Teaching and Teacher Education 28 (2012) 968–978

Contents lists available at SciVerse ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being

教师关于学习评价反馈的信念：提升学生健康学习的支持服务

Gavin T.L. Brown*, Lois R. Harris, Jennifer Harnett

School of Learning, Development, and Professional Practice, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

HIGHLIGHTS

- ► *Teachers' Conceptions of Feedback* inventory validated with 518 New Zealand teachers.
- ► 9 Belief and 4 practices factors were invariant for primary and secondary teachers.
- ► 5 Beliefs predicted feedback practices, with small to medium effects.
- ► Teachers endorsed feedback for learning progress rather than well-being.
- ► Espoused conceptions of feedback were consistent with assessment for learning policy.

ARTICLE INFO

Article history:
Received 6 October 2011
Received in revised form
11 May 2012
Accepted 11 May 2012

Keywords:
Attitudes
Values & beliefs
Feedback
Teachers
Survey research
Structural equation model

ABSTRACT

教师关于反馈的信念目录　　统计不变

Teachers' understandings of feedback probably influence the type and quality of feedback that they provide. The beliefs of 518 New Zealand practicing teachers about feedback were evaluated with the *Teachers' Conceptions of Feedback* (TCoF) inventory and related to practices these teachers considered to be feedback. Nine feedback factors and four practices factors were found and models were statistically invariant between primary and secondary teachers. New Zealand teachers' understandings of feedback were strongly focused on improving learning instead of enhancing student well-being. Similar factors are expected in other contexts, though agreement rates should reflect local policy priorities and cultural values.

1. Introduction

反馈在促进学习中的重要性

Many international experts consider feedback to be an important element of Assessment for Learning (Black, Harrison, Lee, Marshall, & Wiliam, 2003; Clarke, 2003; Hattie, 2009; Sadler, 1989, 1998), with Hattie and Timperley (2007, p. 102) calling it "among the most critical influences on student learning". It can increase learner satisfaction and persistence (Kluger & DeNisi, 1996), and contribute to students adopting more productive learning strategies (Vollmeyer & Rheinberg, 2005). However, what counts as 'good feedback' is contested (Shute, 2008), with feedback considered the element of formative assessment "most laden with a legacy of bad practice and misguided views" (Clarke, 2003, p. 3). Feedback, when provided inappropriately, can lead to negative effects. Kluger and DeNisi (1996) found that feedback actually decreased student performance in a third of the studies analysed.

Nevertheless, feedback continues to be endorsed worldwide as a powerful strategy for teachers of all subjects and grade levels (Leahy, Lyon, Thompson, & Wiliam, 2005). This attitude towards feedback is consistent with the Assessment for Learning (AfL) strategy which focuses 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 & Wiliam, 1998). AfL policy reforms often aim to increase student evaluation of their own progress using rubrics, targets, and pedagogical interactions with their teachers, with the overall goal of developing self-regulating learners (Leahy et al., 2005).

Despite the power that teachers commonly exercise over the delivery of feedback, there has been little research to date investigating teachers' conceptions of feedback, with most work examining their enacted practices (e.g., Torrance & Pryor, 1998; Tunstall & Gipps, 1996). Conceptions consist of beliefs, attitudes, and intentions

* Corresponding author. Tel.: +64 9 623 8899x48602; fax: +64 9 623 8827.
E-mail address: gt.brown@auckland.ac.nz (G.T.L. Brown).

0742-051X/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tate.2012.05.003

that people have (Brown, 2008; Thompson, 1992) and are significant contributors to behaviour (Ajzen, 2005). The beliefs teachers have about educational processes matter since they have been demonstrated to contribute meaningfully to the actions that teachers take (Pajares, 1992; Rubie-Davies, Flint, & McDonald, 2011; Woolfolk Hoy, Davis, & Pape, 2006). Hence, in jurisdictions that rely heavily on an Assessment for Learning approach to educational assessment (e.g., New Zealand), the beliefs that teachers have about the nature and purpose of feedback will likely affect the quality of AfL implementation. Studying understandings and practices of feedback in contexts that already promote AfL may shed light on how AfL might be enacted in international settings considering the adoption of this assessment strategy. This paper contributes to the field by presenting data from a national survey study of New Zealand teachers' espoused conceptions of feedback. These conceptions were then mapped to practices teachers perceive as part of feedback to students. 教师对于反馈的概念

反馈的概念：一个代理(教师、同伴、书,家长,自己,经验)根据相关的理解或功能所提供的信息。

2. Understanding feedback

In their review of international literature, Hattie and Timperley (2007, p. 81) defined feedback as "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding," highlighting that feedback legitimately comes from non-teacher sources. When examining research about feedback, opinion varies about *who* should provide feedback, *how and when* it is best delivered, *what* the content of feedback should be, and *why* it should be provided. As these debates have been reviewed extensively by others (Butler & Winne, 1995; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Shute, 2008), this section will briefly review the main ideas relevant to feedback in compulsory schooling and examine research available on teacher understandings of feedback.

2.1. Who should provide feedback? 谁应该提供反馈

Traditionally, teachers have been responsible for giving feedback. However, during the last two decades, with the international rise of student-centred pedagogy and Assessment for Learning policies, there is increasing agreement that students are legitimate sources of feedback (Andrade, 2010; Black et al., 2003; Strijbos & Sluijsmans, 2010). Peer- and self-assessment practices encourage students to identify learning objectives and understand the criteria used to judge their work, with the goal of increasing self-regulation (Andrade, 2010). Using students as a source of feedback can potentially mitigate teacher feedback problems related to timeliness and frequency (Andrade, 2010) and perceived psychological risks for students (van Gennip, Segers, & Tillema, 2010). While teacher feedback is conventionally regarded as more accurate, Topping (2010) has suggested that peer and student feedback is no less reliable and valid than teacher feedback. However, students still require training in these practices (Andrade, 2010; Brown & Harris, in press; Gielen, Peeters, Dochy, Onghena, & Struyven, 2010) and the nature of this training will be influenced, in part, by the classroom teacher's understandings of feedback.

The validity and effectiveness of feedback from peers and the self is dependent on interpersonal relationships and psychological issues related to self-disclosure and trust (Cowie, 2009; Peterson & Irving, 2008; van Gennip et al., 2010), requiring students to take on the complex role of assessor (Topping, 2010). Research has indicated that some students and teachers question the validity and reliability of the feedback received through these practices (e.g., Harris & Brown, 2010; Harris, Harnett, & Brown, 2009; Peterson & Irving, 2008; Ross, 2006), with studies showing that, especially among younger students, peer feedback cannot be expected to be as accurate as expert feedback (Gielen et al., 2010). van Gennip et al. (2010) found that while student 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. Hence, part of the challenge of implementing such practices is convincing students and other stakeholders that pupils can be effective assessors of their own work.

同伴和自我的反馈信效度依赖于个体的人际关系、信任度以及自我表露，需要学生承担复杂的评估者的角色

2.2. How and when is feedback best delivered?

语文老师

How and when feedback is delivered (e.g., written, spoken, grades or scores) also affects its influence on student learning. Within an Assessment for Learning framework, verbal teacher feedback is often privileged; Cowie and Bell (1999) described this process as "interactive formative assessment," where teachers notice, recognise, and respond to student thinking in an unplanned and spontaneous manner during teacher–student interactions *within the learning process*. Nevertheless, some researchers argue that written comments are preferable to 反馈的时效性问题：学习任务，学习者水平，学习过程 students can revisit them (Kluger & DeN

Feedback's timing is an important variable (Hattie & Timperley, 2007; Shute, 2008), with consensus that it is best provided to learners *during*, not after, the learning process. Hattie and Timperley (2007) suggested that the optimal timing of feedback might vary based on the feedback's content. While simple error correction may be most effective if provided immediately, for feedback relating to processes (e.g., reading fluency) or complex tasks, delaying the feedback allows students time to attempt the task without interruption. Shute (2008) identified that while high-achieving learners may benefit from delayed feedback, low-achieving learners may need more immediate feedback. Thus, there is considerable debate within the literature as to how and when feedback is best delivered, with teachers likely to have divergent views. 成绩作为反馈的问题也是有争议的

The issue of grades as feedback is also contentious, with some arguing they negatively affect student motivation, distracting learner attention away from more constructive narrative feedback (e.g., Assessment Reform Group, 2002). The perceived negative effects of grades have led some to advocate practices, such as 'comment only marking', to eliminate this kind of evaluation entirely (Black et al., 2003). Hattie and Timperley (2007) noted that stakeholders seldom use grades diagnostically; therefore, the problem may lie in how grades are used rather than in grades themselves. Furthermore, regardless of stakeholder opinions, grading-type feedback is normally required in most education settings. 反馈应该包含什么和为什么应该提供反馈 利益相关者

2.3. What should feedback contain and why should it be given?

Theoretically and empirically based models describe multiple types of feedback content, each with differing purposes or outcomes (Askew & Lodge, 2000; Butler & Winne, 1995; Hargreaves, 2005; Hattie & Timperley, 2007; Shute, 2008; Tunstall & Gipps, 1996). Hattie and Timperley's (2007) review of feedback literature identified four types of feedback, along with factors mediating their effectiveness: Feedback Task (i.e., whether work was correct or incorrect), Feedback Process (i.e., comments about the processes or strategies underpinning the task), Feedback Self-regulation (i.e., reminders to students about strategies students can use to improve their own work), and Feedback Self (i.e., non-specific praise and comments about effort). Research indicates self-regulation feedback leads to greater student engagement, effort, and self-efficacy, making it the most powerful type (Butler & Winne, 1995; Hattie & Timperley, 2007). However, task feedback is the type most

frequently provided to students in classrooms (Black & Wiliam, 1998; Harnett, 2007; Hattie & Timperley, 2007). 有争议的

The role of praise in feedback is particularly contentious. Previous research (e.g., Black & Wiliam, 1998; Brophy, 1981; Hattie & Timperley, 2007; Sadler, 1989; Wiliam, 1999) has found that teachers predominantly provide feedback in the form of "low-level" praise or criticism of students' verbal responses and written work. Praise alone is rarely effective because it lacks sufficient information to move students forward in their learning (Cowie, 2005; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Shute, 2008). However, Black and Wiliam (2009) have argued that praise can enhance motivation and that judgements (even negative ones) can be viewed as challenges to overcome. There is also evidence that some teachers believe praise is instrumental in improving student self-esteem (Irving, Harris, & Peterson, 2011). Deciding when and how to use praise is a complex issue for teachers as they attempt to promote student well-being alongside academic learning.

There are contrasting reasons for providing feedback. Educationally, feedback aims to enhance student learning, moving students along a growth pathway, as opposed to protecting their sense of personal well-being (Boekaerts & Corno, 2005). There is evidence that some teachers also provide feedback for affective reasons, hoping to encourage student persistence and effort or to mediate negative evaluations of student work (Black & Wiliam, 2009; Irving et al., 2011). However, it is doubtful that this affective-psychological rationale for feedback is effective in improving students' learning outcomes because students need teachers as a source of accurate information about their performance (Pajares & Graham, 1998). Furthermore, teachers are often required by administrators or external stakeholders to provide certain types of feedback (e.g., grades, norm-referenced reports), regardless of teachers' beliefs or even best practice research.

When examining the effects of feedback, several factors influence its use and mediate its effect on student learning. Foremost is students' ability to interpret and use the feedback and their motivation to do so (Sadler, 2010). Additionally, the type, content, timing, complexity, and accuracy of the feedback contribute to its effectiveness. These factors are often under the control or supervision of the classroom teacher; consequently, understanding teacher thinking seems critical if AfL reforms are to be implemented.

2.4. What do teachers believe? 很少有研究明确探究教师关于反馈的性质和目的信念

Few studies have explicitly examined teacher beliefs about the nature and purpose of feedback. Irving et al. (2011) found the New Zealand teachers in their study described three types of feedback: spoken or written comments about learning, grades or marks, and spoken or written comments about behaviour or effort. Teachers identified four main purposes for these three types of feedback: improving student learning (e.g., providing information about weaknesses in student work and how to correct them), reporting and compliance (e.g., giving grades, hinting to students about their final results), and encouraging students (e.g., praise, feedback about effort). The fourth purpose was that feedback, in certain circumstances, served no function whatsoever because students did not act on it, making it irrelevant. Narrative feedback given with a final grade was most strongly related to irrelevance because the teachers believed most students ignored such comments.

O'Quin (2009) surveyed 308 middle school teachers in Louisiana and reported that the endorsement of feedback as a strategy to enhance student learning predicted the practices teachers described in their definitions of feedback. This relationship was much stronger for teachers of alternative as opposed to regular classrooms ($\beta = .43$; .09 respectively). Interestingly, endorsement of the notion that feedback was done because it was required did not have a statistically significant relationship to the kinds of feedback the Louisiana teachers described in their definitions. This would indicate that teachers' beliefs that feedback improves learning might determine their feedback practices more than external factors. Clearly, studies with different teacher populations and in differing policy contexts are required.

3. Methods 大规模调查问卷

To contribute to knowledge about teacher conceptions of feedback, this paper reports the results of a large-scale questionnaire survey of New Zealand teachers' conceptions of feedback. A non-experimental survey of a nationally representative sample of practicing teachers was used to test the validity of a theoretically devised set of constructs concerning teachers' conceptions of feedback. This section will begin by describing New Zealand's assessment context and its place in the global Assessment for Learning movement before providing details about the research questions and hypotheses, survey instrument, sampling, data collection, and data analysis techniques used.

3.1. Feedback within a New Zealand context 在新西兰国家中的反馈研究 学习评价框架

New Zealand has a relatively unique assessment environment, likely conducive to strong teacher endorsement of formative feedback. While Assessment for Learning has been strongly endorsed by educators in many internationally contexts (e.g., the United Kingdom with Black et al., 2003), New Zealand is one of very few that have formally adopted it as a dominant assessment framework. The New Zealand Ministry of Education (2010, p. 5) stated, "We have a deliberate focus on the use of professional teacher judgment underpinned by Assessment for Learning principles rather than a narrow testing regime". 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, with learning outcomes and objectives specified across multiple levels.

At the time of this survey, there was no compulsory, state mandated assessment regime prior to Year 11. Hence, for teachers within this study, all assessment practices were voluntary and low stakes, making it possible for teachers to implement a range of feedback practices without the threats presented by nationally mandated testing and accountability programs. However, while the primary sector commonly uses informal assessments and standardised tests, usually for the purpose of improving instruction and student learning, secondary school assessment is often focused on preparing for or implementing the high-stakes student qualifications system (i.e., the National Certificate of Educational Achievement) which begins formally in the third year of secondary schooling when students are about age 15 (Croft, Strafford, & Mapa, 2000; Crooks, 2010; Hill, 2000). Consequently, the influence of high-stakes testing on feedback practices is likely to become stronger as students progress through high school, even though it is not actually implemented until Year 11.

3.2. Research questions and hypotheses 教师反馈的信念

Two research questions guided the development and use of the Teacher Conceptions of Feedback (TCoF) survey instrument and subsequent measurement models of teacher definitions and conceptions of feedback: 随后的测量模型

实践教师对反馈持什么样的观念他们联系这样的术语实践什么？

1. What conceptions of feedback do practicing teachers hold and what practices do they associate with this term?
2. What is the relationship between teachers' *conceptions* of feedback and the *practices* they associate with feedback?

j教师关于反馈的观念与他们在实践中应用反馈之间的关系

Question 1 was answered by finding the factorial structure which best-represented teacher thinking about feedback conceptions and practices. Thus, we first sought to establish whether our ten conceptions of feedback factors could be found in the teacher responses. Then, we explored the structure of the practices that teachers perceived as feedback to ascertain how these could be aggregated. Question 2 was answered by examining the regression paths between the conceptions of feedback factors to the perceived feedback practices factors in a structural equation model. Since we expected that Teacher Conceptions of Feedback (TCoF) factors would have meaningful relationships to the practices teachers perceived as feedback, we developed five hypotheses about those relationships. Specifically,

结构方程模型

5个假设

H1. Feedback conceptions related to improvement, task, process, self-regulation, and timeliness would predict teacher formative feedback practices.
H2. Feedback conceptions related to student sources would predict non-teacher feedback practices (e.g., use of peer and self-feedback).
H3. Feedback conceptions related to encouragement and the self would predict praise and self-esteem building feedback practices.
H4. Accountability conceptions of feedback would predict feedback practices designed to inform parents (e.g., reports).
H5. The conception that feedback is irrelevant or useless would not predict any feedback practices.

建立自尊

问责的反馈观念预示反馈实践旨在告知父母

3.3. Instruments 工具

教师关于反馈详细内容的观念

3.3.1. Teachers' conceptions of feedback inventory

Because no self-report survey inventory existed to measure teacher conceptions of feedback, the Teachers' Conceptions of Feedback (TCoF) questionnaire was devised (Harris & Brown, 2008), drawing primarily on work by Hattie and Timperley (2007) and Irving et al. (2011). Items related to ten feedback constructs were drafted. The first four factors related to Irving, Harris, and Peterson's four purposes of feedback (i.e., irrelevance, improvement, reporting and compliance, and encouragement). The next four factors were related to Hattie and Timperley's (2007) four feedback types (i.e., task, process, self-regulation, and self). The final two factors were related to questions arising from the feedback literature. Factor 9 related to the validity of self- and peer feedback, while Factor 10 related to feedback's timing. While designed as 'independent' factors, intercorrelations were expected, as some factors were conceptually similar (e.g., the encouragement purpose and the self-type). These ten factors are listed below with a sample item provided for each:

Purposes 目的

没有相关性

Irrelevance: Feedback is pointless because students ignore my comments and directions.
Improvement: Students use the feedback I give them to improve their work.
Reporting and compliance: At my school, teachers are expected to give both spoken and written feedback to students.
Encouragement: The point of feedback is to make students feel good about themselves.

Types 类型

Task: My feedback is specific and tells students what to change their work.
Process: I give students opportunities to respond to my feedback.
Self-regulation: Feedback is about helping students evaluate their own work.
Self: My feedback includes comments on the effort students put into their work.

Other

Peer and self-feedback: Students are able to provide accurate and useful feedback to each other and themselves.
Timeliness of feedback: Students should not have to wait for feedback.

Respondents used a six-point, positively-packed agreement rating scale known to generate discrimination in contexts of social desirability (Brown, 2004). Responses were coded: strongly disagree = 1, mostly disagree = 2, slightly agree = 3, moderately agree = 4, mostly agree = 5, and strongly agree = 6.

3.3.2. Practices perceived as feedback checklist 实践视为反馈清单

Additionally, teachers were presented with a list of 17 different feedback practices commonly used in New Zealand schools. The teachers were asked to respond to the prompt, "When I think of feedback, I think of the following practices: (tick all that apply)". These options are listed in Table 4. This technique is an adaptation of previous studies with secondary students (Irving, Peterson, & Brown, 2007, 2008), which indicated that selection patterns could be identified with cluster analysis and confirmatory factor analysis demonstrated good quality of fit to the data. This analytic approach provides an effective way to place feedback practices into conceptually meaningful groupings. In this way, insights into the relationship of purposes to practices that define feedback could be examined. Exploratory factor analysis of the responses was used to simplify these into conceptually related groups of practices.

探索性因素分析的反应被用来简化这些概念上相关的实践。

参与者情况介绍

Survey forms were sent out to New Zealand primary and secondary schools selected to create a representative national sample considering school size, region, and socio-economic strata. When forms were returned blank, they were resent to a school with similar stratification. School principals distributed questionnaires to volunteer teachers who returned their questionnaires directly to the research team in postage paid envelopes. In total, 1492 teacher surveys were delivered to 457 schools.

Over 500 (n = 518) teachers returned valid, completed questionnaires, constituting a 35% return rate. Of these, 72% were female (n = 374) and 82% were of New Zealand European ethnicity (n = 422). These proportions are consistent with the 2004 Teacher Census (New Zealand Ministry of Education, 2005) which had 80% of respondents identify as European/Pākeha[1]; 82% of primary and 58% of secondary teachers were female. Just over $^3/_4$ had taught for six or more years with 56% having taught more than 10 years. Approximately half (52%) described themselves as a teacher with no additional responsibilities (e.g., department head, dean, director, manager, or subject specialist). Just over half of the sample worked in primary schools (n = 276, 53%), 238 (46%) were secondary teachers, and four were unclassified.

3.5. Analysis

验证性因素分析

Confirmatory factor analysis was used to recover the 10 constructs from the original TCoF design. Exploratory factor

探索性因素分析

[1] Pākeha is the indigenous Māori word for white people.

analysis was utilised to identify the dimensionality of the Feedback Practices instrument and structural equation modelling was used to evaluate the relationship of TCoF to definitions of feedback. All cases with >7 missing values for the TCoF were dropped from analysis and missing values in the balance of data (average 1.5% missing per item) were imputed using the expectation maximisation procedure (Dempster, Laird, & Rubin, 1977). 负载系数：>0.3

In accordance with convention, items were accepted as belonging to their intended factor only when their loading was >.30 and if cross-loadings, as indicated by modification indices, were low (Bandalos & Finney, 2010). Maximum likelihood confirmatory factor analysis of the variance–covariance Pearson correlation matrices, using AMOS software (Arbuckle, 2008), was used throughout to test and trim the measurement and structural models. In line with suggested practice (Cheung & Rensvold, 2002; Fan & Sivo, 2007; Marsh, Hau, & Wen, 2004), models with statistically non-significant χ^2 per df, GAMMA hat > .90, root mean square errors of approximation (RMSEA) < .08, and standardised root mean residuals (SRMR) close to .06 were considered sufficiently close to the data to not be rejected. The greatest lower bound (glb) scale reliabilities were estimated using TiaPlus (Heuvelmans, 2010); the glb provides a more accurate estimate of the lower bound of internal reliability than Cronbach's alpha (Sijtsma, 2009). Nonetheless, since the various scales are intercorrelated, a superior method for determining the sufficiency of the scales is the much stricter model testing involved in establishing that confirmatory factor analysis and structural equation models fit the data (Bryant & Yarnold, 1995).

Structural equation modelling was used to test the hypotheses about how feedback conceptions related to feedback practice definitions. In addition, nested, multi-group invariance testing (Byrne, 1989) was utilised to determine whether the various models were equivalent according to the level of employment. Equivalence of responding requires identical pathway configurations (i.e., RMSEA < .05), equivalent factor to item regression weights and intercepts, equivalent 2nd-order factor to 1st-order factor regression weights and intercepts, and equivalent covariances between the 2nd-order factors. Differences in the CFI of more than .01, as each parameter is constrained to be equivalent across groups, indicate that the contrasting groups responded differently to an instrument (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000).

The effect of structural relationships is determined by selecting statistically significant predictors and by examining the amount of variance explained by those predictors. Interpretation of the variance explained by regression paths (f^2) is guided by Cohen (1992).

4. Results

教师应对TCoF库存的结构

4.1. The structure of teacher responses to the TCoF inventory

层次结构模型

Since the questionnaire was designed to have ten factors, first, we tested whether the ten factors existed as proposed using a hierarchical model in which a 2nd-order general TCoF factor predicted the 10 feedback factors. After removing items with strong modification indices and merging the Encouragement and Self constructs (now named Encouragement), an acceptably fitting measurement model of nine factors predicted by a general TCOF factor was found (k = 38 χ^2 = 1626.22; df = 656; χ^2/df = 2.48, p = .12; CFI = .81; gamma hat = .91; RMSEA = .053, 90% CI = .050–.057; SRMR = .062) (Table 1).

The process used to create this well-fitting model of teacher conceptions of feedback resulted in some simplification of the complexity in the intended factors. For example, the Irrelevance factor focused on students ignoring teacher feedback and excluded

Table 1
Teachers' conceptions of feedback factors and items.

Factors and items	Loading
I Conception – irrelevance (students ignore)	
21. Feedback is pointless because students ignore my comments and directions	.66
2. Students rarely make changes in their work in response to my feedback	.63
42. I seldom give written feedback because students throw it away	.39
52. Time spent giving feedback is wasted effort	.34
II Conception – improvement (student use)	
2. Students use the feedback I give them to improve their work	.62
12. I can see progress in student work after I give feedback to students	.67
33. Students use comments I give them to revise their work	.62
53. Giving students feedback is important because it helps them learn	.42
III Conception – accountability (expected)	
13. Parents can tell how well their child is learning from my feedback	.56
34. At my school, teachers are expected to give both spoken and written feedback to students	.42
62. Feedback practices at my school are monitored by school leaders	.39
IV Conception – encouragement + self-type (praise)	
49. Feedback should be full of encouraging and positive comments	.72
58. Teachers should always include praise in their feedback about student work	.62
35. The goal in giving feedback is to protect and enhance the student's self-esteem	.54
45. Good feedback praises students	.54
66. My feedback includes comments on the effort students put into their work	.54
24. The point of feedback is to make students feel good about themselves	.43
V Task type (task)	
55. My feedback helps students decide what to include and/or exclude in their work	.73
36. My feedback is specific and tells students what to change their work	.58
15. My comments help students create the kind of work I expect from them	.51
VI Process type (process)	
68. I organise time in class for students to revise, evaluate, and give themselves feedback about their own individual work	.64
56. In feedback, I describe student work to stimulate discussion about how it could improve	.61
47. I give students opportunities to respond to my feedback	.59
64. Feedback is a two-way process between my students and me	.54
VII Self-regulation type (SR)	
38. My feedback reminds each student to self-assess his or her own work	.62
57. My students generate ideas about improving their learning independent of me	.61
17. Feedback is about helping students evaluate their own work	.58
69. My students analyse their own work with little direction from me	.45
7. I encourage students to correct/revise their own work without my prompting	.43
VIII Peer & self (PASA)	
40. Students are able to provide accurate and useful feedback to each other and themselves	.73
50. Students can be critical of their own work and can find their own mistakes	.62
59. Peers are the best source of feedback	.43

Table 1 (*continued*)

Factors and items	Loading
IX Timeliness (prompt)	
41. I give students feedback immediately after they finish	.65
29. Students should not have to wait for feedback	.63
10. I aim to deliver feedback to students within two days of receiving their work	.55
67. Quality feedback happens interactively and immediately in the classroom while students are learning	.55
20. Feedback that takes more than a week to get to the student is useless	.50

Note. Item numbers refer to order presented in inventory; loading values are standardized beta regression weights.

items related to students preferring grades or marks to written comments. The Improvement factor excluded items related to teacher intentions or values for feedback, focussing on items describing students using the feedback they received. Reporting and Compliance contained statements indicating feedback should inform parents of student progress, but excluded items about grading practices per se. The Encouragement factor included statements suggesting that providing students with praise would boost self-esteem, but did not include items indicating that encouragement motivates greater student effort. The Task factor focused on giving students information about aspects of their work that could be improved rather than on accuracy or specific error correction. The Process factor focused on allowing students to engage actively in responding to feedback. The Self-regulation factor included items about student autonomy and agency in evaluating their own work. The Peer and Self-Feedback factor focused on students actively giving themselves and each other feedback, not on its accuracy or reliability. The Timeliness factor included items relating to the importance of prompt response to student work. The implications of this narrowing of each feedback factor are discussed later in the paper.

教师如何看待反馈

The values of the regression paths from the general conception of feedback (i.e., the second-order factor) indicate how much each factor is part of the way teachers think about feedback. Feedback was most strongly defined by the factors Process (β = .99), Self-regulation (β = .93), Improvement (β = .92), and Reporting and Compliance (β = .90). Three constructs were moderately predicted (i.e., Peer and Self-Feedback β = .75, Task β = .72, and Timeliness β = .65). Encouragement was weakly predicted (β = .30) and the Irrelevance construct was negatively predicted (β = −.48). These paths show that the nine factors formed three general groups in terms of how strongly they expressed teachers' thinking about feedback.

An important criterion of factor analysis is that there is evidence that the factors exist independently of each other; otherwise, there is no need to have so many factors. Strong inter-correlations between the factors would indicate that the various conceptions of feedback do not actually form separate entities in teachers' thinking. However, the TCoF factor inter-correlations were weak to moderate (Table 2) with the strongest inter-correlation between the Process and Self-regulation constructs ($r = .67$). The Irrelevance factor was negatively correlated with all other constructs. This level of inter-correlation suggests that teachers' thinking about each factor is relatively independent.

4.1.1. Stability of model by level of teaching

In order to claim that the TCoF elicited similar responses from both primary and secondary teachers, the statistical parameters of the model needed to be equivalent for both groups. Separating the data into two groups caused the residual variances for two factors (i.e., Process and Reporting & Compliance) to become negative. These were fixed to .005 because the observed values were less than one standard error below zero, indicating their true value was not negative. The RMSEA fit index was .04, indicating that the configuration of paths was identical for the two groups. Inspection of the change in χ^2 and CFI for each nested level of equivalence showed that the TCoF measurement model, described above, was invariant across level of teaching.

Table 2
TCoF factor inter-correlations.

TCoF scales	Teacher conceptions of feedback								
	1	2	3	4	5	6	7	8	9
1. Irrelevance	(.66)								
2. Improvement	**−.42**	(.71)							
3. Reporting & compliance	−.17	**.46**	(.48)						
4. Peer & self-feedback	−.15	.39	.33	(.82)					
5. Task	−.22	**.53**	.36	.22	(.64)				
6. Process	−.29	**.60**	**.47**	**.50**	**.45**	(.72)			
7. Self-regulation	−.18	**.58**	**.40**	**.53**	**.42**	**.67**	(.72)		
8. Timeliness	−.06	.39	.27	.35	.27	**.47**	**.43**	(.63)	
9. Encouragement	.08	.13	.26	.14	.21	.24	.22	.16	(.74)

Note. glb estimate of scale reliability reported on diagonal; values > .40 marked in bold.

It becomes possible to compare the mean scores for each factor after establishing the statistical equivalence of responding to the TCoF items for the two groups of teachers. The mean scores for the nine factors ranged from less than '2 = mostly disagree' for Irrelevance to almost '5 = mostly agree' for Improvement (Table 3). There were statistically significant ($p < .01$) and practically significant ($d > .50$) differences in mean score between primary and secondary teachers with the former endorsing six of the nine factors much more than the latter. Differences were trivial ($d \leq |.21|$) for Irrelevance, Task, and Encouragement.

4种类型的反馈实践

4.2. Four types of practices perceived as feedback

无教师反馈、教师的形成性反馈
保护性评价反馈、家长报告

After cluster analysis, a four-factor [redacted] feedback, Teacher Formative feedback, Teacher Protective Evaluative feedback, and Parent Reporting) of teachers' self-reported feedback practices was tested in a hierarchical measurement model. This model had good fit characteristics ($k = 17$; $\chi^2 = 321.08$; $df = 115$; $\chi^2/df = 2.79$, $p = .09$; CFI = .90; gamma hat = .96;

分层测量模型

Table 3
TCoF and feedback practices scale statistics by level.

Scale	Total	Primary	Secondary	Mean score differences		
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	$F_{(1, 512)}$	*p*	Cohen's *d*
TCoF						
1. Irrelevance	1.77 (.61)	1.71 (.57)	1.83 (.65)	5.43	.02	−.20
2. Improvement	4.78 (.66)	4.95 (.60)	4.59 (.66)	41.24	<.01	.57
3. Reporting & compliance	4.06 (.88)	4.32 (.90)	3.77 (.76)	55.98	<.01	.66
4. Peer & self-feedback	3.80 (.84)	4.10 (.78)	3.46 (.77)	85.18	<.01	.82
5. Task	4.59 (.79)	4.66 (.79)	4.51 (.79)	4.46	.04	.19
6. Process	4.42 (.89)	4.68 (.83)	4.14 (.88)	52.04	<.01	.63
7. Self-regulation	4.15 (.71)	4.32 (.69)	3.97 (.68)	34.57	<.01	.51
8. Timeliness	4.06 (.91)	4.38 (.80)	3.72 (.90)	77.58	<.01	.78
9. Encouragement	3.47 (.85)	3.55 (.83)	3.37 (.86)	5.51	.02	.21
Practices						
Non-teacher	.75 (.32)	.81 (.27)	.67 (.35)	28.36	<.01	.45
Teacher formative	.90 (.15)	.89 (.15)	.92 (.15)	4.45	.04	−.20
Protect & evaluate	.51 (.36)	.44 (.37)	.60 (.33)	28.23	<.01	−.45
Parents	.78 (.37)	.78 (.36)	.78 (.38)	.02	.90	.00

Note. Cohen's *d* values are positive when primary mean is >secondary mean. Level information missing for 4 teachers, $n = 514$.

RMSEA = .059 [90% CI = .051–.067]; SRMR = .058). Item and scale details are provided in Table 4.

The Non-teacher group included practices in which students and family members gave feedback in lieu of the teacher. The Teacher Formative group consisted of practices associated with in-class interaction between teachers and students where guidance was given as to how work could be improved. The Teacher Protective-Evaluation group embedded praise to the student within feedback that evaluated student work. The Parent Reporting group involved the teacher communicating with parents about the student.

As before, a comparison of the statistical features of this model was carried out between primary and secondary teachers. Inspection of changes in χ^2 and CFI showed that the model was equivalent for configuration of paths, regression weights from factors to items, and covariances between the factors. Therefore, comparison of mean scores between groups was permissible. A high proportion of teachers selected each item (i.e., mean selection rates varied from 51 to 90%) (Table 3). Mean score differences by sector were small to moderate with primary teachers selecting Non-Teacher feedback practices more often, while secondary teachers included Teacher Protective-Evaluation practices more frequently in their definitions.

4.3. Relationship of TCoF factors to perceived practices factors

The relationship between the Teacher Conceptions of Feedback nine-factor hierarchical model and the Teacher Feedback Practices four-factor hierarchical model were tested in a structural equation model. Paths were tested in accordance with the five hypotheses given earlier and statistically non-significant paths were trimmed from the model (Fig. 1). The resulting trimmed model had good fit to the data ($k = 56$; $\chi^2 = 2886.23$; $df = 1467$; $\chi^2/df = 1.97$, $p = .13$; CFI = .82; gamma hat = .91; RMSEA = .043 [90% CI = .041–.046]; SRMR = .061). Inspection of changes in χ^2 and CFI showed that the model was equivalent for configuration of paths, regression weights from factors to items, and covariances between the factors, indicating sufficient similarity between primary and secondary teachers to allow comparisons between groups.

Table 4
Feedback practices factors and items.

Practices of feedback used to define feedback	Loading
Non-teacher (FBP1) glb = .79	
1. Advice or comments that students give each other	.91
2. Suggestions or comments from other students	.87
3. Comments students give to themselves	.81
4. Comments parents give to their own child	.80
Teacher formative (FBP2) glb = .59	
5. Spoken comments	.57
6. Detailed written comments	.45
7. Instant responses to students' classroom work	.44
8. Hints, tips, and reminders written on student work	.36
9. Information on the quality of work relative to standards, norms, or expectations	.31
10. Discussions with students about their work	.26
Teacher protective evaluation (FBP3) glb = .82	
11. Grades, scores, or marks on student work	.84
12. Ticks or crosses on student work	.84
13. Stickers, stamps, or smiley faces on student work	.82
14. Praising students for how hard they have worked	.81
15. Giving correct answers when students answer incorrectly	.79
Parent reporting (FBP4) Cronbach α = .75[a]	
16. Parent–teacher conferences	.91
17. Reports to students' parents	.89

Note. glb = greatest lower bound.
[a] glb not estimable with two items.

4.3.1. Hypothesis 1

The Teacher Formative feedback factor was predicted by the Improvement factor (β = .26), resulting in a small additional variance explained ($\Delta R^2 = .04$, $f^2 = .04$). Contrary to expectations, all other anticipated predictors (i.e., task, process, self-regulation, and timeliness) were not statistically significant, perhaps due to their relatively high inter-correlations. Accordingly, because teachers expect students to use their feedback, they appear to think that feedback practices which they control (i.e., teacher-centric) are used for the explicit purpose of improving the quality of student learning outcomes.

4.3.2. Hypothesis 2

The Non-Teacher feedback factor was predicted by the Peer- and Self-Feedback (β = .31) and Process (β = .20) factors, resulting in a medium amount of additional variance explained ($\Delta R^2 = .21$, $f^2 = .27$). Hence, teachers' understanding of feedback as a set of practices in which they are not involved was predicated by an understanding that feedback requires student self- and peer-interaction, especially around how work should be carried out.

4.3.3. Hypothesis 3

The Teacher Protective-Evaluation feedback factor was predicted by the Encouragement conception of feedback (β = .21), resulting in a small additional variance explained ($\Delta R^2 = .11$, $f^2 = .12$). Consequently, the practice of giving praise while giving grades or scores appeared to be linked to the idea that feedback is meant to be encouraging and supportive of student emotional commitment and engagement in learning.

4.3.4. Hypothesis 4

The Parent Reporting factor was predicted by the Reporting and Compliance factor (β = .14), resulting in a trivial difference in variance explained ($\Delta R^2 = .01$, $f^2 = .01$). As might be expected, feedback practices around reporting to parents were related to endorsement that feedback requires reporting. However, the small amount of additional variance in the self-reported practice suggested that these practices might be driven by considerations independent of teacher conceptions of feedback (e.g., school policy requirements or expectations related to teacher professionalism).

4.3.5. Hypothesis 5

Unsurprisingly, none of the feedback practice factors was predicted by the Irrelevance factor. It is only natural that practices teachers associated with feedback should not be related to any sense that feedback is irrelevant. If feedback were irrelevant, then it should not be practised at all.

5. Discussion

5.1. Summary of results

小学、初中教师对反馈的态度基本相同

These data suggest that New Zealand teachers endorsed feedback factors associated with the policy framework of using assessment and feedback to improve learning. Furthermore, primary and secondary teachers provided largely equivalent responses to the questionnaire; with only small differences in factor mean scores. The structural paths from the conceptions of feedback factors to the feedback practice definitions, while generally weak, indicated that there were conceptually meaningful relations between teachers' conceptions of feedback and the practices they associate with it.

教师对反馈的观念与其在实践中的应用有一定的联系

The teachers' focus was on involving students in generating and using feedback to improve their work and develop autonomy, rather than in providing task-oriented information to students.

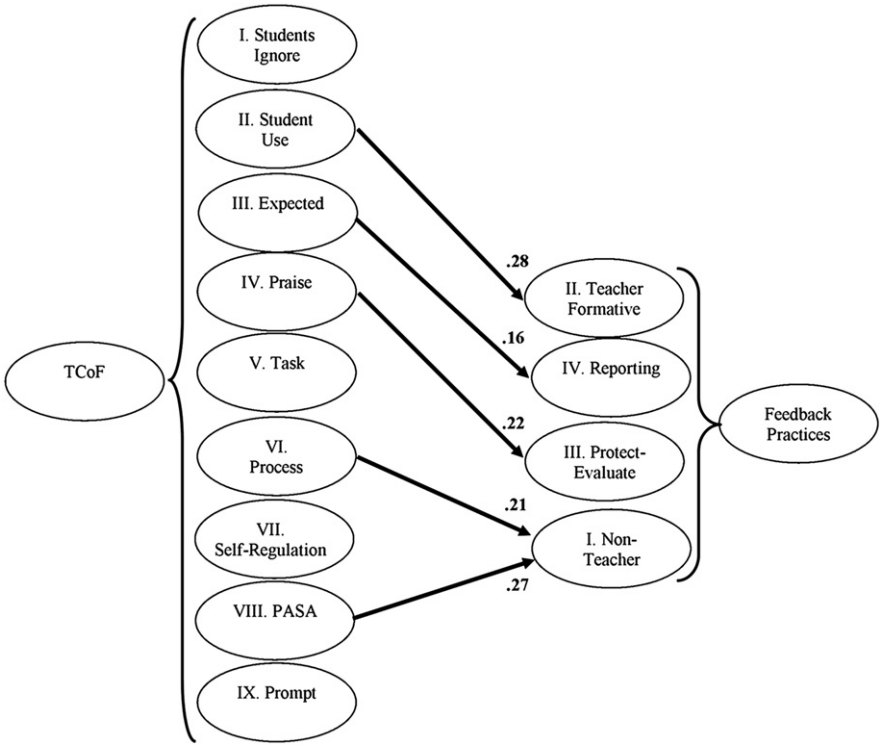

Fig. 1. Schematic representation of the relationship of TCoF factors to feedback practices factors. Note. Inter-factor correlations, items, and factor paths omitted for simplicity. Values are standardised beta regression weights.

Encouragement and protection of student self-esteem were considered as only minor aspects of this learning-oriented conception of feedback. Teacher responses in this sample indicated they did not believe in using feedback simply to enhance students' well-being (i.e., praise for effort or increased self-esteem); except in the case of evaluative reporting.

5.2. Implications for the TCoF inventory

The factors obtained in this study were consistent with the planned factors, though somewhat narrower in scope. The narrowing of these factors is unfortunate and partially attributable to the processes involved in creating pools of items with strong inter-correlations. The current factors do not reflect fully the nuances and complexity of feedback processes and constructs evident in the research literature. This study does not reveal whether teachers' understandings of feedback are as sophisticated and complex as those found in academic literature, nor is it clear whether any self-report survey could ascertain such 'fine-tuned' information. Future research, using focus groups or interviews, may identify language that teachers use to express complex views about feedback and these data would be useful in creating additional statements that tap into their conceptions. 教师可以识别反馈类型

Despite the simplification of some factors, the results of this study demonstrate that teachers can identify the distinctions proposed in Hattie and Timperley's (2007) feedback typology (i.e., Task, Process, Self-regulation, and Self). However, the inter-correlations do suggest that strong distinctions among process, task, and self-regulation feedback are probably not made in practice. The teachers clearly identified the irrelevance, improvement, and encouragement purposes identified by Irving et al. (2011). Furthermore, given teacher requirements to report student progress and results to parents, school leaders, and other external stakeholders, it is unsurprising that the feedback for compliance and reporting purposes was also identified. Additionally, there is some evidence of convergent and divergent validity for the TCoF factors because of the logical nature of the statistically significant paths between conceptions of feedback and practices perceived as feedback.

Future studies with this inventory would be much stronger with additional methods of determining feedback practices (e.g., classroom observations, inspection of written feedback to students, student reporting of feedback received) to ascertain the extent to which teachers' self-reported endorsement of the learning-orientation is reflected in their practice. In the interim, it would appear the TCoF inventory is capable of identifying a wide range of teacher beliefs about the nature of feedback, and continued use of the inventory as an adjunct to professional development or as a research tool appears warranted. The relatively weak regression weights from the second-order factor (Conceptions of Feedback) to Encouragement and Irrelevance factors suggest that these beliefs are influenced by other factors. Future studies could integrate examination of teacher feedback beliefs and practices in light of external factors, such as grade level, student performance, assessment characteristics, or teacher personality. Student ratings of teacher assessment behaviours have been shown to be good predictors of teacher practices (van der Schaaf, Stokking, & Verloop, 2008) and these data would be a useful adjunct to future research. Furthermore, it is highly likely that teacher conceptions of feedback are inter-twined with their beliefs concerning assessment, teaching, learning, and curriculum (Brown, 2008). For this reason, future studies would do well to examine explicitly the relationship of feedback conceptions with these other domains.

6. Conclusion

This study has implications internationally for teaching practice and the implementation of Assessment for Learning reforms. Although the results are situated in New Zealand, an educational

context that prioritises Assessment for Learning policy and practice, feedback remains an essential pedagogical practice across all teaching and learning situations even in contexts where Assessment for Learning is not the dominant paradigm. Hence, the TCoF inventory has applicability in all teaching and teacher education contexts because providing feedback to students is an integral part of teaching.

对教学实践的影响

6.1. Implications for teaching practice

It is reassuring that the teachers in this study endorsed learning-oriented conceptions of feedback rather than focussing on student well-being, despite the conventional perception that a pastoral care orientation dominates teacher responses to students (Pajares & Graham, 1998). This growth-orientation has been associated with improved learning outcomes and greater self-regulation among learners (Boekaerts & Corno, 2005). Notwithstanding possible discrepancies between teachers' espoused views and their actual practices, it would appear that the rhetoric and logic of Assessment for Learning is well established in the conceptions of practicing teachers in New Zealand.

We anticipate, however, that this perspective is distributed reasonably robustly in teacher populations around the world as a consequence of pre- and in-service teacher professional development. Almost by definition, teachers are interested in improving the learning of children and adolescents and it is expected that teachers would endorse a learning-orientation for feedback. While it may seem that Assessment for Learning reforms are best implemented in societies where teachers' *pre-existing* attitudes and understandings about feedback are aligned with the emphases of the reform movement, there is evidence that teachers can be taught both the skills and rationales underpinning effective feedback (Brookhart, Moss, & Long, 2010; Cooper & Cowie, 2010; Dixon & Haigh, 2009).

支持理论

Nonetheless, these results represent *espoused* theories (Argyris & Schon, 1974) that the participants have used to explain their actions, rather than their actual practices. Previous studies (Eraut, 2000; Hammerness, Darling-Hammond, & Bransford, 2005; Harnett, 2007; Turner-Bisset, 1999) have concluded that there is frequently a discrepancy between what teachers believe or claim they are doing and what they are actually doing. Consequently, although the teachers in the present study *professed* a preference for learning-oriented feedback it remains to be seen whether this is how feedback is conducted in real classrooms, as studies have identified gaps between the intended and actual quality of feedback to students (e.g., Parr & Limbrick, 2010).

While use of the TCoF inventory will not directly improve teachers' actual feedback practices, the data gained from this instrument could help professional development providers better tailor programs to align with teacher–participant thinking. It is possible that professional development interventions around formative feedback work only with and for teachers who already believe in learning-oriented feedback. Thus, still to be examined are factors that contribute to *changing* teachers' beliefs concerning feedback from 'reporting grades' to 'supporting learning.' The TCoF inventory may provide a useful tool for establishing baseline beliefs about feedback as a means of evaluating whether teachers participating in professional development are actually modifying their beliefs or simply building on pre-existing conceptions that are already aligned with the proposed reforms.

Additionally, pre-service teacher educators should not take for granted that students already believe in and practice the types of feedback that the practicing teachers surveyed here endorse. Accordingly, use of the current inventory in a pre-service context may help identify student teachers' existing beliefs, and ascertain their need for professional learning about effective feedback. Regardless, pre-service teachers would doubtless benefit from instruction in strategies and techniques for giving growth-oriented feedback to students as well as how to provide 'accountability-type' reporting to administrators or parents.

In a more speculative vein, it is worth considering how school students might learn to implement student-based feedback practices as advocated by the AfL reform. It seems likely that students will not learn to give themselves or their peers effective feedback without experiencing growth-oriented feedback from their own teachers. Indeed, if teachers use feedback to protect students from psychological harm in the context of evaluation, this may obscure from students, especially low performing ones, their real learning needs. This suggests that new teachers themselves need to experience growth-oriented feedback, including information about weaknesses and steps to improve, from their own pre-service educators. Therefore, it seems to be incumbent on pre-service teacher educators to equip new teachers with theory about and experience of high-quality feedback. Otherwise, it is possible that future teachers will resort to giving vague, unhelpful feedback such as "doing well", "could be doing better", or "not working hard enough".

学习改革评价的影响

6.2. Implications for assessment for learning reforms

Inspired by a British response to high-stakes evaluation assessment practices (e.g., Assessment Reform Group, 2002; Black & Wiliam, 1998), the Assessment for Learning policy reforms (e.g., Leahy et al., 2005) have travelled globally (Berry, 2011; Black & Wiliam, 2005). The reform seeks to shift assessment understandings and practices from primarily a summative and evaluative focus to a formative and pedagogical one. Feedback within the Assessment for Learning reform movement is meant to be very much focused on in-the-moment interactions in which teachers provide information that facilitates and prompts improvement in student learning. It also seeks to involve students in giving themselves and each other feedback that supports learning. The results of this survey show that the priorities in teachers' conceptions of feedback are consistent with the official New Zealand government policy concerning the use of assessment and feedback. That policy requires teachers to use assessment interactively to adjust teaching, engage students in self-regulating their own learning, and provide formative feedback that informs next steps in teaching and learning. Certainly, based on this survey, we would not expect professional development in Assessment for Learning to effect major changes in teacher beliefs about feedback since, at least in New Zealand, it would appear teachers already endorse these values.

The more innovative use of non-teacher feedback sources appears to depend in part on individual teacher beliefs, rather than automatically occurring in response to an Assessment for Learning policy. If teachers are not persuaded that there is educational value in students generating feedback, then they are not likely to acknowledge and implement this practice. In contrast, it is worth considering that teacher conceptions of feedback did not explain additional variance in perceived feedback practices around reporting to parents and giving formative feedback, which are part of expected practice in New Zealand. Consequently, the results may not reflect teachers' personal beliefs since such practices are expected from all teachers.

These New Zealand results raise questions as to whether teachers in other societies, especially where consequences for assessment are much higher, would have similar conceptions of feedback. While teaching universally seems to involve instruction, monitoring, and giving feedback to students about how they could

do their work better, grading practices and the use of student-led feedback are much less uniform. For example, reporting with test scores and/or grades appears to be much more prevalent in societies that privilege examinations (e.g., Hong Kong) or make use of high-stakes accountability systems to press for learning improvements (e.g., *No Child Left Behind* in the USA). There may even be different feedback practices in societies where there is a strong emphasis on not questioning the authority of teachers.

Assessment policies are also probable factors in shaping teacher conceptions of feedback. For example, in low-stakes assessment environments, teachers' beliefs about assessment for improvement were weakly correlated with assessment for school accountability ($r = .43$ in Queensland; $r = .47$ in New Zealand) (Brown, 2011; Brown, Lake, & Matters, 2011). In contrast, in environments where assessment has much higher consequences for students and schools, these beliefs had quite different associations. Research with teachers in Hong Kong and China found a strong correlation between assessment for improvement and assessment for accountability ($r = .80$) (Brown, Hui, Yu, & Kennedy, 2011). Likewise, in the United States, perhaps because of high-stakes consequences for assessment results (e.g., *The Elementary and Secondary Education Act* 1965; *No Child Left Behind* 2001), teacher responses to assessment have been predominantly negative (Hamilton, 2003), and especially so in contexts where large populations of 'at-risk' students are present (Deneen & Brown, 2011; O'Quin, 2009). It seems, then, in environments in which strong external accountability pressures exist, that assessment for improvement is strongly entangled with demonstrating quality on the external measures. This suggests that assessment policy frameworks have an impact on teachers' conceptions of the role of feedback. In high-stakes testing environments, feedback may be more about maximising performance on accountability testing than it is about learning improvement.

Consequently, we suspect that while endorsement of feedback as learning-oriented will be relatively important in all societies, the level of endorsement of some factors in the TCoF (e.g., Praise, Expected, Peer and Self-Assessment) will be consistent with the dominant priorities of the society in which teachers are employed (i.e., ecologically rational) (Rieskamp & Reimer, 2007). Therefore, we expect that the TCoF factors will be identified in most jurisdictions, though it is likely factor mean scores and intercorrelations will be replicated only in jurisdictions with similar low-stakes, formative assessment frameworks. If policy reform or teacher education seek to bring about a greater willingness and capability of teachers to provide formative, learning-oriented feedback, it may be necessary to make changes to the general policy context to reduce negative consequences for schools through assessment for accountability. Further, it will no doubt require strong educational leadership to make sure that accountability pressures do not thwart teacher efforts to provide learning-oriented, formative feedback.

References

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

Andrade, H. (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.

Arbuckle, J. L. (2008). *AMOS (version 17.0.0) [computer program]*. Crawfordville, FL: Amos Development Corporation.

Argyris, C., & Schon, D. A. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.

Askew, S., & Lodge, C. (2000). Gifts, ping-pong and loops – linking feedback and learning. In S. Askew (Ed.), *Feedback for learning* (pp. 1–17). London: Routledge.

Assessment Reform Group. (2002). *Testing, motivation and learning*. Cambridge, UK: Faculty of Education, University of Cambridge.

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.

Berry, R. (2011). Assessment reforms around the world. In R. Berry, & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (pp. 89–102). Dordrecht, NL: Springer.

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

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

Black, P., & Wiliam, D. (2005). Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices. *The Curriculum Journal, 16*(2), 249–261.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: a perspective on assessment and intervention. *Applied Psychology: An International Review, 54*(2), 199–231.

Brookhart, S. M., Moss, C. M., & Long, B. A. (2010). Teacher inquiry into formative assessment practices in remedial reading classrooms. *Assessment in Education: Principles, Policy & Practice, 17*(1), 41–58. doi:10.1080/09695940903565545.

Brophy, J. (1981). Teacher praise: a functional analysis. *Review of Educational Research, 51*(1), 5–32.

Brown, G. T. L. (2004). Measuring attitude with positively packed self-report ratings: comparison of agreement and frequency scales. *Psychological Reports, 94*(3), 1015–1024. doi:10.2466/pr0.94.3.1015-1024.

Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York: Nova Science Publishers.

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. Student self-assessment. In J. H. McMillan (Ed.), *The SAGE handbook of research on classroom assessment*. Thousand Oaks, CA: Sage, in press.

Brown, G. T. L., Hui, S. K. F., Yu, W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: a tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research, 50*(5–6), 307–320. doi:10.1016/j.ijer.2011.10.003.

Brown, G. T. L., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: the impact of policy priorities on teacher attitudes. *Teaching and Teacher Education, 27*(1), 210–220. doi:10.1016/j.tate.2010.08.003.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). Washington, DC: APA.

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

Byrne, B. M. (1989). Multigroup comparisons and the assumption of equivalent construct validity across groups: methodological and substantive issues. *Multivariate Behavioral Research, 24*(4), 503–523.

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

Clarke, S. (2003). *Enriching feedback in the primary classroom: Oral and written feedback from teachers and children*. London: Hodder & Stoughton.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.

Cooper, B., & Cowie, B. (2010). Collaborative research for assessment for learning. *Teaching & Teacher Education, 26*, 979–986. doi:10.1016/j.tate.2009.10.040.

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

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, NC: Information Age Publishing.

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education: Principles, Policy and Practice, 6*(1), 101–116.

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

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B, 39*(1), 1–38.

Deneen, C. C., & Brown, G. T. L. (2011, October). *The persistence of vision: An analysis of continuity and change in conceptions of assessment within a teacher education program*. Paper presented to the 37th annual meeting of the International Association for Educational Assessment, Manila, Philippines.

Dixon, H., & Haigh, M. (2009). Changing mathematics teachers' conceptions of assessment and feedback. *Teacher Development: An International Journal of Teachers' Professional Development, 13*(2), 173–186. doi:10.1080/13664530903044002.

Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology, 70*(1), 113–136.

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

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.

Hamilton, L. (2003). Assessment as a policy tool. *Review of Research in Education, 27*(1), 25–68.
Hammerness, K., Darling-Hammond, L., & Bransford, J. (2005). How teachers learn and develop (with Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K.). In L. Darling-Hammond, & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 358–389). San Francisco: Jossey-Bass.
Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education, 35*(2), 213–224.
Harnett, J. A. (2007). *Changing learning conversations: An action research model of reflective professional development*. Unpublished doctoral dissertation, Massey University, Palmerston North.
Harris, L. R., & Brown, G. T. (2008). *Teachers' conceptions of feedback (TCoF) inventory*. Unpublished test, University of Auckland, Measuring Teachers' Assessment Practices (MTAP) Project, Auckland, NZ.
Harris, L. R., & Brown, G. T. L. (2010, May). *"My teacher's judgement matters more than mine": Comparing teacher and student perspectives on self-assessment practices in the classroom*. Paper presented to the SIG-classroom assessment at the American Educational Research Association annual conference, Denver, CO.
Harris, L. R., Harnett, J., & Brown, G. T. L. (2009). "Drawing" out student conceptions of assessment: using pupils' pictures 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.
Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.
Heuvelmans, T. (2010). *TiaPlus (version build 309) [computer software]*. Arnhem, NL: CITO.
Hill, M. (2000). Dot, slash, cross: how assessment can drive teachers to ticking instead of teaching. *SET: Research Information for Teachers, 1*, 21–25.
Irving, S., Harris, L., & Peterson, E. (2011). 'One assessment doesn't serve all the purposes' or does it? New Zealand teachers describe assessment and feedback. *Asia Pacific Education Review, 12*(3), 413–426. doi:10.1007/s12564-011-9145-1.
Irving, S. E., Peterson, E. R., & Brown, G. T. L. (2007, August). *Student conceptions of feedback: A study of New Zealand secondary students*. Paper presented to the biennial conference of the European Association for Research in Learning and Instruction (EARLI), Budapest, Hungary.
Irving, S. E., Peterson, E. R., & Brown, G. T. L (2008, July). *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.
Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment: minute by minute, day by day. *Educational Leadership, 63*(3), 18–24.
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.
New Zealand Ministry of Education. (2005). Report on the findings of the 2004 teacher census. Retrieved 10.12.10, from. http://www.educationcounts.govt.nz/publications/schooling/teacher_census.
New Zealand Ministry of Education. (2010). Ministry of education position paper: assessment [schooling sector]. Retrieved 20.11.10, from. http://www.minedu.govt.nz/theMinistry/PublicationsAndResources/AssessmentPositionPaper.aspx.
O'Quin, C. R. (2009). *Feedback for students: What do teachers believe?* Unpublished doctoral dissertation, the Consortium of Southeastern Louisiana University and University of Louisiana Lafayette, Hammond, LA.
Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Review of Educational Research, 62*, 307–332.
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.
Parr, J. M., & Limbrick, L. (2010). Contextualising practice: hallmarks of effective teachers of writing. *Teaching and Teacher Education, 26*(3), 583–590. doi:10.1016/j.tate.2009.09.004.
Peterson, E. R., & Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*(3), 238–250.
Rieskamp, J., & Reimer, T. (2007). Ecological rationality. In R. F. Baumeister, & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 273–275). Thousand Oaks, CA: Sage.
Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation, 11*(10), Available online. http://pareonline.net/getvn.asp?v=11&n=10.
Rubie-Davies, C. M., Flint, A., & McDonald, L. G. (2011). Teacher beliefs, teacher characteristics, and school contextual factors: what are the relationships? *British Journal of Educational Psychology*, . doi:10.1111/j.2044-8279.2011.02025.x.
Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education, 35*(5), 535–550.
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*(1), 77–84.
Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189.
Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*(1), 107–120. doi:10.1007/S11336-008-9101-0.
Strijbos, J., & Sluijsmans, D. (2010). Unravelling peer assessment: methodological, functional, and conceptual developments. *Learning and Instruction, 20*(4), 265–269.
Thompson, A. G. (1992). Teachers' beliefs and conceptions: a synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York: Macmillan.
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.
Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Buckingham, UK: Open University Press.
Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: a typology. *British Educational Research Journal, 22*(4), 389–404.
Turner-Bisset, R. (1999). The knowledge bases of the expert teacher. *British Educational Research Journal, 25*(1), 39–55.
Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(4), 4–70.
van der Schaaf, M. F., Stokking, K. M., & Verloop, N. (2008). Teacher beliefs and teacher behaviour in portfolio assessment. *Teaching and Teacher Education, 24*, 1691–1704. doi:10.1016/j.tate.2008.02.021.
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.
Vollmeyer, R., & Rheinberg, F. (2005). A surprising effect of feedback on learning. *Learning and Instruction, 15*(6), 589–602.
Wiliam, D. (1999). Formative assessment in mathematics: part 2: feedback. *Equals, 5*(3), 8–11.
Woolfolk Hoy, A., Davis, H., & Pape, S. J. (2006). Teacher knowledge and beliefs. In P. A. Alexander, & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed.). (pp. 715–737) Mahwah, NJ: LEA.