A GEARed peer feedback model and implications for learning facilitation

Shu-Chen Huang

Routledge
Taylor & Francis Group

A GEARed peer feedback model and implications for learning facilitation

Shu-Chen Huang

Foreign Language Center, National Chengchi University, Taipei, Taiwan

ABSTRACT

This study proposes a model to view and analyse college students' written peer feedback after they assess each other's formative task assignment and provide comments. The model was developed by: (a) revisiting learning-oriented assessment and feedback theories, (b) reviewing and summarising existing peer feedback analysis approaches as detailed in published empirical studies, and (c) analysing a small set of primary learner peer feedback data. In the proposed GEARed model of peer feedback, each complete piece of comment on one improvable unit may have four cognitive elements, i.e. identification of the **G**ap between performance and goal, **E**xplanation of the gap, gap-bridging **A**spirations, and being **R**esourceful to reach beyond the current knowledge/skill level. Two social-affective-interactive elements also play a part, i.e. being **e**ncouraging to the peer feedback receiver and **d**oubtful toward the provider self. Each feedback comment may contain any or all of these incremental elements rather than being assigned only to one category. It is argued that the absent or unresolved aspects in feedback leave room for the receiver to reflect and may suggest instructional actions for teachers. Finally, different patterns of GEARed component combinations are hypothesised and their implications for learning and teaching are discussed.

Introduction

Qualitative peer feedback, a follow-up activity of task assignment and a natural by-product of peer assessment, has become a popular pedagogical practice and a focus of study in higher education. Its complexity in scope, coverage, administration and data analysis poses challenges for researchers and practitioners. One such challenge is similar to classroom assessment being likened to a 'hodge-podge' (Brookhart 2004), pointing to the lack of common terminologies and agreed-upon data analysis approaches to help aggregate study efforts and accumulate findings in the field. To address this problem, a few meta-analyses published in the past two decades have focused specifically on peer assessment in classroom learning contexts, and tried to lay down and refine the scope and processes of peer assessment (Gielen, Dochy, and Onghena 2011; Topping 1998; Van Zundert, Sluijsmans, and van Merriënboer 2010). The scope of these reviews, however, is at the more macro level, trying to take stock of a myriad of factors involved in peer assessment, while many individual studies focus on just a few of them.

Often researchers resort to the analysis of peer feedback data to answer their questions about peer assessment, and such analysis is usually laborious and troublesome. Lacking at the more micro level

in our toolbox of peer assessment research is an approach to analyse peer feedback data that can be used across educational disciplines and contexts, that can help shed light on the data being analysed, and that can aggregate and advance our understanding of peer assessment and feedback. Under the same objective of consolidating research efforts in peer assessment, and more specifically and narrowly in peer feedback, the current study zeroed in on learners' written peer feedback and attempted to build a framework for investigating and understanding its nature, components and significance by: (a) revisiting relevant theories, (b) examining methodologies of peer feedback analysis in existing studies, and (c) analysing a small set of empirical learner peer feedback data. The questions that guided this endeavour were as follows.

(1) Based on assessment-oriented learning and assessment feedback theories, how does peer feedback contribute to learning under formative assessment situations?
(2) By referring to published empirical studies, what are the existing approaches to peer feedback data analysis? What insights could they offer? What problems should be resolved?
(3) With a peer feedback analysis framework emerging from the above investigations, how could a small set of primary peer feedback data be interpreted and help shape a peer feedback model?

The model is proposed in the hope that it will help us better understand peer feedback and learning. It will be presented and explained with real peer feedback examples from the primary data-set, followed by discussions of the significance and implications drawn from analysis of the data using this model.

Theories of formative assessment and feedback

Black and Wiliam's (2009) two-dimensional theory of formative assessment, or assessment for the purpose of learning, features one dimension of cognitive activity in which: (1) criteria of learning are defined, (2) the gap between the criteria and actual performance is identified, and (3) the gap between the two is bridged so as to complete learning. The other dimension involves participants in a classroom all being mobilised as assessors. Ideally, the teacher first models the assessment steps. Under the ideal of a socio-constructivist paradigm, learners are then commissioned to practice assessment with peers and learn to assess and assess to learn. Eventually, individual learners acquire and internalise the assessment ability and become independent learners in the relevant subject area.

In the centre of formative assessment lies the critical information of feedback which is supposed to guide learners in the gap-filling journey. Hattie and Timperley's (2007) meta-analysis based feedback model summarises characteristics of effective feedback. Aiming to reduce discrepancies between actual and desired performance, good feedback tells learners: (a) 'where they are going' by feeding up to the criteria, (b) 'how they are doing' by feeding back to the performance, and (c) 'where to next' by feeding forward to the next step of learning. In this model, teacher and student roles are clearly distinguished, featuring the teacher as the feedback provider and students as receivers. However, as we know more about formative assessment and feedback, learners are no longer expected to just passively receive feedback. Nicol and Macfarlane-Dick's (2006) widely-cited seven principles of good feedback project an ideal image of self-regulatory learners who proactively participate in assessment and respond to feedback. Although many peer feedback activities appear to be directed for the receiver, recent studies seem to unequivocally suggest that the giver benefits much more (Nicol, Thomson, and Breslin 2014; Reinholz 2016).

Peer assessment and feedback have been used as an important way to improve learning, especially at the tertiary level when learners are cognitively and affectively more mature. Topping (1998) validates the efficacy of peer assessment through a meta-analysis and sets out a peer assessment typology to consolidate research efforts and synthesise findings. His effort was later updated by Gielen, Dochy, and Onghena (2011), who refine Topping's typology into a five-cluster peer assessment 'inventory' to capture the diversity and cover the particularities of peer assessment design. Van Zundert, Sluijsmans, and van Merriënboer (2010), also by way of literature review, focus on peer assessment processes and

ascertain relations among variables that foster effective peer assessment. They conclude that training can enhance peer assessment skills and such skills are related to students' thinking style and academic achievement.

In a peer assessment model, Sluijsmans et al. (2004) define peer assessment skills in three sequential aspects: (1) defining assessment criteria, (2) judging the performance of a peer, and (3) providing feedback for future learning, which is almost identical with Black and Wiliam's (2009) theory of formative assessment. Particularly for the third aspect, Sluijsmans et al.'s (2004) model further breaks down peer feedback provision into three parts: (1) formulate discrepancies in a peer-assessment report, (2) formulate points for improvement, and (3) reflect on points of improvement to the peer.

This review suggests that the fundamentals of peer assessment and feedback still lie in the gap-bridging intention that is supported by knowledge of learning criteria and assessment of actual performance, both of which entail lots of training in the subject matter. It has to be noted that learners, by engaging in peer assessment and feedback give-and-take, are not to become experts like their teachers, but to engage in the task so as to actively learn.

Existing approaches to peer feedback analysis

To understand existing approaches to analysis of peer feedback data, I focused on empirical studies at tertiary level which reported details on how written peer feedback was analysed. First, Van Zundert, Sluijsmans, and van Merriënboer (2010), a literature review for the period between 1990 and 2007 focusing on peer assessment processes and effectiveness, was taken as a starting point. By screening major databases for relevant publication, Van Zundert et al. found 26 articles and divided them into four outcome variable types: psychometric qualities, domain-specific skill, peer assessment skill, and student attitudes toward peer assessment. Among them, peer assessment skill is the category directly related to the current investigation and, within this category, three studies detailing approaches to peer feedback analysis were selected.

To update and expand this sample, I followed Van Zundert et al.'s procedures and did another round of screening in the Web of Science by selecting the following four databases: Social Science Citation Index (SSCI), SCI-Expanded, A&HCI, and ESCI. Selection criteria included: (a) document type: journal article, (b) language: English, (c) publication year: 2008–2017, and (d) domain: educational research. Some high-frequency but irrelevant results, such as those studies in the field of medicine, related to journal peer review, case study, at primary and secondary levels, or with electronic platforms such as MOOCs as the focus were excluded. Keywords applied included peer feedback, peer comment, peer assessment, peer review, peer evaluation and peer rating. Titles, abstracts and then the content of the initial 597 articles were read to find empirical studies at tertiary level that reported with details on how written peer feedback was analysed.

Fifteen articles were retained and are summarised in Table 1 and marked with an asterisk in the reference list. For each study, the country, discipline and learner task were highlighted, followed by number of participants and data sets in the units the author(s) specified. The categories of peer feedback were listed in the middle, with relations among categories marked by brackets of up to three levels. The frames and approaches of analysis were noted in the righthand column.

Frames of reference in viewing peer feedback data

Across the 15 studies, peer feedback was viewed in either a task criteria frame, an assessment function frame or a combination of both. Earlier studies adopted a task criteria frame as evidenced in the categories which peer feedback was assigned to. For example, Lin, Liu, and Yuan (2001) checked learner's peer feedback to see if it addressed the task criteria of relevance, thoroughness and sufficiency of references. Such a typology, taken directly from the rubrics or requirements for learners' assessed work, could be mapped to the object of peer feedback and was straightforward. However, task criteria were not only discipline specific; they are mostly context and situation specific. With such a frame, it is

Table 1. Summary of Peer Feedback Typologies.

Study	Data-set	Peer feedback typology	Analysis approach
Lin, Liu, and Yuan (2001) Taiwan Computer Science Summary writing	58 Computer Science majors' peer feedback	6 criteria: relevance, thoroughness, sufficiency of references, theoretical clarity, discussion clarity, significance of conclusion. (in contrast to holistic general comments in control group)	Frame: criteria; Analysis: holistic scalar rating (1–10), not weighted
Sluijsmans et al. (2004) The Netherlands Discovery Learning Lesson plan	93 sophomores' peer feedback	8 variables: use of criteria, naïve word use, consequent structure, being critical, giving a conclusion, posing questions, giving a mark, giving points for improvement	Frame: criteria and assessment; Analysis: categorical (1/0 × 4) and scalar (1-4 × 7) rating, partially weighted
Davies (2006) United Kingdom Computing Essays	67 students; 528 marking; 4298 comments	10 commenting categories: readability, aimed at correct level, personal conclusion, referencing, research and use of the Web, content and explanations, examples and case studies, overall report quality, introduction and definitions, report presentation and structure (each determined as over or under comment)	Frame: criteria; Analysis: categorical; not weighted
Nelson and Schunn (2009) United States History Essays	24 papers × 6 reviews → 1073 idea units (each from a few words to several sentences)	3 types of feedback: {summary, praise, problem/solution [global, local; affective language (mitigation-compliment, mitigation-other, neutral); localization (+/–); type (problem, solution, both); explanation of the problem (+/–); explanation of the solution (+/–)]}	Frame: assessment; Analysis: hierarchical and categorical; not weighted
Cho and Cho (2011) United States Physics Lab reports	72 students, 3889 idea units (a self-contained message on a single issue of peer writing)	6 categories under 2 dimensions: **evaluation** (strength vs. weakness), **scope** (surface – writing mechanics, micro-meaning – within a paragraph, macro-meaning – across paragraphs) Note: For micro and macro, further assign content into: focus, development, validity, organisation, other; but no details provided	Frame: assessment; Analysis: assign a comment into one of six categories (2 dimensions × 3 scopes); not weighted
Patchan et al. (2013) United States Cognitive Science Written papers	38 students; 69 drafts; 165 reviews × 2 dimensions = 330 pieces of feedback → 1138 comments	3 types of feedback: {summary, praise, criticism [problem, solution, or both (2 focuses on content for each problem/solution: high prose, low prose, substance)]}	Frame: assessment; Analysis: hierarchical and categorical; not weighted
Chang (2015) Taiwan EFL Writing Take-home essays	27 sophomores, 22 essays and reviews; Unit of analysis: clause	**Audience awareness:** 2 feedback tone (indirect/peer-to-peer, direct/authoritative), 3 feedback focus (global, local, verb tense), 3 feedback stance [critique – problem only, revision-oriented suggestions (problem identification, explanation, solution); **Affectivity:** 8 social-affective functions (salutation, closing remarks, addressing by first name, addressing by using second person pronoun, compliments, encouragements, reviewer's name), 2 perspective (reader's, reviewer's)	Frame: assessment; Analysis: hierarchical and categorical; not weighted
Cheng, Liang, and Tsai (2015) Taiwan Biology Written reports	47 students × 5 peers × 3 rounds = 705 peer feedback messages	Affective (supporting, opposing), cognitive (direct correction, personal opinion, guidance), meta-cognitive (evaluating, reflecting), irrelevant Note: Each instance of peer feedback (the whole feedback message) could only be classified into one category according to the main idea of the message. (p. 80)	Frame: assessment; Analysis: categorical, mutually exclusive

(Continued)

Table 1. (*Continued*).

Study	Data-set	Peer feedback typology	Analysis approach
Patchan and Schunn (2015) United States Psychology 3-page papers	186 students × 4 reviews → 8288 comments	3 types of feedback; praise (+/−), problem (+/−), solution (+/−); then for problem and solution, check for localization (+/−) and focus (low prose, high prose, substance) Note: Each comment checked against the above types not as mutually exclusive, but as independent features	Frame: assessment; Analysis: hierarchical, NOT mutually exclusive; not weighted
Walker (2015) United Kingdom Technology Module Written Wiki reports	25 drafts and polished work; 48 reviews → 592 comments	6 **criteria** of the evaluation framework (meet brief, factual accuracy, structure, style, tech level, English), 5 assessment **categories** (content, skills development, motivating, de-motivating, other learning), 3 **depths** (indicate, correct/amplify, explain), 4 **tones** (suggesting, stating, exhorting, probing)	Frame: criteria and assessment; Analysis: each comment checked 4 times against 4 dimensions of 3–6 mutually exclusive levels
Allen and Katayama (2016) Japan EFL Writing Essays	6 students (in 3 dyads)	Suggestions: addition, deletion, substitution; Type: content, organisation, language, format, and unclassifiable (omitted from analysis)	Frame: criteria and assessment; Analysis: categorical and mutually exclusive
Ion, Barrera-Corominas, and Tomas-Folch (2016) Spain Pedagogy and Social Education Group project work	160 students → 637 feedback units	7 **task-development aspects**: on objectives, learning outcomes, transfer to other learning, balanced, enhance reflection, stimulate engagement, referring to competencies; 4 **formal aspects**: linguistic correctness, structural aspects, use of norms, ability to produce changes; 2 **motivational aspects**: self-esteem, assertiveness (mutually exclusive), Likert 1–4 less-a lot	Frame: assessment; Analysis: 3 aspects with 7, 4, 2 levels; scalar (1–4 Likert scale)
Noroozi, Biemans, and Mulder (2016) The Netherlands Biotechnology Argumentative essays	189 students' peer feedback	7 assignment criteria for peers to comment on: Intuitive opinion, claims in favour of the topic, scientific facts in favour of the topics, claims against the topic, scientific facts against the topics, integration of pros and cons, conclusions; 1 added dimension for peer review: suggestions for improving writing	Frame: assessment and criteria; Analysis: each review holistically checked against 8 standards of 3 levels (0, 1, 2)
Novakovich (2016) Canada Essay writing Essays	2 groups × 21 students × 3 essays; a minimum of 20 comments per draft required	4 levels: Naïve comments – too general to be helpful, usually encouraging comments, 1 point; Editing comments – lower-order subordinate skills, 2 points; Critical comments – evaluate the work and explain its effect on the reader; 3 points, and Directive comments – suggest ways to improve work, 4 points	Frame: assessment; Analysis: each comment assigned to one of four levels with ascending weights
Pham and Usaha (2016) Vietnam EFL writing Essays	32 sophomore English majors × drafts 1–3 in 2 feedback rounds → 897 comments	2 areas: global (content, organisation), local (words, usage, grammar, spelling, punctuation); 2 nature: revision and non-revision oriented; 3 type: clarification-unity, clarification-specific of idea, explanation; Others as statements	Frame: criteria and assessment; Analysis: categorical

difficult to establish a common language for peer feedback researchers across educational contexts. More importantly, to understand peer feedback, we need to know more than whether learners have attended to certain task criteria in their peer feedback.

Except for a few earlier instances, as can be seen in Table 1, most studies adopted either an assessment function frame or a combination of assessment and criteria frame. Nelson and Schunn's (2009) typology is an example of assessment frame. It looks at whether peer feedback contains praise, identifies problem, provides solution, offers explanation and so on. Such a frame helps us examine different aspects of peer feedback and has higher potential for improving the teaching and research of peer feedback. An example of a combined frame is Sluijsmans et al. (2004). In addition to examining whether peer feedback attended to the use of criteria in peers' work, Sluijsmans et al. examined whether peer feedback was critical, posed questions, gave points for improvement, and so on.

Core peer feedback categories: cognitive and affective

Taking stock of the peer feedback categories seemed a daunting task at first. The numbers of categories from each study as shown in Table 1 range from a handful to about 20, with lots of subtle variations from one study to another. With such complexity, theories of formative assessment and general intentions of implementing peer feedback could provide a useful framework for discussion.

First, there has to be a set of underlying standards for the work to be assessed. These standards, or rubrics, or criteria, or expectations, could be explicitly stated as in typologies with a criteria frame or in a combined frame, or implicitly held by learners as a prerequisite, as in many studies using an assessment frame. With such standards established, although sometimes not articulated, the peer feedback providers were then able to check the work being assessed against these standards and determine if there is a discrepancy between the standards and actual performance. If not, learning is successful and there was no peer feedback or there may be a praise. If a discrepancy is observed, this peer feedback may, in the form of opinions or suggestions, probe with questions, resolve it by pointing out the gap, explain why it is a gap, bridge it, and further explain why the gap is bridged this way. Or, any or all of the above elements could be missing in the peer feedback. As a result, categories such as praise, questioning, evaluation for strength and weakness, criticism (with or without explanation) and solution (with or without explanation) are commonplace and form the core of peer feedback categories.

These types of peer feedback delineate the cognitive processes involved when learners assess peers' work and provide comments in the hope to help revise it for the better. Beyond cognition, there are categories having to do with interpersonal relationship and interaction. These peer feedback types are labelled as motivating, demotivating, affective (supporting, opposing), affective language (mitigation-complement, mitigation-other, neutral), motivational (self-esteem, assertiveness) or various tones (indirect/peer-to-peer, direct/authoritative, suggesting, stating, exhorting, probing). Praise, to signal nonexistence of gap between performance and expectations, could also be assigned to this category. Such affective aspects could also be found in teacher feedback, but may be especially significant when learners are learning through the not-so-familiar and inherently critical assessor stance and, moreover, will be at the receiving end after they have formulated and delivered peer feedback to others.

Other peer feedback categories: context-specific and teacher-role related

Some peer feedback types, such as those related to the scope or structure of an assignment, even with an assessment frame, can be context-bound and are not necessarily transferrable to other tasks in other situations. For example, many typologies expect that when learners furnish peer feedback, they attend to global and local aspects of writing (e.g. Cho and Cho 2011; Pham and Usaha 2016), which may not apply to non-writing tasks. Similarly, Noroozi, Biemans, and Mulder's (2016) 'suggestions for improving writing' (26) has to be based on a writing task. As legitimate as these peer feedback types are in their own teaching/learning contexts, they could be inapplicable for, say, a mathematics task.

Naturally, some peer feedback typologies look up to assessment feedback guidelines written for teachers, but the embedded expectation for peer feedback quality may seem too high to be realistic for students. For example, Ion, Barrera-Corominas, and Tomas-Folch (2016), following Nicol's (2011) design for written feedback, include in their peer feedback typology task-development aspect qualifiers such as 'transfer to other learning', 'balanced', 'enhance reflection', 'stimulate engagement' and 'referring to competencies' (4). These features may not be justifiable as part of a universal peer feedback typology.

Peer feedback units and coding decisions

Though not specified explicitly, units of peer feedback in Lin, Liu, and Yuan (2001) are the entirety of peer feedback by individual students. Each learner's peer feedback was holistically checked against six criteria and rated. Yet units of analysis in most other studies were much smaller. Except for Chang (2015), whose unit was a clause, most studies segment peer feedback into pieces of comments, or idea units, that represent 'a self-contained message on a single issue' (Cho and Cho 2011, 634), and this unit can range 'from a few words to several sentences' (Nelson and Schunn 2009, 386). These segmented units are then assigned to categories, sometimes repeatedly under different dimensions. Some of these dimensions are parallel (e.g. Novakovich 2016) and others are hierarchical (e.g. Patchan et al. 2013), in that assignment of data into one dimension (such as praise) marks an end but assignment into another dimension (such as criticism) entails further decisions on assignments to be made.

Once segmented, peer feedback data is assigned to pre-established or evolving categories. Coding assignments is mostly not absolute and could be subject to interpretation. Therefore, mutual exclusion among not-so-distinctive categories could pose challenges to researchers. For example, in Cheng, Liang, and Tsai's (2015) analysis where codes were mutually exclusive, it was stated that 'each instance … could only be classified into one category according to the main idea of the message' (80). The distinction between 'affective – supporting' and 'affective – opposing' may be clear-cut. While it may be more difficult to decide between 'affective – opposing', 'cognitive – direct correction' and 'cognitive – guidance', a decision must be made. When the types are mutually exclusive, they had better be very different from each other, or a degree of subjectivity would be involved and may influence the analysis results.

Mutual exclusion does not have to be a rule though. Patchan and Schunn (2015), following the coding schemes established since Nelson and Schunn (2009) and modifying them from one study to the next, made an important revision in their analysis method. This change, as stated by the authors, was that 'praise, problem, and solution were considered independent features rather than mutually exclusive' (602), as the latter had been in their earlier studies. As an illustration of their revised coding scheme, one piece of data was coded five times as 'no praise; problem; no solution; localised; low prose', showing us a peer feedback scenario in which a performance gap was located and identified, but was not completed with a suggested solution. Findings about peer feedback of this kind, if put together, may provide teachers and researchers with a bigger picture of what the average learner can do and fail to do, such as being able to find a problem but not explain why. But by focusing on comparing peer feedback across types, and not considering these types as parts of a whole, existing analyses seem to see only the trees and lose sight of the woods.

Weighting: some peer feedback better than others?

Following coding, results are usually summed so frequencies or percentages can be compared among categories or within and across dimensions. There are also times when data are weighted. For example, Novakovich (2016), with presumptions that peer feedback should contribute to the betterment of existing work and that suggesting improvement is better than explaining critical comments, which in turn is better than general encouraging comments, awards ascending points to the 'naïve', 'editing', 'critical' and 'directive' comments observed.

These presumptions are plausible for training good peer feedback providers, especially after we have learned from research about problems with teacher feedback in that it is usually unclear,

incomprehensible, hard to understand, and does not offer guidance with constructive suggestions (Australian Council for Educational Research 2009; Bailey and Garner 2010; Price et al. 2010). However, studies have also repeatedly pointed out that aiming for more thorough feedback may be futile and missing the mark (Boud and Molloy 2012; Nicol and Macfarlane-Dick 2006; Orsmond and Merry 2011). Teachers know well how open-ended questions can usually engage active learners better than direct telling (Engin 2012). What really matters may be to engage feedback receivers in evaluating and improving their own performance with the help of peer feedback. After all, a passive feedback recipient who accepts and copies all suggestions without reflecting on or challenging the peer feedback may not be so desirable. By the same token, expecting learners to furnish perfect feedback may serve the provider but not the receiver. It may not be the mandatory goal of peer feedback activities all the time and may risk losing sight of the role played by the receiver, whose work is ultimately his/her own responsibility. As Walker (2015) concluded in her study: 'The use recipients made of the feedback depended very little on the characteristics of the feedback received, but did vary strongly across the recipients' (232).

To sum up, in order to establish a peer feedback model that has potential to be used across contexts and help us understand peer feedback better, there seem to be the following guidelines: (1) use an assessment frame of reference; (2) while task criteria may be a prerequisite at the peer feedback stage and do not have to be included, the gap finding and bridging efforts and the explanations revolving around these efforts should form the core of a peer feedback model; (3) the socio-affective and interactive aspects of peer feedback should be addressed; (4) the most common unit of peer feedback comments, which is a complete unit of meaning on one improvable aspect regardless of its length or linguistic structure, may serve most analysis purposes; (5) within cognitive or affective aspects, checking data against peer feedback components may be better than assigning them to mutually exclusive categories; (6) apply different weighting only when it is necessary for the purpose of research.

Primary peer feedback data

In addition to top-down theories and empirical studies, bottom-up primary peer feedback data helped to verify and adjust theoretical reasoning.

Context and participants

At a university in northern Taiwan, 17 students enrolled in an essay writing course, one of the twenty different kinds of elective College English III, for two credits, meeting two hours per week for a total of eighteen weeks. This particular section was taught by the researcher in Spring 2016. Students gave their informed consent to participate and were trained to comment on each other's writing. After in-class marking exercises and teacher modelling of feedback provision, peer feedback opportunities were arranged in between essay drafts and formal submissions, so the drafts could be improved before being submitted for teacher comments and grades. The submissions had to be accompanied by a revision report, in which the author explained where the essay had been revised from the draft and how it was related to the peer feedback received. More details about how assessment abilities were taught and facilitated in an earlier version of this course can be found in Huang (2016).

Data collection

Data was collected from the third assignment, a comparative essay, when learners had practiced peer feedback a few times. Instead of face-to-face peer feedback in the previous two assignments, students were instructed to post their drafts on Moodle, and then provide written feedback to their assigned partners. Each feedback had to address the task criteria of content, organisation, lexis and grammar. Two sets of learner data were discarded because, for one, the revised version was a complete rewrite and, for the other, the rewrite was a different genre.

At the end of the semester, the three students whose overall improvement was most obvious were invited for individual interviews with the teacher/researcher. They were asked about the course experience in general and given their part of the peer feedback data and asked to reflect on the peer feedback and essay revision experiences. Their accounts helped to interpret the data in terms of how students approached and perceived the peer feedback and revision tasks. The interviews were conducted in the learners' mother tongue, Mandarin Chinese, recorded and transcribed. Excerpts included in this report were translated by the author. The entire data-set included the following:

- 15 essay drafts
- 15 sets of written peer feedback
- 15 revised essays submitted for grades and teacher comments
- 15 revision reports
- three individual stimulated recall interviews with three selected learners on their writing and peer feedback experience

Data analysis procedures

The teacher/researcher, with assistance from a research assistant who was also the teaching assistant for the course, analysed the data in two phases. Informed by relevant theories and the literature review on peer feedback analysis methods, the proposed model and its labels took shape out of the first phase - after a few iterations of data review, references to supplementary data, discussions of problems observed, and deletions and mergers of categories. In the second phase after the framework of the model was finalised, the entire feedback set was encoded anew by the established model, again separately by the researcher and the assistant, with very minor discrepancies resolved through discussions.

The GEARed peer feedback model

A GEARed model of peer feedback is proposed. The six possible components of a peer feedback unit are:

G – identifying the **G**ap observed between current performance and the learning goal

E – **E**xplaining and elaborating on the gap to make it fully understood

A – **A**spiration to bridge the gap, with various degrees of success in opinion or suggestion

R – being **R**esourceful, trying to reach beyond current capability and seek outside assistance in the gap-bridging act

e – showing **e**ncouragement toward the receiver

d – expressing **d**oubts and cautions for comments provided toward the provider self

The unit of analysis - a feedback comment, following Walker (2015), was 'a statement relating to a particular shortcoming or praiseworthy item, whether that statement occupies part of a sentence, a whole sentence or several sentences' (236). Explanations and solutions associated with the said particular shortcoming or praiseworthy item, whenever present, were included in the same peer feedback unit so as to preserve the whole picture. Each unit was checked separately against the four capitalised elements for cognitive components. Ideally each component builds upon the previous one(s) to develop into a piece of comprehensible and full-fledged comment. Each peer feedback unit was checked again for the lower-case elements which represent two affective components in peer feedback. Theoretically, all six components can be present in one peer feedback unit at the same time. But in the real-world data, more often than not, different components are omitted for various reasons, such as common background knowledge between the peer feedback giver and receiver and hence no need for the mention, or sheer negligence.

An overview of the peer feedback data using the GEARed model

A total of 94 comments relating the presence of a gap and another 31 with no gap were segmented in the peer feedback data. Each peer feedback unit was checked against the GEARed components. Table 2 shows the number and percentage of each GEARed component. For comments associated with gaps, changes made in the revised essays because of these feedback comments occurred 69 times; i.e. 73% of feedback was incorporated into subsequent work. The affective components were embedded in the language the feedback providers used. As shown in Table 2, e was always associated with cases when there was no gap, and **d** occurred 16% of the time when a gap was present.

The building blocks of peer feedback and their relationship

Each peer feedback element in the GEARed model is discussed in this section with feedback comment examples and student accounts from the stimulated recall interviews.

G: Peer feedback providers did or did not identify something that could be improved. In this English as a foreign language (EFL) writing context, these gaps could be referring to smaller units in the essay, such as word choice or local grammatical problems, or sentences that could be improved, or larger units such as the organisation or content that would require more substantial revision, as in these two excerpts:

(1) The first two sentences in your second paragraph … (followed by a direct quote)
(2) But I think the expenses for watching movies at home should not be limited to rental or down-loading fees.**E:** After identifying the gap, or an area that could be improved, 55% of the feedback comments explained further, the other 45% did not. Most explanations were brief and some involved clarification requests, such as these two examples:

(3) Using this word is not appropriate here because it means …
(4) Do you mean xxx? If so, you should …

Some explanations seemed to be vaguer, but, according to the students interviewed, vague explanations could still serve the purpose. In such cases, the feedback receiver picked up most of the unfinished work and tried to figure out why it was a gap and how to solve the problem:

(5) It's abrupt when you suddenly switch to …
(6) I don't understand this.**A:** Eighty-eight per cent of the feedback comments contained the 'A' element, aspirations for improvement with ideas. They were related to local revisions such as word choice and grammar correction, or to global revisions such as organisation and content, usually above the sentence level. While ideas for local problems were mostly quite specific and closed-ended, global issues were more often accompanied by open-ended suggestions, such as:

(7) I think you can briefly introduce what companies A and B are so that readers may have necessary background knowledge to understand your comparison.
(8) You have discussed three differences. I think the second one could stand alone as one paragraph.

Table 2. Composition of peer feedback (PF) components from 15 learners.

GEARed components	PF units	PF on each cognitive parts				Revision made	PF on affect	
		G	E	A	R		e	d
With gap	94	92	51	83	8	69	0	15
Percentage	100	98	54	88	9	73	0	16
Without gap	31	31	27	0	0	0	31	0
Percentage	100	100	87	0	0	0	100	0

(9) I think you could elaborate more in the concluding paragraph, maybe by referring back to the hook at the beginning and saying more.**R:** About 9% of the comments involved the element of being resourceful. This part exemplified the situation when the feedback provider tried to reach out beyond his/her current proficiency level and obtain credible assistance to deal with the problem at hand. An element not derived from theories or existing methodologies, this peer feedback feature surfaced from learner data and signified an important step of the peer feedback provider towards self-regulated learning in the gap-bridging effort:

(10) After consulting the Oxford dictionary, I found that the preposition in your first paragraph for the phrase 'play a crucial role' should be 'in'.

(11) The word 'venture', according to the dictionary, means 'a business project which involves taking risks'. I think you should use 'startup' instead.

(12) I checked a collocation dictionary. There is no such usage as 'rank slow', just 'high/low rank'.**e:** There were 31 occurrences of comments showing no gap in performance. Compared to 94 with a gap, that is a ratio of 1 to 3. Eighty-seven per cent of these were compliments, i.e. 27 out of 31, indicating why it was praiseworthy. According to interviewees, these praises were very important to them because 'then you know you didn't perform so poorly'. Although being labelled as 'naïve' and considered not helpful in improving the work (Novakovich 2016; Sluijsmans et al. 2004), praise and its value may have been underestimated and may lie at a less utilitarian, yet perhaps equally or more important, level.

d: Another interesting phenomenon not implicated by theories was observed: the expression of doubt about proposed solutions. This was similar to Nelson and Schunn's (2009) earlier peer feedback types of 'affective language: mitigation – compliment, mitigation – other, and neutral'. This occurred an average of one instance per peer feedback provider in this data-set. Expressions of doubt were mostly attached to suggestions and deliberately hedged by saying 'just my personal opinion', 'for your reference' or 'just one minor point for you to consider'. Whether it was culturally influenced modesty, a surface courtesy or an educated prudence may deserve further study.

Observed patterns of peer feedback component combinations

By checking for the presence of each element in peer feedback comments, six combinations were found, as shown in Table 3. For example, G_ _ _ represents the kind of comments that had only identified the gap; GE_ _ tags comments that had identification plus explanation; G_A_ refers to those comments that identified the gap and aspired to bridge it, but lacked an explanation and outside resources. From the table, the most common combinations were G_A_ and GEA_, representing 37 and 39% of the data respectively.

G_ _ _: This was rare. But when the problem was minor and self-evident, no explanation or suggestion seemed necessary:

(13) There are two 'also' in the first sentence of your fourth paragraph.**GE_ _:** This type of comment explained the gap identified but offered no suggestion. When the expected improvement required more substantial work, the peer feedback provider was likely to leave it to the author. Student interviewees felt that this was not a problem because it was the author's job to revise and address the issue identified:

Table 3. Distribution of different GEAR patterns.

	G_ _ _	GE_ _	G_ A_	GEA_	GEAR	_ _A_	Total
	G_ _ _	GE_ _	G_ A_	GEA_	GEAR	_ _A_	Total
Instances	5	7	35	37	8	2	94
Percentage	5	7	37	39	9	2	100

(14) I am kind of lost in the fourth paragraph. The first part is about a unique function that Dcard has. But when you switch to PTT, you talk about anonymous mechanism instead of the same unique function. Moreover, I think PTT has anonymous mechanism as well. We use nicknames instead of real names on PTT. Isn't this a kind of anonymous function?**G_A_:** In 39% of the comments, explanation was omitted but suggestions were offered. Student interviewees reported that explanations were not always necessary since both the peer feedback provider and receiver shared common background knowledge:

(15) I think it would be more appropriate you removed 'the' in 'the expanding of American culture' and changed it to 'expansion of American culture'.**GEA_:** This combination in which a gap was identified and explained, and suggestions for improvement offered can be considered typical. It also ranked highest among all combinations in the current data-set:

(16) Based on what I understand, 'The French fries McDonald's sold were thin, crisp and more salty'. You are comparing it to KFC's, so you should change all adjectives to comparatives, making it 'thinner, crispier and saltier'.**GEAR:** In eight instances, an identified gap was accompanied by an explanation, a suggestion, and reference to resources:

(17) The word 'tool' in your first paragraph refers to 'an instrument such as a hammer, screwdriver, saw, etc. that you hold in your hand and use for making things, repairing things, etc'. (Oxford dictionary). So it may not be the right word here, 'equipment' may be more suitable.**_ _A_:** Theoretically, each piece of comment should start with G. G is the foundation for feedback comments to develop. Without a clear G, feedback may look vague. But in the data-set, two comments were found without a clearly perceptible G, but involved A (aspirations to bridge a gap):

(18) You should use wordings such as 'First', 'Secondly' to start each paragraph.
(19) I just think that you can consider rephrasing all your points above and put them in the last paragraph to make the solution stronger.

Although one can infer the gap that the peer feedback providers meant, i.e. in (18) the shortcoming was that points not introduced clearly enough and in (19) the weakness was that the solution was not strong, the data analysis decision was to keep the literal meaning without inferring too much. These two instances may showcase real-world situations in which the peer feedback provider fails to point out a clear gap, yet still provides something for the peer feedback receiver to ponder upon.

Significance of the GEARed model

By using the GEARed model, peer feedback comments were examined both by components and in their entirety. Peer feedback comments varied in their degree of comprehensiveness. G, identifying a gap or otherwise, seemed to be the foundation of feedback, although in very rare cases it was absent. E, elaborating on the nature of the gap, although conventionally desirable when teachers point out problems in student work, seems to be something that can be bypassed when a problem is obvious or when both parties hold mutual background knowledge. Element A, aspirations to bridge the gap, was not necessary in the case of a praise. Given a gap, A could be built upon G and E, or additionally followed up by R, or simply follow G without E and R, or stand alone as in _ _A_. In cases in which peer feedback activities were designed as a means to an end, A is a highly desirable component as it can benefit the provider more than the receiver.

It has to be noted that not all A's were successful and A could be successful in various forms. For willing and engaged feedback receivers, incomplete A's may provide more learning opportunities than an A that does it all. R is a part less commonly discussed in formative assessment literature. In the context where this study was carried out, writing resources such as dictionaries of various kinds and corpus concordances had been introduced as a part in the syllabus. Learner assessors readily applied

such budding knowledge to help them solve the problems at hand, which in theory, could be seen as trying to reach beyond the zone of proximal development (ZPD, Vygotsky 1987), and being on the way to becoming independent learners with self-assessment and problem-solving abilities.

Taken together, the percentages in Table 2 help depict a feedback profile unique to this data-set, as shown in Figure 1. Each bar in the GEAR chart indicates what peer feedback providers had contributed in different phases of GEAR to the gap-bridging act, despite their success being to various degrees. For the receiver, the work was not completed here. Under an instructional design to facilitate learners, the feedback receiver is supposed to take over and fill out the remaining part in each bar to present an improved work that stretched beyond his/her ZPD.

On the affective side, e, encouraging comments, seemed to be important to peer feedback receivers, although in past feedback studies this part has been considered to be unhelpful, labelled as 'naïve', or weighed much less in feedback scoring because it did not make the work better. But student informants showed how compliments could boost morale for them, especially when critiques constitute the major part of peer feedback. The d, expressing doubts, was unanimously pointed toward the providers themselves. Such characteristic may be mostly absent in teacher feedback, but for learners it seemed to demonstrate an educated cautious attitude, which could be considered healthy and facilitative to learning.

Hypothetical GEARed patterns and their pedagogical significance

The combinations of GEAR components shown were not exhaustive. There may be combinations such as G_ _R, G_AR, or other anomalies such as that of _ _A_ existing in real classrooms. They may have different implications for educators. A few theoretically possible patterns are suggested and discussed here, as shown in Figure 2.

Figure 2(a) features a kind of feedback high in G and E, but low in A and R. In this hypothetical case, learners need help in pointing out the gap and explaining it more fully. Once they understand why it is a gap, they can bridge the gap and reach for resources on their own. Figure 2(b) is a case wherein the gap is more self-explanatory and the teacher's job is focused on demonstrating how to improve the work and use resources. Figure 2(c) features a scenario in which the gap is merely pointed out and resources are referred to, leaving it up to the learner to figure out why it is a gap and try to bridge it by using the resources provided or referred to. Different patterns of peer feedback may elicit different learning behaviours.

We need more empirical data to be examined against the proposed GEARed model to help us better understand feedback and how it works in different contexts.
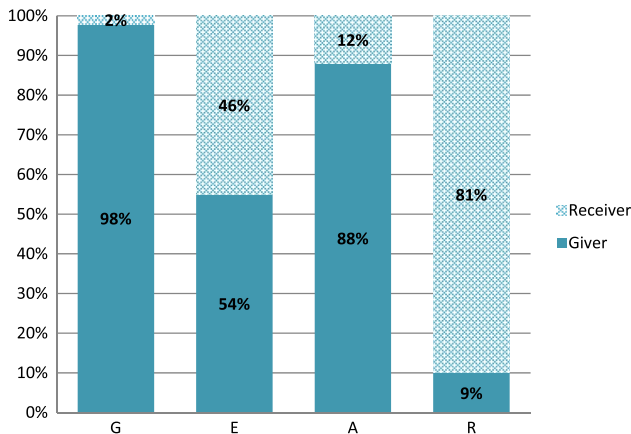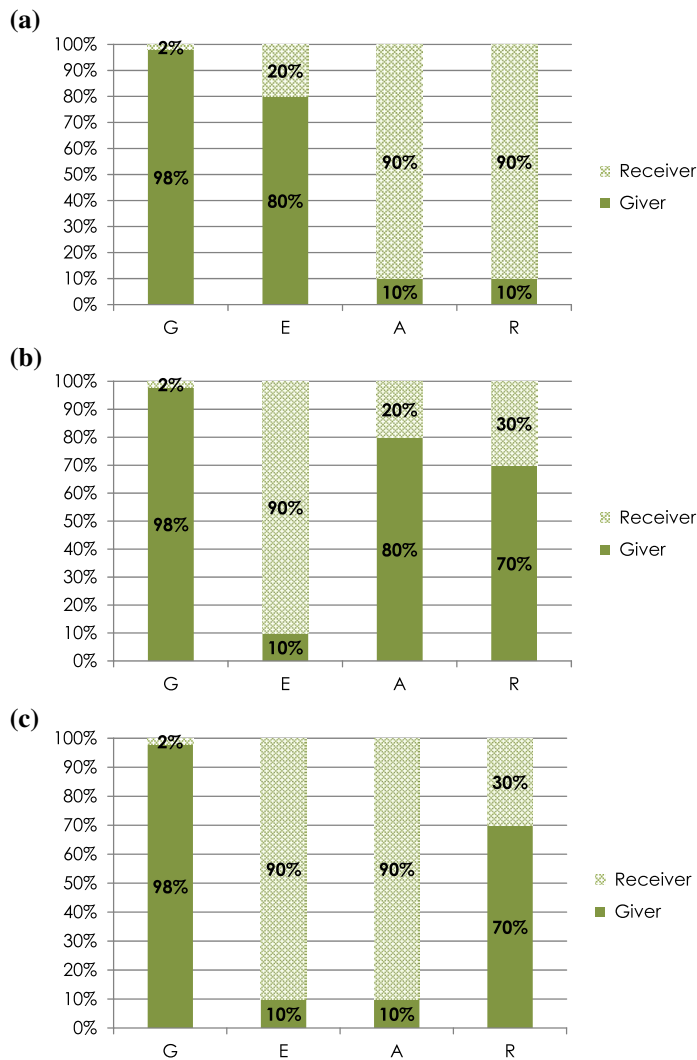

Figure 1. Peer feedback pattern observed.

Figure 2. Hypothetical feedback patterns for different learning needs.

Conclusion, implications for future research and limitations

In conclusion, this GEARed model was generated with the components G, E, A, e mostly based on theories and empirical studies and R and d informed by primary data. Its significance lies in viewing peer feedback by examining its different components side by side and drawing teaching implications from the distribution and combinations of these peer feedback components. The proposed GEARed model addressed two areas that were somewhat neglected in past peer feedback research, one on the role of peer feedback receiver and the other on viewing peer feedback as the whole of its key characteristics, highly related to each other, and not separating them.

The lonely journey of improving one's work is different with the presence of peer feedback. The classroom setting has value because there are teachers and peers to help, people who have the same goal and interests and who would care. They can help, but it's not their job. The good thing about peer feedback is that it serves both the giver and the receiver, with the effect contingent on their respective level of involvement. As shown in the students' accounts, peer feedback served as a catalyst to point to

what the author had not noticed originally. Once the author was called upon to notice by peer feedback comments, even when there was vague or zero explanation, a learner who cared to improve the work would take over and learn and improve.

In the past when we questioned the effectiveness of teacher feedback (Bailey and Garner 2010; Price et al. 2010), we seemed to expect a perfect job from teachers. Later we realised that even a perfect job on the giver's side will not suffice; feedback has to engage the receivers for it's how the receiver responds that ultimately matters (Barker and Pinard 2014; Bloxham and Campbell 2010; Nicol 2010), so peer feedback has been used as a means to engage learners. But perhaps it is easier to examine the tangible peer feedback data provided by the giver than the intangible functioning in the feedback receiver's mind, much emphasis has been put on the data and hence the giver and we are tempted to ask for more from the giver, as we had with teachers. The GEARed model made this case clear by pointing to the unfilled upper parts of the bars in Figures 1 and 2 that could and should be picked up by the receiver. As reported by student informants, a flawed peer feedback comment could still help if the receiver cared to think about it and work further on it. In this way, peer feedback served the receiver perfectly and the purpose of an imperfect peer feedback is fulfilled.

By the same token, we may lose sight of the whole picture if peer feedback components are examined separately and not as a whole. With the whole picture obtained through peer feedback data analysed in a GEARed model lens, such as the information revealed by the average peer feedback giver as shown in Figure 1 or the different hypothetical scenarios in Figure 2, the teacher, aided by his/her contextual knowledge of the learners and their learning, is able to diagnose learners' needs better and provide more focused guidance. Instead of giving meticulous explanations and demonstrating all solutions or expecting peer feedback givers to do so, a teacher well informed by the GEARed profile of his/her learners' peer feedback can forego the E in Figure 2(a) or the R in Figure 2(c) and target specifically at the needed areas and help learners more effectively.

The value of the proposed model may be more conceptual than pragmatic. The model does not offer any quick solution to teachers' day-to-day work, but a perspective to view peer feedback and reflect upon facilitation to learners. Limited by the small data-set from a foreign language writing context and that most learner tasks referred to in the literature included in Table 1 are writing assignments, the GEARed model's applicability needs to be subject to more scrutiny with more empirical data in the future. Moreover, the three student informants giving their accounts revealed how active learners viewed and processed peer feedback. Accounts from less active and average learners may also be needed in future studies.

Acknowledgement

Disclosure statement

Funding

Notes on contributor

Shu-Chen Huang has been teaching English as a foreign language (EFL) in Taiwan at the tertiary level for more than 20 years. Her research interests include EFL pedagogy, learning motivation and formative assessment. She has published in journals such as Assessment and Evaluation in Higher Education, Language Testing and Language Assessment Quarterly.

References

ACER (Australian Council for Educational Research). 2009. Engaging Students for Success: Australasian Student Engagement Report. http://www.acer.edu.au/documents/AUSSE_ASERReportWebVersion.pdf.

*Allen, D., and A. Katayama. 2016. "Relative Second Language Proficiency and the Giving and Receiving of Written Peer Feedback." *System* 56: 96–106.

Bailey, R., and M. Garner. 2010. "Is the Feedback in Higher Education Assessment worth the Paper It is Written on? Teachers' Reflection on Their Practices." *Teaching in Higher Education* 15: 187–198.

Barker, M., and M. Pinard. 2014. "Closing the Feedback Loop? Iterative Feedback between Tutor and Student in Coursework Assessments." *Assessment & Evaluation in Higher Education* 39: 899–915.

Black, P., and D. Wiliam. 2009. "Developing the Theory of Formative Assessment." *Educational Assessment, Evaluation and Accountability* 21: 5–31.

Bloxham, S., and L. Campbell. 2010. "Generating Dialogue in Assessment Feedback: Exploring the Use of Interactive Cover Sheets." *Assessment & Evaluation in Higher Education* 35: 291–300.

Boud, D., and E. Molloy. 2012. "Rethinking Models of Feedback for Learning: The Challenge of Design." *Assessment & Evaluation in Higher Education* 38: 698–712.

Brookhart, S. M. 2004. "Classroom Assessment: Tensions and Interactions in Theory and Practice." *Teachers College Record* 106: 429–458.

*Chang, C. Y.-H. 2015. "Teaching Modelling on EFL Reviewers' Audience-Aware Feedback and Affectivity in L2 Peer Review." *Assessing Writing* 25: 2–21.

*Cheng, K.-H., J.-C. Liang, and C.-C. Tsai. 2015. "Examining the Role of Feedback Messages in Undergraduate Students' Writing Performance during an Online Peer Assessment Activity." *The Internet and Higher Education* 25: 78–84.

*Cho, Y. H., and K. Cho. 2011. "Peer Reviewers Learn from Giving Comments." *Instructional Science* 39: 629–643.

*Davies, P. 2006. "Peer Assessment: Judging the Quality of Students' Work by Comments rather than Marks." *Innovations in Education and Teaching International* 43: 69–82.

Engin, M. 2012. "Trainer Talk: Levels of Intervention." *ELT Journal* 67: 11–19.

Gielen, S., F. Dochy, and P. Onghena. 2011. "An Inventory of Peer Assessment Diversity." *Assessment & Evaluation in Higher Education* 36: 137–155.

Hattie, J., and H. Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77: 81–112.

Huang, S.-C. 2016. "No Longer a Teacher Monologue—Involving EFL Writing Learners in Teachers' Assessment and Feedback Processes." *Taiwan Journal of TESOL* 13: 1–31.

*Ion, G., A. Barrera-Corominas, and M. Tomas-Folch. 2016. "Written Peer Feedback to Enhance Students' Current and Future Learning." *International Journal of Educational Technology in Higher Education* 13: 1–11.

*Lin, S. S. J., E. Z. F. Liu, and S. M. Yuan. 2001. "Web-Based Peer Assessment: Feedback for Students with Various Thinking-Styles." *Journal of Computer Assisted Learning* 17: 420–432.

*Nelson, M. M., and C. D. Schunn. 2009. "The Nature of Feedback: How Different Types of Peer Feedback Affect Writing Performance." *Instructional Science* 37: 375–401.

Nicol, D. 2010. "From Monologue to Dialogue: Improving Written Feedback Processes in Mass Higher Education." *Assessment & Evaluation in Higher Education* 35: 501–517.

Nicol, D. 2011. "Good Designs for Written Feedback for Students." In *McKeachie's Teaching Tips: Strategies, Research, and Theory for College and University Teachers*. 13th ed, edited by W. J. McKeachie and M. Svinicki, 108–124. Cengage Learning, Belmont USA: Wadsworth.

Nicol, D., and D. Macfarlane-Dick. 2006. "Formative Assessment and Self-Regulated Learning: A Model and Seven Principles of Good Feedback Practice." *Studies in Higher Education* 31: 199–218.

Nicol, D., A. Thomson, and C. Breslin. 2014. "Rethinking Feedback Practices in Higher Education: A Peer Review Perspective." *Assessment & Evaluation in Higher Education* 39: 102–122.

*Noroozi, O., H. Biemans, and M. Mulder. 2016. "Relations between Scripted Online Peer Feedback Processes and Quality of Written Argumentative Essay." *The Internet and Higher Education* 31: 20–31.

*Novakovich, J. 2016. "Fostering Critical Thinking and Reflection through Blog-Mediated Peer Feedback." *Journal of Computer Assisted Learning* 32: 16–30.

Orsmond, P., and S. Merry. 2011. "Feedback Alignment: Effective and Ineffective Links between Tutors' and Students' Understanding of Coursework Feedback." *Assessment & Evaluation in Higher Education* 36: 125–136.

*Patchan, M. M., B. Hawk, C. A. Stevens, and C. D. Schunn. 2013. "The Effect of Skill Diversity on Commenting and Revisions." *Instructional Science* 41: 381–405.

*Patchan, M. M., and C. D. Schunn. 2015. "Understanding the Benefits of Providing Peer Feedback: How Students Respond to Peers' Texts of Varying Quality." *Instructional Science* 43: 591–614.

*Pham, V. P. H., and S. Usaha. 2016. "Blog-Based Peer Response for L2 Writing Revision." *Computer Assisted Language Learning* 29: 724–748.

Price, M., K. Handley, J. Millar, and B. O'Donovan. 2010. "Feedback : All That Effort, but What is the Effect?" *Assessment and Evaluation in Higher Education* 35: 277–289.

Reinholz, D. 2016. "The Assessment Cycle: A Model for Learning through Peer Assessment." *Assessment & Evaluation in Higher Education* 41: 301–315.

*Sluijsmans, D. M. A., S. Brand-Gruwel, J. J. G. van Merriënboer, and R. L. Martens. 2004. "Training Teachers in Peer-Assessment Skills: Effects on Performance and Perceptions." *Innovations in Education and Teaching International* 41: 59–78.

Topping, K. 1998. "Peer Assessment between Students in Colleges and Universities." *Review of Educational Research* 68: 249–276.

Vygotsky, L. S. 1987. *The Collected Works of L. S. Vygotsky: Problems of the General Psychology, including the Volume Thinking and Speech*. Vol. 1. Edited by R. Rieber and A. S. Carton. New York: Plenum Press.

*Walker, M. 2015. "The Quality of Written Peer Feedback on Undergraduates' Draft Answers to an Assignment, and the Use Made of the Feedback." *Assessment & Evaluation in Higher Education* 40: 232–247.

Van Zundert, M., D. Sluijsmans, and J. van Merriënboer. 2010. "Effective Peer Assessment Processes: Research Findings and Future Directions." *Learning and Instruction* 20: 270–279.