback Practices An inter-correlated three factor solution with good fit was found (n=182; χ^2=133.38, *df*=74, χ^2/df=1.80, p=0.18; CFI=0.94; gamma hat=0.96; RMSEA=0.067, 90 % CI=0.048–0.085, SRMR=0.069). The feedback practices (Appendix 2) were: Teacher Evaluation (four items), Teacher Help (six items), and

Table 2 Feedback category characteristics

Categories	Characteristics
Form of feedback	
Teacher feedback	Written feedback from teacher
	Spoken or nonverbal feedback from teacher
Student-led feedback	Spoken or non-verbal from self
	Spoken or non-verbal feedback from peer
	Written feedback from peer
	Written feedback from self
Spoken feedback	Spoken or nonverbal feedback from teacher
	Spoken or nonverbal feedback from self
	Spoken or nonverbal feedback from peer
	Feedback from parents
Written feedback	Written feedback from teacher
	Written feedback from peer
	Written feedback from self
	Written feedback source ambiguous
	Grades, results, scores, outcomes, reports
	Ticks and crosses
	Smiley face, stickers, stamps, rewards, certificates
Content of feedback	
Task feedback	Describing or comparing performance
	Suggestions for improvement, feed-forward
	Surface learning features (spelling, grammar, basic facts, recall)
	Presentation (neatness, layout, speed)
	Deep learning (understanding, explaining, extending, clarifying)
Self feedback	Praise and encouragement
	Effort and identifying improvement
	Behaviour
Emotional impact	
Positive	Smiling student
	Smiling teacher
	Positive symbols
Negative	Sad, angry or upset student
	Sad, angry or upset teacher
	Negative symbols
Student response to feedback	
Accepts	Student accepts feedback
Rejects	Student rejects feedback
Shares	Student shares feedback results with peers;
	Student shares feedback results with parents
Acts on	Student action or intention to act on feedback from others;
	Student reflects, compares results, and/or proposes own action;

Table 2 (continued)

Categories	Characteristics
	Student questions teacher about feedback or requests clarification

Interpersonal (four items). Students selected, on average, about two thirds of the items within the Teacher Help and Evaluation Practices; in contrast, they only selected just under half the interpersonal practices (Table 4). The Teacher Help factor was strongly inter-correlated with both other factors, while Interpersonal and Teacher Evaluation were moderately inter-correlated, providing some divergent validation evidence.

The inter-correlations of the three conceptions of feedback to the three feedback practices also provided some validation evidence (Table 5). Divergence was seen in the weak inverse correlation between Teacher Evaluation practices and Comments for Improvement, while convergence was seen in the two Interpersonal factors. The remaining inter-correlations were statistically not significant.

6.2 Drawings of Feedback

The images drawn by students were classified according to four major categories. These were Form of Feedback, Content of Feedback, Emotional Impact, and Student Response to Feedback. Each category had multiple subcategories and agreement indices were based on whether each drawing was assigned to the same subcategory within each major category.

The form of feedback category was dominated by written feedback, grades of various sorts, and assorted feedback from the teacher (see Fig. 1 and Table 6). While no one form was identified by the majority of students, only 10 students showed no form of feedback and the three most frequent were present in at least a third of drawings. Overall, teacher sources were explicitly present in 140 (73 %) individual drawings; however, many of the other representations of feedback (e.g. grades, ticks, and crosses) were also likely of teacher origin.

In terms of content, feedback was portrayed overwhelmingly as descriptive information that compared performance to standards, criteria, or expectations (76 %; see Fig. 2 and Table 7). The next most frequent categories were suggestions for improvement and praise. Feedback that focused on surface characteristics of work or presentation

Table 3 Student conceptions of feedback descriptive statistics

			Factor inter-correlations		
Factor	*M*	*SD*	I	II	III
I. Comments for improvement	4.97	0.77	–		
II. Interpersonal feedback	4.03	0.95	0.35	–	
III. Negative feedback	1.92	0.78	−0.81	-0.12^{ns}	–

ns not statistically significant

Table 4 Feedback practices descriptive statistics

			Factor inter-correlations		
Factor	*M*	SD	I	II	III
I. Teacher help	0.68	0.33	–		
II. Teacher evaluation	0.65	0.38	0.73	–	
III. Interpersonal	0.45	0.38	0.73	0.52	–

accounted for just 11 (6 %) drawings. Given the relatively low focus on praise and effort (37 and 29 %), it can be concluded that the students perceived feedback predominantly in terms of task description and growth-oriented steps for improvement.

In terms of the Emotional Impact category, the drawings were dominated by positive emotions (see Fig. 3 and Table 8). Students smiled in 44 % of drawings, teachers smiled in 26 %, and positive symbols were seen in 9 % of drawings. In contrast, negative emotions were seen from students in 7 % of drawings and from teachers in just 3 % of drawings. This stands in stark contrast to the much darker pictures of assessment drawn by New Zealand students in an earlier study (Harris et al. 2009; Harnett et al. 2012), where negative depictions outnumbered positive ones.

When examining how students drew the actions they took in response to feedback, there was a high proportion of students (81, 42 %) who did not include any element of responding (Table 9). However, nearly half drew the student accepting the feedback (49 %) and nearly a quarter (23 %) showed the student taking or intending to take an action based on the feedback (see Fig. 4). Only two students drew pictures showing rejection of feedback, though 13 showed the student feeling negative as a response. Only 17 portrayed sharing the feedback with a peer (10) or an adult (7). Together, these suggested that feedback is still a very personal, albeit positive, construct that is understood to be about improved learning.

6.3 Differences by school sector

While sample size precluded analyses by individual teacher and ethnicity, there were a number of statistically significant differences by school sector, identified by examining the mean scores for the feedback practices and conceptions, alongside the frequencies of student inclusion of particular categories of imagery within their drawings (see

Table 5 Inter-correlations conceptions of feedback and feedback practices factors

	Feedback practices factors		
Conceptions of feedback factors	Teacher evaluation	Teacher help	Interpersonal
Comments for improvement	−0.18*	−0.08	−0.05
Interpersonal feedback	−0.10	0.09	0.28**
Negative feedback	0.14	0.04	0.05

N=182

*$p<0.05$; **$p<0.01$

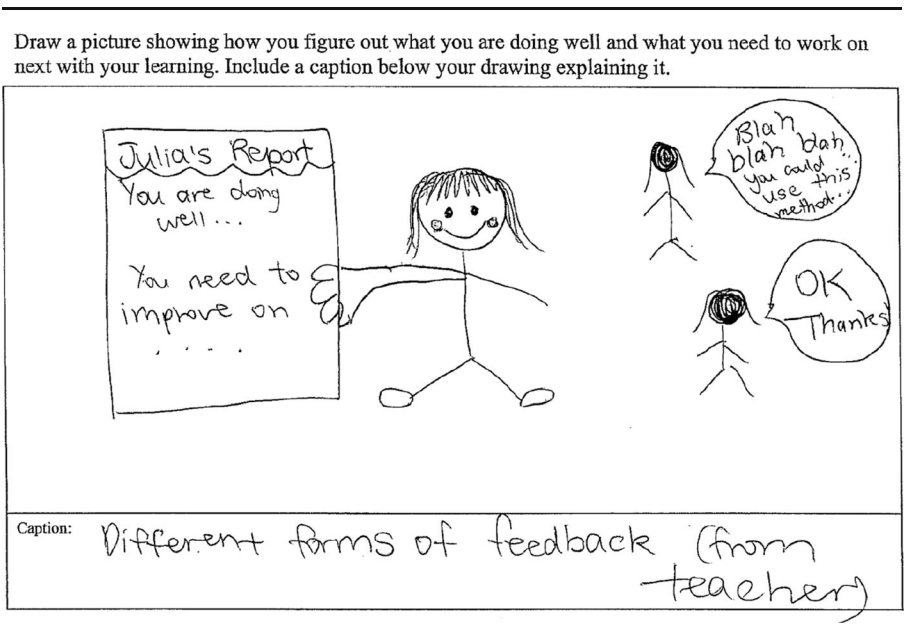

Fig. 1 Student drawing of teacher feedback

Table 10). Primary students were more likely to represent feedback as student-led, spoken, task-oriented, and self-oriented than secondary students. Thirty-five primary and intermediate students did not have explicit teacher sources, while only 18 secondary students did not include feedback explicitly said to be from the teacher, suggesting that primary and intermediate students were likely receiving feedback from more diverse sources than secondary students. The primary students endorsed the Comments to Improve and Interpersonal conceptions of feedback more than secondary students, with large and moderate effect sizes respectively. In contrast, secondary

Table 6 Frequencies within the form of feedback category

Form of feedback	*N*	%
Written feedback from teacher	91	47
Grades, results, scores, outcomes, reports	85	44
Spoken or nonverbal feedback from teacher	67	35
Ticks and crosses	37	19
Spoken or non-verbal feedback from self	29	15
Spoken or non-verbal feedback from peer	19	10
Smiley face, stickers, stamps, rewards, certificates	16	8
No form	12	5
Spoken feedback (source ambiguous)	10	6
Written feedback source ambiguous	5	3
Feedback from parents	3	2
Written feedback from peer	2	1
Written feedback from self	1	1

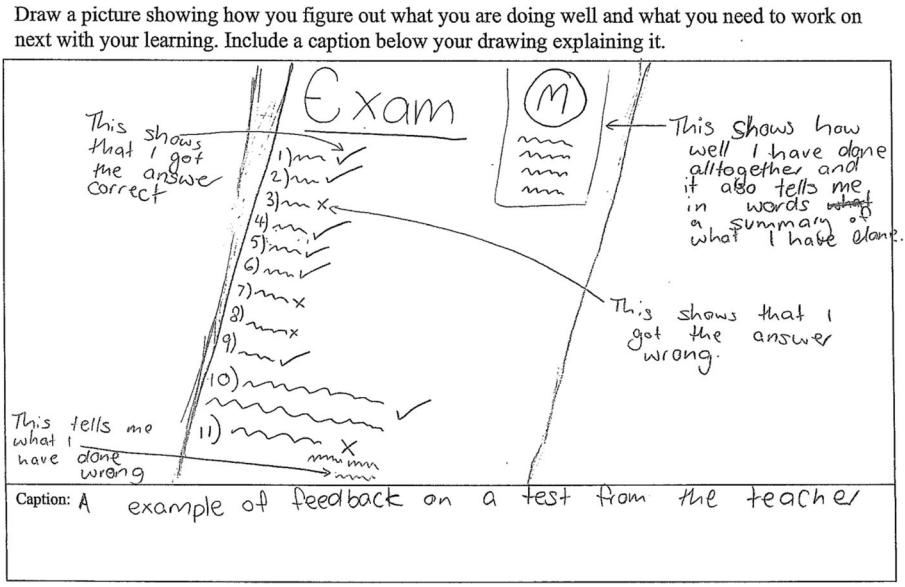

Fig. 2 Descriptive feedback

students endorsed Teacher Evaluation Feedback Practices and a Negative conception of feedback more than primary students.

6.4 Inter-correlations between survey and drawn responses

Further validation evidence of the picture analysis can be seen in the inter-correlation of the three Student Conceptions of Feedback factors and three Feedback Practices factors with the 12 major traits of the drawings (see Table 11); only significant results are discussed. When examining the relationships between the Student Conceptions of Feedback factors and the drawings, Comments for Improvement factor was positively correlated with both Student-led Feedback and Task Feedback aspects in the drawings.

Table 7 Frequencies within content of feedback category

Content of feedback	*N*	%
Describing or comparing performance	147	76
Suggestions for improvement, feed-forward	76	40
Praise, encouragement	72	37
Illegible or incomplete written comment	67	35
Effort, identifying improvement	29	15
No content	22	11
Surface learning features (spelling, grammar, basic facts, recall)	11	6
Presentation (neatness, layout, speed)	10	5
Deep learning (understanding, explaining, extending, clarifying)	10	5
Behaviour	5	3

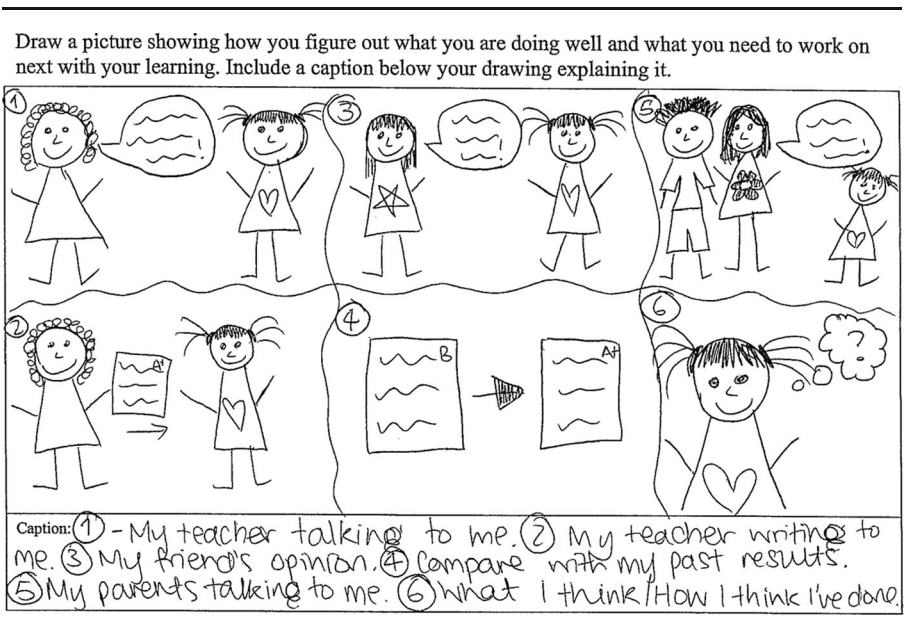

Fig. 3 Positive affect with feedback

Interpersonal Feedback was positively correlated with Student-led and Spoken Feedback and negatively correlated with Teacher Feedback. The Negative Feedback factor was negatively correlated with Student-led and Written Feedback aspects.

When examining the relationships between the Feedback Practice factors and the drawings, Teacher Evaluation was negatively correlated with Student-led Feedback, Spoken Feedback, and Positive Affect. Teacher Help was positively correlated with Teacher Feedback and negatively correlated with Student-led Feedback, Self-feedback,

Table 8 Frequencies within emotional impact category

Emotional impact	*N*	%
Smiling student	85	44
No affect	62	32
Smiling teacher	51	26
Neutral faced student	29	15
Neutral faced teacher	22	11
Positive symbols	18	9
Sad, upset, or angry student	14	7
Symbol for progress or sequence (e.g. ladder, flowchart)	7	4
Negative symbols	5	3
Sad or upset teacher	5	3

Table 9 Frequencies within student response to feedback category

Student response to feedback	*N*	%
Student accepts feedback	94	49
No response	81	42
Student action or intention to act on feedback from others	44	23
Student responds positively or feel good	36	19
Student reflects, compares results, and/or proposes own action	28	15
Student responds negatively or feels bad	13	7
Student shares feedback results with peers	10	5
Student questions teacher about feedback or requests clarification	8	4
Student shares feedback results with adult	7	4
Student rejects feedback	2	1

Positive Emotions, and Shares Feedback. Interpersonal was negatively correlated with Student-led feedback, Positive Emotions, and Shares Feedback.

While the majority of these correlations were to be expected, there were a few unpredicted results. Emotional Impact was negatively correlated with all three types of feedback practices, meaning students who selected higher numbers of items from the feedback practices checklist were less likely to include positive affective imagery in their drawings. No clear explanations exist for this result. It is possible that students experience many diverse types of feedback, with potentially conflicting messages. This

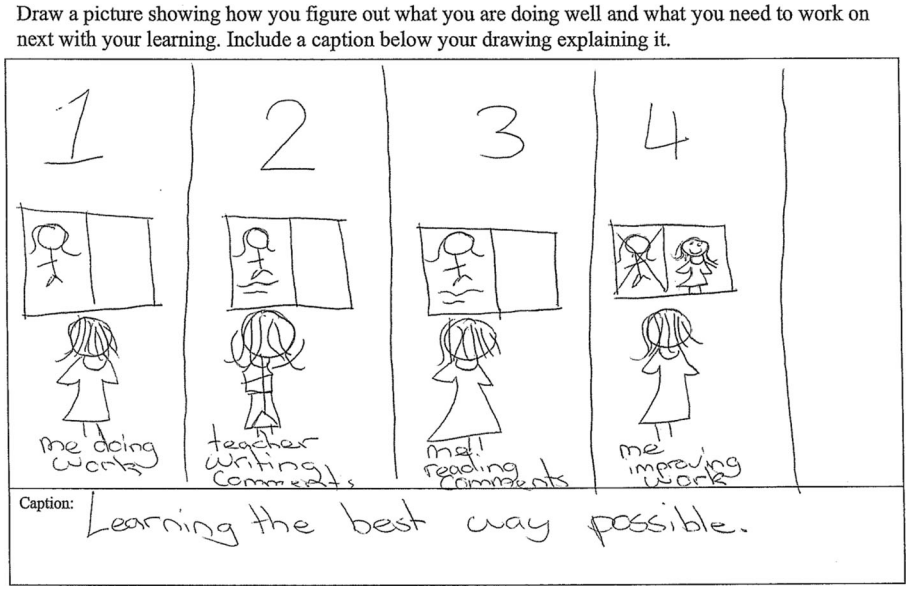

Fig. 4 Student accepting and responding to feedback

Table 10 Statistically significant sectoral differences in feedback practices, feedback conceptions, and drawn pictures by type of analysis

Data sources	Primary	Secondary			
	Mean score analysis of variance				
	M (*SD*)		*F*	*P*	Cohen's *d*
Data sources	Primary	Secondary			
Feedback practices checklist					
Teacher evaluation	0.55 (0.39)	0.77 (0.33)	16.61	<0.001	0.61
Conceptions of feedback factors					
Comments to improve	5.19 (0.69)	4.70 (0.77)	21.02	<0.001	0.67
Interpersonal	4.22 (0.81)	3.82 (1.06)	8.66	0.004	0.43
Negative	1.76 (0.77)	2.11 (0.76)	9.93	0.002	0.46
	Equivalent distribution by sector				
Drawing categories	(max. *n*=104)	(max. *n*=88)	χ^2	*p*	Cramer's V
Student-led feedback	37	6	22.68	<0.001	0.34
Spoken feedback	71	33	18.18	<0.001	0.31
Task feedback	93	66	6.97	0.008	0.19
Self feedback	55	30	6.82	0.009	0.19

may lead to confusion and less positive experiences. However, future studies are certainly required to understand this interaction properly. Likewise, the negative correlation between the Interpersonal Feedback factor and the Student-Led Feedback category in the drawings was not anticipated. This may have occurred because

Table 11 Inter-correlations of drawing traits to feedback practices and conceptions

	Feedback practices[a]			Conceptions of feedback[b]		
Drawing categories	Negative	Teacher evaluation	Interpersonal	Teacher help	Comments for improvement	Interpersonal
Teacher feedback	−0.08	−0.15*	−0.09	0.12	0.21**	0.04
Student-led feedback	0.33**	0.23**	−0.22**	−0.46**	−0.29**	−0.19**
Spoken feedback	0.14	0.19**	−0.02	−0.24**	−0.08	−0.02
Written feedback	0.04	−0.07	−0.20**	−0.13	−0.14	−0.10
Task feedback	0.25**	0.06	−0.13	−0.07	−0.01	0.12
Self feedback	−0.14	−0.18*	−0.06	0.10	0.07	−0.09
Positive affect		−0.19*	−0.20**	0.12	0.01	−0.10
Shares feedback	−0.09	−0.17*	−0.17*	0.07	0.02	−0.02

[a] *n*=182

[b] *n*=192

*p<0.05; **p<0.01

relatively small numbers of students included these representations in their drawings compared to how many nominated interpersonal practices on the checklist or because while Student-led Feedback only included drawings of peer- and self-assessment, Interpersonal Feedback contained comments from parents and non-specified others.

7 Discussion

This study adds to our understanding of how students experience and understand feedback. Consistent with previous studies with New Zealand secondary students (e.g. Peterson and Irving 2008), the source and content of feedback were strongly associated with the teacher giving spoken comments, written comments, or grades of various sorts in both fixed and free-response instruments. While some may argue that this creates an inappropriate dependence on the teacher (e.g. Carless 2011), it would appear that students have accepted that feedback from the teacher, linked to analysis and evaluation of their work, is positive, which is consistent with previous studies (Brown et al. 2009a, b). It seems students are committed to the notion that feedback contributes to growth in learning and that the relative expertise of the teacher is a legitimate source of usable and valid feedback.

On the other hand, it could be argued that, notwithstanding New Zealand's formative assessment policy framework, these data suggest feedback is still mainly practiced in a traditional, teacher-centric fashion and the student responses seen here may be simply a reflection of the ecological reality of New Zealand schools. Previous studies with New Zealand secondary students have found that they predominantly defined assessment in terms of teacher-controlled, test-like practices (Brown et al. 2009a) and that, relative to secondary students, primary and intermediate students endorsed more strongly affective, improvement, and external attribution conceptions of assessment and much less strongly the irrelevance conception of assessment (Brown and Harris 2013). While the dominance of teacher-centric, formative feedback is understandable, and possibly necessary, there is evidence that student-generated feedback is present in some classrooms, especially within primary schools, where mean endorsement of the Interpersonal Conception was stronger than for secondary students. While student mean agreement was lower for the Interpersonal Conceptions factor than the Comments for Improvement factor featuring teacher feedback, the students in this study still moderately agreed with this factor. That peer- and self-feedback were present in the drawings and student selections at all, in contrast to their absence in an earlier study of assessment drawings (Harris et al. 2009) and in contrast to the relatively low rate of selection among secondary students for informal-interactive assessment practices including PASA (Brown et al. 2009a), suggests that these practices do have traction and are accepted, especially by primary students, as legitimate sources of feedback. A comparative study in a high-stakes jurisdiction would likely show fewer student-led formative practices. The way both teacher and student-led formative practices are depicted positively in the drawings suggests that the goals of the Assessment *for* Learning policy are deeply entrenched in the minds of New Zealand students. This appears to be equally true for secondary students where the National Certificate for Educational Assessment is very much focused on delivering feedback to students to inform them as to how they can improve their performance. However, the dominance

of teacher-led practices indicates there is not universal commitment to a truly student led form of Assessment *for* Learning (Brown et al. 2012). Certainly, there will and should be debates as to the proper place of the classroom expert (i.e. the teacher) in any Assessment *for* Learning protocol.

In contrast to previous studies of New Zealand primary and secondary students' assessment conceptions, the results here are remarkably positive regarding the emotional and behavioural consequences of feedback. It seems that while students may struggle with the consequences of assessment, they are positive about the benefit of feedback for their learning, even when considering results (e.g. marks, scores, grades) as feedback. This is a very favourable result and is one that teachers should look to exploit. Survey studies in New Zealand have suggested that while students might not enjoy being tested by their teacher, there was a strong association between testing and improved performance (Brown et al. 2009a, b). This study demonstrates that part of students' tolerance of testing may be the promise of feedback giving directions for improved learning. Hence, teachers need to actively shoulder the burden of providing feedback from every assessment opportunity and encourage students to view grades, marks, and scores as feedback rather than as value or personal judgments. Needless to say, there is a long journey ahead to persuade students that feedback can come legitimately from themselves and their peers. Because other studies have pointed to the need to create psychological safety as a pre-condition for effective student generated feedback (e.g. Brown and Harris 2013; Cowie 2009), developing a supportive classroom culture may be a good first step to the meaningful implementation of formative peer- and self-assessment practices.

This study has only partially replicated previously reported results for the Feedback Practices and Conceptions of Feedback instruments. While the Student Conceptions of Feedback inventory's previously reported seven factor structure was not able to be replicated (Irving et al. 2008), this may have occurred due to this study's smaller sample size or because it incorporated the viewpoints of younger students. Nonetheless, there is conceptual integrity between the current factors and those reported in the SCoF-III. For example, the Interpersonal feedback factor consists of five of the six Peer factor items, with an additional two items not used in SCoF-III. The Negative feedback factor found here contains four of the six Ignore items and three of the five items from Trust Teacher factor, albeit with reverse loadings. The Comments for Improvement factor consists of items from five SCoF-III factors: five of the six Motivation items, both Enjoyment items, two of the Parent items, two Standards items, and one Teacher Trust item. It is clear that there is conceptual integrity around three core ideas between this study and previous results: feedback is from peers, feedback is ignored, and feedback motivates and enables my learning. What is still missing from this instrument are quality items relating to self-assessment feedback, as the items in the current version of the instrument only includes statements comparing the value of self-feedback with other forms of feedback. The instrument would benefit by developing items which ask students to evaluate the utility of self-feedback independent of other feedback sources so self-feedback's relationship to other types of feedback could be better examined.

In terms of the Feedback Practices, a noticeable difference has to do with the feedback practice item 'facts and answers to the questions' which loaded (>0.30) onto two factors (i.e. Interpersonal and Negative) and was subsequently removed from analysis. This item is clearly interactive, but seems to be negatively evaluated by students. Further investigation is needed to better understand how or why students

might consider getting the answers as a negative type of feedback, although it is also possible that this finding is an artefact of the wording of the item. A second important difference is that this structure refines the Feedback from Others factor (Irving et al. 2008) by removing teacher-based items. The current result locates the practices in three logical notions—commentary for improvement from adult teachers and parent/care-givers; error or correctness evaluations; and student-generated feedback sources.

Nonetheless, this study is only a starting point for understanding how feedback is experienced by students. While it does triangulate data from three separate instruments, it has utilised a sample that is relatively small and not nationally representative. The study took place in a policy environment that values and prioritises Assessment for Learning and so it is unlikely that results would be replicated in contexts where high-stakes consequences for both schools and students are attached to assessments. Further, despite triangulation, all three methods depend on student self-perceptions, lacking independent corroboration from observations of practice or teacher perspectives. The depictions of the content of teacher feedback to students within the pictures are obviously based on student perception and their ability to recreate the kinds of comments their teacher might make to them without relying on stereotypes. Hence, studies that directly analyse the content of spoken and written feedback students receive alongside student self-reports would be valuable contributions to the field.

Also, while this study was unable to examine what students do with the feedback they receive, the pictures displayed many students at least acknowledging and accepting feedback. However, it is not able to shed light as to when or why they might choose to accept and potentially act on this feedback. Further research is needed to understand better how students make use of feedback as part of a process of self-regulated learning. For example, a 'draw-a-picture' prompt that specifically asks students to draw what they do after they receive feedback may help shed more light on uses of feedback. Interviews in which students explain their understandings of feedback they have generated or received and their plans for next steps, corroborated with their actual revisions, would also be helpful in unpacking the reasons why student may choose to act on or ignore feedback. Additionally, more research is needed to examine the effects on motivation and learning that occur when students act on feedback that they have misinterpreted or that is incorrect (e.g. inaccurate comments). Since it is what students actually do with feedback that makes it a potential powerful tool for improving learning, these processes merit further investigation.

Acknowledgments The authors acknowledge financial support from The University of Auckland Faculty Research Development Fund grants (#3609273/9215). All data were collected and analysed in accordance with conditions set down by The University of Auckland Human Participants Ethics Committee approvals (HPEC #2008/006, #2008/053). Additionally, we thank the participating students for sharing their ideas about feedback with us. A previous version of this paper was presented to the Classroom Assessment SIG at the April 2012 Annual Meeting of the American Educational Research Association in Vancouver, BC.

Appendices

Appendix 1

Table 12 Conceptions of feedback factors

Conceptions of feedback item	Loading
Comments for improvement	
I pay attention to comments from my teacher	0.78
I enjoy getting comments on my work	0.74
Comments on my work give me information on how well I am doing	0.69
Comments from my teachers make it clear how to improve	0.69
I look at the comments and marks on my work to see what I did wrong	0.68
Comments on my work change the way I learn and study	0.67
I make active use of the comments I get from my teacher	0.67
I look forward to getting comments from my teacher about my work	0.65
When someone talks to me about my work, it makes me try harder	0.57
I know I have done well when my score or grade is better than last time	0.55
My parents/guardians provide helpful comments about my learning	0.52
Good grades will help me get the job I want	0.45
My parents/guardians expect teachers to comment on my work and say how I can improve	0.44
Interpersonal feedback	
Feedback from my classmates help my learning	0.84
Comments from my classmates about my work really help me	0.82
I learn better when my friends comment on my work	0.77
I can trust comments about my work that come from my classmates	0.64
I actively use other people's comments to help me improve	0.58
When I can't understand something, I ask another student for help	0.55
I do better when I work on something new with my friends	0.44
Negative feedback	
I can trust the comments about my work from my teacher (reverse)	−0.73
Reports sent to my parents/guardians are honest (reverse)	−0.68
Teacher comments don't tell me anything useful	0.66
The most useful comments about my work are from my teacher (reverse)	−0.62
Teachers give me honest and believable comments on my work (reverse)	−0.59
I ignore bad grades or comments	0.54
I ignore the comments teachers make about my work	0.50
Teachers' comments on my work are often hard to understand	0.49

Appendix 2

Table 13 Feedback practices

Feedback practices items	Loading
Teacher Help	
Teachers providing me with information on how to improve	0.73
Spoken comments from my teacher	0.70
Teachers writing comments about my work	0.64
Hints, tips and reminders written on my work	0.64
Teachers talking to me about my work	0.59
Teachers suggesting what I need to improve on	0.52
Teacher Evaluation	
Ticks and crosses on my work	0.81
Teachers telling you what you got wrong so you can get it right	0.79
Reports for my parents/guardians	0.71
Grades and marks	0.60
Interpersonal	
My class mates making suggestions or giving comments	0.76
Advice or comments that other people give me	0.72
Comments I give myself	0.64
Things my parents/guardians tell me	0.61

References

Andrade, H. L. (2010). Students as the definitive source of formative assessment: academic self-assessment and the self-regulation of learning. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). New York: Routledge.

Brown, G. T. L. (2004). Measuring attitude with positively packed self-report ratings: Comparison of agreement and frequency scales. *Psychological Reports, 94*, 1015–1024.

Brown, G. T. L. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters, 3*, 45–70

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), SAGE Handbook of Research on Classroom Assessment (pp. 367–393). Los Angeles: SAGE.

Brown, G. T. L., Harris, L. R., & Harnett, J. A. (2012). Teacher beliefs about feedback within an Assessment for Learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education, 28*(7), 968–978. doi:10.1016/j.tate.2012.05.003.

Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. F. (2009a). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction, 19*(2), 97–111. doi:10.1016/j.learninstruc.2008.02.003.

Brown, G. T. L., Peterson, E. R., & Irving, S. E. (2009b). Beliefs that make a difference: Adaptive and maladaptive self-regulation in students' conceptions of assessment. In D. M. McInerney, G. T. L. Brown & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about assessment for learning*. (pp. 159–186). Charlotte, NC US: Information Age Publishing.

Brown, G., Irving, S. E., & Keegan, P. J. (2008). *An introduction to educational assessment, measurement, and evaluation* (2nd ed.). Auckland, NZ: Pearson Education.

Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 93–114). New York: Routledge.

Bell, P. (2001). Content analysis of visual images. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 10–34). Los Angeles: Sage.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74.

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

Brown, J. (2007). Feedback: the student perspective. *Research in Post-Compulsory Education, 12*(1), 33–51.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research, 65*(3), 245–281.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31*(2), 219–233.

Carless, D. (2011). *From testing to productive student learning: implementing formative assessment in Confucian-Heritage settings*. London: Routledge.

Carless, D., & Lam, R. (2012). The examined life: perspectives of lower primary school students in Hong Kong. *Education 3–13: International Journal of Primary, Elementary and Early Years Education.*. doi:10.1080/03004279.2012.689988.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255.

Clarebout, G., Depaepe, F., Elen, J., & Briell, J. (2007). *The use of drawings to assess students' epistemological beliefs. Paper presented at the Biannual Conference of the European Association for Research on Learning and Instruction*. Hungary: Budapest.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46.

Courtney, M. G. R. (2013). Determining the number of factors to retain in EFA: using the SPSS R-Menu v2.0 to make more judicious estimations. Practical Assessment Research & Evaluation, 18(8). http://pareonline.net/getvn.asp?v=18&n=18

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

Croft, A. C., Strafford, E., & Mapa, L. (2000). *Stocktake/evaluation of existing diagnostic tools in literacy and numeracy in English*. Wellington: NZCER.

Crooks, T. J. (2010). Classroom assessment in policy context (New Zealand). In B. McGraw, P. Peterson, & E. L. Baker (Eds.), *The international encyclopedia of education* (3rd ed., pp. 443–448). Oxford: Elsevier.

Diem-Wille, G. (2001). A therapeutic perspective: the use of drawings in child psychoanalysis and social science. In T. V. Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 119–133). London: Sage.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research, 42*(3), 509–529.

Fraser, D. (2001). Developing classroom culture: setting the climate for learning. In C. McGee & D. Fraser (Eds.), *The professional practice of teaching* (2nd ed., pp. 15–34). Palmerston North: Dunmore Press.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304–315.

Goetz, T., Preckel, F., Pekrun, R., & Hall, N. C. (2007). Emotional experiences during test taking: does cognitive ability make a difference? *Learning & Individual Differences, 17*(1), 3–16. doi:10.1016/j.lindif.2006.12.002.

Haney, W., Russell, M., & Bebell, D. (2004). Drawing on education: using drawings to document schooling and support changes. *Harvard Educational Review, 74*(3), 241–271.

Harnett, J. A., Harris, L. R., & Brown, G. T. L. (2012). *Student feelings about assessment and feedback: Why evaluative feedback may avoid the negative affective characteristics students associate with assessment.* Paper presented at the 5th Educational Psychology Forum, Auckland, New Zealand.

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: case studies into teachers' implementation. *Teaching and Teacher Education, 36*, 101–111. doi:10.1016/j.tate.2013.07.008.

Harris, L. R., Harnett, J., & Brown, G. T. L. (2009). Assessment from students' perspectives: using pupil drawings to examine their conceptions of assessment. In D. M. McInerney, G. T. L. Brown & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 53–83). Charlotte, NC: Information Age Publishing.

Harris, L. R., Harnett, J. A., & Brown, G. T. L. (2013). *Exploring the content of teachers' feedback: What are teachers actually providing to students?* Paper presented at the American Educational Research Association Annual Meeting, San Francisco.

Hattie, J. (2009). *Visible learning: a synthesis of meta-analyses in education*. London: Routledge.

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

Hattie, J. A., & Brown, G. T. L. (2008). Technology for school-based assessment and assessment for learning: development principles from New Zealand. *Journal of Educational Technology Systems, 36*(2), 189–201.

Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: reconsidering the role of assessment feedback in student learning. *Studies in Higher Education, 27*(1), 53–64.

Hyland, F. (2003). Focusing on form: student engagement with teacher feedback. *System, 31*(2), 217–230.

IBM. (2011). *Amos 20 [computer program] (version Build 817)*. Meadville: Amos Development Corporation.

Irving, S. E., & Peterson, E. R. (2007). *Student conceptions of feedback (SCoF) inventory (version 3) test*. Auckland: University of Auckland.

Irving, S. E., Peterson, E. R., & Brown, G. T. L. (2008). *Feedback and academic achievement: The relationship between students' conceptions of feedback and achievement*. Paper presented at the 6th Biennial Conference of the International Test Commission, Liverpool, UK.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Lam, T. C. M., & Klockars, A. J. (1982). Anchor point effects on the equivalence of questionnaire items. *Journal of Educational Measurement, 19*(4), 317–322.

Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: assessment for learning or assessment of learning? *Assessing Writing, 12*(3), 180–198.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320–341.

McKillop, C. (2006). Drawing on assessment: using visual representations to understand students' experiences of assessment in art and design. *Art, Design & Communication in Higher Education, 5*(2), 131–144. doi: 10.1386/adch.5.2.131/6.

Ministry of Education. (2010). Ministry of Education Position Paper: Assessment [Schooling Sector]. http://www.minedu.govt.nz/theMinistry/PublicationsAndResources/AssessmentPositionPaper.aspx

Mizutani, S. (2009). The mechanism of washback on teaching and learning (unpublished doctoral dissertation). The University of Auckland, Auckland, NZ

Nesbit, P. L., & Burton, S. (2006). Student justice perceptions following assignment feedback. *Assessment & Evaluation in Higher Education, 31*(6), 655–670.

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

Pajares, M. F., & Graham, L. (1998). Formalist thinking and language arts instruction: teachers' and students' beliefs about truth and caring in the teaching conversation. *Teaching & Teacher Education, 14*(8), 855–870.

Peterson, E. R., & Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*(3), 238–250.

Picker, S., & Berry, J. (2000). Investigating pupils' images of mathematicians. *Educational Studies in Mathematics, 43*, 65–94.

Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: the students' perspective. *Assessment & Evaluation in Higher Education, 33*(2), 143–154.

Rea, L. M., & Parker, R. A. (1992). *Designing and conducting survey research*. San Francisco: Jossey-Boss.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.

Strijbos, J., & Sluijsmans, D. (2010). Unravelling peer assessment: methodological, functional, and conceptual developments. *Learning and Instruction, 20*(4), 265–269.

Tjeerdsma, B. L. (1997). Comparison of teacher and student perspectives of tasks and feedback. *Journal of Teaching in Physical Education, 16*(4), 388–400.

Topping, K. J. (2010). Peers as a source of formative feedback. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 61–74). New York: Routledge.

Topping, K. (2013). Peers as a source of formative and summative assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 395–412). Los Angeles: SAGE.

Tunstall, P., & Gipps, C. (1996a). Teacher feedback to young children in formative assessment: a typology. *British Educational Research Journal, 22*(4), 389–404.

Tunstall, P., & Gipps, C. (1996b). 'How does your teacher help you to make your work better?' Children's understanding of formative assessment. *Curriculum Journal, 7*(2), 185–203.

van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: the role of interpersonal variables and conceptions. *Learning & Instruction, 20*(4), 280–290.

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education, 31*(3), 379–394.

Wheelock, A., Bebell, D., & Haney, W. (2000a). Student self-portraits as test-takers: variations, contextual differences, and assumptions about motivation. Teachers College Record, ID Number: 10635

Wheelock, A., Bebell, D., & Haney, W. (2000b). What can student drawings tell us about high-stakes testing in Massachusetts? Teachers College Record, ID Number: 10634