Educ Asse Eval Acc (2009) 21:347–367
DOI 10.1007/s11092-009-9082-2

"I really need feedback to learn:" students' perspectives on the effectiveness of the differential feedback messages

Anastasiya A. Lipnevich · Jeffrey K. Smith

Received: 9 November 2008 / Accepted: 17 August 2009 /
Published online: 26 August 2009

Abstract The current study examined students' perceptions of the effects of different forms of instructional feedback on their performance, motivation, and emotion. Forty-nine students attending an eastern US university participated in focus group discussions. The groups explored students' reactions to grades, praise, and computer versus instructor provided feedback, as well as students' views of the ideal feedback. Students named detailed comments as the most important and useful form of feedback. Grades were deemed to be unnecessary if the goal of an activity was to learn. Students proposed that low grades elicit negative affect and damage the students' sense of self-efficacy, and high grades decrease motivation and lessen students' perceived need to improve. Praise was reported to positively affect emotion, but not to be directly conducive to learning.

Keywords Formative assessment · Feedback · Grades · Praise · Source of information · Focus groups

1 Introduction

In the majority of educational settings, a significant portion of a teacher's time is devoted to the assessment of students (Crooks 1988; Orrell 2006). Assessment is

A. A. Lipnevich (✉)
Educational Testing Service, Rosedale Road, R-18, Princeton, NJ 08541, USA
e-mail: a.lipnevich@gmail.com

J. K. Smith
University of Otago—EARU, P.O. Box 56, Corner of George and Union, Dunedin, New Zealand
e-mail: jeffreyksmith@gmail.com

used for both formative and summative purposes, and in many instances, these purposes are combined. That is, the assessment is used to help students learn, but also contributes to the students' grade for a course. Thus, on one hand, assessment provides students with the necessary information that would allow them to adjust their learning strategies, get rid of existing misconceptions, correct mistakes, increase (or sustain) motivation, and, ultimately, move them closer to the desired goals. On the other hand, it serves as the primary mechanism for giving students grades. Research has shown that the blending of these purposes of assessment very often has an adverse effect in terms of student response to assessment feedback (Hattie and Timperley 2007; Kluger and DeNisi 1996; Shute 2007; Torrance 1993; Wiliam and Thompson 2007). Possible reasons for such a seeming contradiction are hidden in specific attributes of evaluative practices, with the quality and type of feedback being among the most important.

Our understanding of differential effects of assessment in general and feedback in particular on students' learning is crucial to ensure optimal outcomes of education. The purpose of the investigation presented here was to better understand how students react to various aspects and approaches to the assessment feedback they receive. This study is a follow-up to a larger, controlled experiment that looked into the effects of assessment feedback. In the component of the study reported here, we listened directly to the voices of students through a series of focus groups, conducted after the conclusion of the experiment, presented elsewhere (Lipnevich and Smith 2008) and summarized below. Our focus in the overall study was to look at several factors involving assessment feedback. First, we wanted to look at the effects of providing students with detailed feedback on an essay examination where students had the opportunity to revise their work after receiving feedback. Second, we wanted to look at the source of the information and its effect on students' reaction to feedback. Third, we wanted to look at the effects of receiving or not receiving a preliminary grade, and fourth we wanted to look at the effects of receiving or not receiving a statement of praise in the feedback.

1.1 The effects of grades on receipt of feedback

The most common type of feedback that students receive in a typical classroom is grades, and more often than not, a letter grade or a numeric score by itself (Marzano 2000; Oosterhof 2001). Grades provide a convenient summary of students' performance, and inform all interested parties of students' achievement. However, it is clear that the grade as a summary of achievement is fundamentally summative in nature (Airasian 1994; Smith and Gorard 2005).

One of the main conclusions Black and Wiliam (1998) draw from their review of literature on formative assessment is that descriptive feedback, rather than letter grades or scores, leads to the highest improvements in performance. Moreover, evidence from several studies that investigated the effect of differential feedback on learning suggests that using grades to improve learning is simply not effective. For example, Butler and Nisan (1986) compared effects of constructive feedback and grades. The researchers concluded that grades emphasized quantitative aspects of learning, depressed creativity, fostered fear of failure, and weakened students' interest. Quite opposite to this pattern, no negative consequences followed from the

use of task-specific individualized comments. In a later study, Butler (1988) found comments specifically tailored to students' performance resulted in a significant increase in scores (by almost 30%) on a task. The group that received only grades showed a significant decline in scores, as did the group that received both grades and comments. Analysis of students' reports of interest in performing the task demonstrated a similar pattern, with interest being undermined for both graded conditions. Interestingly, high achievers in all three feedback regimes sustained a high level of interest, whereas low achievers in the graded groups evidenced dramatic declines (Butler 1988).

Other studies investigating the impact of grades on students' learning present evidence that is in agreement with Butler. For example, in an experiment conducted by Grolnick and Ryan (1987), students who were told they would be graded on how well they learned a social studies lesson had more trouble understanding the main point of the text than did students who were told that no grades would be involved. Even on a measure of rote recall, the graded group remembered fewer facts a week later.

The explanations of negative effects of grades on students' performance vary. Butler and Nisan (1986) and Butler (1988) propose that normative grades inform students' about proficiency relative to others, whereas individualized comments create clear standards for self-evaluation specific for the task. The researchers discuss these results in terms of cognitive evaluation theory, and posit that even if feedback comments are helpful for students' work, their effect can be undermined by the negative motivational effects of the normative feedback, that is, by giving grades and scores (Butler 1988).

Elawar and Corno (1985) look at their findings through the lens of cognitive theory and research, which emphasize the importance of deep processing when acquiring complex information. Comments provided by teachers turn students' attention to relevant, specific information, stimulate mental elaboration, and as a result boost performance. Grades, perceived as reinforcers and punishers, which are believed to be controlling and lacking specificity, lead to inhibition of students' cognitive processes and slow the progress of learning.

The argument that grades are detrimental to students' performance is commonly heard, but it is not the only one in the field of assessment. In an attempt to refute a commonly voiced urge to abolish grades, Marzano (2000) states that the most important purpose for grades is to provide feedback to students and, if referencing for grading is content-specific, letter grades and numerical scores will lead to an increase in students' performance. He postulates that if students have a clear understanding of the requirements of the task and if grading is based on students' achievement and effort only, students can increase their level of knowledge and understanding based on grades alone. Guskey and Bailey (2001) take a similar stance on the issue of grades. They suggest that if grading is done properly, an increase in students' academic attainment will follow.

Overall, although the review of the studies on grading are not supportive of its use in facilitating learning, there has not been extensive recent research that has inquired into the effects of grades alone or in combination with other types of feedback on students' performance. Students' reactions to grading have not yet been explored either.

1.2 The effects of praise on receipt of feedback

Praise has been defined as "favorable interpersonal feedback" (Baumeister et al. 1990, p. 131), or "positive evaluations made by a person of another's products, performances or attributes "(Kanouse et al. 1981, p. 98). This type of feedback is probably the second most common kind (with the first being grades) that students' receive from their teachers, and it runs the gamut from simple "You did a great job!" statements to much more elaborate and personalized positive references to students' performance. Generally, praise is believed to have beneficial effects on students' self-esteem, motivation, and performance. As a result, teachers are encouraged to use praise as a reinforcer of a desired behavior (Dev 1997). However, quite similarly to the research on grading, the conclusions concerning the impact of praise on students' performance are not consistent.

There are two opposing views on the effect of praise on students' learning. One camp of researchers and educators claims that normally, a feedback message containing praise enhances motivation and leads to improvement of individuals' performance (Cameron and Pierce 1994; Dev 1997; Pintrich and Schunk 2002). Shanab et al. (1981) investigated the influence of praise on motivation and found that praise during a puzzle-solving task led undergraduates to spend more time on the task and to rate their interest as higher than participants in a control condition who received neutral feedback. Similarly, meta-analytic studies examining the effects of praise on motivation have shown that positive statements have a tendency to increase intrinsic motivation across a variety of dependent measures (Cameron and Pierce 1994; Deci et al. 1999). This effect, however, is not always strong, varies for different age groups, and often has been derived in the course of methodologically flawed studies (Henderlong and Lepper 2002; Lepper et al. 1999).

There are also examples of the negative impact of praise on students' learning. Baumeister et al. (1990) present evidence that praise can both impede and facilitate individuals' performance. Their analyses showed that positively framed feedback improved students' performance on a pure effort task, but consistently led to impairment in skilled performance. Additionally, the researchers found that both task-relevant and task-irrelevant praise resulted in performance decrements. The researchers proposed several possible mechanisms by which praise could impede successful task completion. The most logical and parsimonious explanation (as deemed by the authors) was that praise made individuals self-conscious and led to disruption of skilled performance. Apparently, attention to the self, resulting from praise, robs cognitive resources that would otherwise be committed to the task. Only if a task is automated, and fewer resources are needed for its completion, would praise have a neutral or positive effect on performance. Therefore, the assumption that praise focuses attention on self, and not the task, seems to be the most plausible explanation of the negative effect of praise on performance.

In sum, there is ample evidence providing support for claims at both ends of the praise spectrum. However, this evidence is inconclusive, and to our knowledge, there are no studies that attempted to gather students' reactions to the receipt of praise.

1.3 The effects of source on receipt of feedback

Computer assisted instruction, use of hypermedia, and sophisticated learning environments have become an ingrained part of modern instructional practices. One of the main functions of many of these complex educational technology systems is to provide students with feedback about their performance. If the effect of teacher-provided feedback is unclear, the impact of computer-provided feedback is even more obscure. There are a few studies showing positive effects of feedback from machines on individuals' performance and affect, but they do not involve authentic learning tasks (Mishra 2006), and are mostly conducted in the area of organizational psychology (Earley 1988; Landauer et al. 2003; Mishra et al. 2001; Nass et al. 1996).

Earley (1988) inquired into a contrast between computerized feedback and feedback provided by the supervisor in a subscription-processing job. The results showed that computerized feedback was more trusted, led to stronger feelings of self-efficacy, to more strategy development, and to better performance compared with identical feedback coming from a supervisor. These findings seem to support the argument of those researchers who believe that computers are perceived by individuals as neutral tools, and consequently, unbiased sources of information (Lajoie and Derry 1993; Lepper et al. 1993). Because machines do not elicit affective responses from individuals, cognitive resources get directed towards tasks resulting in an increase in performance. Feedback provided by the supervisor could have directed participants' attention to evaluating the intentions of the supervisor and their implications for goals of the self, whereas the computerized feedback directed attention to the task and to the task details.

A more recent study conducted by Mishra (2006), investigated the effects of feedback provided by computer. Analysis of the results showed that computer-provided feedback made a significant difference in the participants' motivation and affect. Praise provided by the computer had a uniform positive impact on participants' motivation and affect. Mishra's (2006) study provides initial answers to questions concerning individuals' reaction to computer-provided feedback. It shows that students do form affective reactions towards feedback provided by the machine, but the nature of the differences between their reactions to computer-provided feedback and their reactions toward human-provided feedback remains unclear.

Thus, there is no consensus in the field on the effects of computer versus instructor provided feedback. The current study intended to glean students' reactions toward what they thought was computer and instructor provided feedback.

1.4 The context for the study

In an attempt to fill the existing gap in the research literature, two complementary studies were designed to systematically explore student reactions to differential feedback messages. The first study was an experiment aimed at teasing out the effects of grades, praise, and source of feedback on students' performance. The second study, discussed in the current article, consisted of a series of focus groups conducted to gain a more comprehensive picture of students' responses to feedback and

to substantiate the findings of the experiment. A selective synopsis of the experimental study follows. For complete description, see Lipnevich and Smith (2008).

The experiment included three conditions with some students not receiving detailed feedback on their performance on an essay examination, other students receiving detailed feedback with an understanding that their feedback came from the course instructor, and a final group of students believing that their feedback was computer-generated. Additionally, the three conditions were crossed with two factors of grade (grade or no grade) and praise (praise or no praise), resulting in a $3 \times 2 \times 2$ design. The dependent measure included an authentic learning task with students working on an essay exam and then coming back a week later to revise it based on the feedback. The exam was a part of a course requirement and, therefore, was expected to be taken seriously by the participants.

The most salient finding of the experiment was that descriptive feedback specific to individual work is critical to students' improvement. Detailed, specific, descriptive feedback which focuses students' attention on their work, rather then the self, is the most advantageous kind of information for students. The benefits of such feedback occur at all levels of performance. Evaluative feedback in the form of grades may be helpful if no other options are available, and can beneficially be accompanied by some form of encouragement. At the same time, grades were shown to decrease the effect of detailed feedback. It appears that this occurs because it reduces a sense of self-efficacy and elicits negative affect around the assessment task.

The experimental study brought more clarity to the existing understanding of the effects of feedback on students' progress of learning. To uncover potential reasons behind certain effects of feedback, a series of follow-up focus group discussions were conducted. The results of the focus group analyses are the focus of the present study.

2 Method

2.1 Participants

Six focus groups, each consisting of eight to nine students attending an eastern US university, were conducted two weeks following the completion of the experiment. For each group, students were selected based on the source of feedback they received and the presence or absence of a grade. Thus, separate focus groups were held for students in the no feedback condition who received a grade and those who did not receive a grade, for students in the instructor condition who received a grade and those who did not, and, finally, for students in the computer feedback condition with the grade and without the grade. Additionally, each group included an equal number of students who did and did not receive praise. The students were advised that for participating, they would receive five points (out of a maximum of 100) towards their final exam score.

2.2 Procedures

A total of six focus groups were conducted, each approximately 40 min long. Assuming that the participants would be more candid in the presence of a person they already knew, the experimenter served as a focus group moderator. A semi-structured discussion

was used to elicit participants' responses to the set of questions. It employed a specification of topics to be covered, but flexibility in terms of the sequence and the phrasing of the questions in the course of the focus group. The semi-structured discussion guide used in each of the six focus groups had the following questions:

- How did you react to the feedback? How did you go about revisions?
- Did you trust your feedback? Did you find it accurate? Did you think it was fair and worthwhile?
- How did your grade help you? Would you have liked to know your grade before you began your revisions?
- Was the praise helpful? Would you have liked to receive praise on your performance?
- What would be your ideal kind of feedback?

2.3 Analysis

All focus group discussions were transcribed verbatim. Formal analysis of the data records began with deductive coding. Categories included students' interpretations of the effects of grades, praise, and feedback source on their performance, their perceptions of the effect of feedback on motivation, affect, and self-efficacy, and their views on what would be ideal feedback. The coding was carried out separately for each of the six focus groups. Summaries of findings were compiled for each group to present a general picture of participants' experiences and their reactions to the specific type of feedback they received. The data were then compared across the six groups, with the analysis being guided by the findings of the experiment. This step allowed for determining similarities and differences for participants who had different experience in terms of feedback.

3 Results

Analysis of the experimental data revealed intriguing patterns in students' responses to differential feedback. To further explore students' reactions, focus group discussions, held shortly after students completed their essay exams, were analyzed with the aim to substantiate the quantitative findings with students' accounts of their experiences. Examination of students' perceptions of various feedback messages provides a more comprehensive picture of their responses to instructional feedback and of its effects on their motivation and emotion. Such analysis helps to identify the kinds of feedback that best support learning and gives voice to the key participants of the educational process—students.

3.1 Instructor feedback with a grade

This group consisted of eight students, all of whom received a grade and personalized comments with the understanding that the feedback was provided by the instructor. In addition to a grade and comments, four of the students received a general statement praising them on their performance.

"I loved the feedback," "the comments were really helpful," "I was like, wow, he's actually giving us feedback. Like, he read all of the essays! He's giving us feedback and it was actually cool" were among the first responses of the focus group participants to the most general question asked: "How did you react to the feedback?" When prompted, the participants expanded and talked about the importance and usefulness of comments for making revisions and ensuring improvement. One of the participants noted "I was relieved when I went through the comments. It felt nice to know exactly what I needed to do." This statement resonates with the views of the participants in this group.

Although there was a consensus about the effectiveness of personalized comments, the participants' reactions to the grade they had received were not uniform for this group. Some students described their main reaction towards their grade as "panic," "shame," "disappointment," and "anger." Students elaborated by saying that receiving a grade which was much lower than they had anticipated was discouraging, and it took time and special effort to convince themselves to move forward with their revisions. The students noted that the presence of comments made it much easier to resume their work on their essays. They knew exactly what to do, and tried to take each suggestion the professor provided and incorporate it into their essay. Interestingly, several participants emphasized that their negative emotion was directed at the instructor. A participant noted: "I got kind of mad at [the instructor]. I thought he was way too hard on me." Two other students concurred that receiving a grade made them think that the instructor was a tough grader, a sentiment leading to escalated feeling of helplessness and worry about their future performance.

Similarly, several participants reported feeling dissatisfied with their performance and feeling embarrassed in front of the instructor: "I thought, gosh, he [the instructor] thinks I am so dumb. I really felt terrible. I felt like I let him down." Apparently, feelings of incompetence were a common reaction after receiving a low grade. It appears the majority of emotional responses described by students could be labeled as negative. Students felt angry at the instructor, ashamed for letting him down, or simply incompetent. However, the availability of detailed comments nearly eclipsed the initial strong reaction, as the participants reported working hard on improving their essays.

In those cases where the grade presented to the students was high, the participants expressed a different view on the matter. A student stated: "I got an 85 and I was satisfied so I just did some spelling corrections and that was it." Two other students agreed that getting a high grade served as an indicator of how much work needed to be done. "I knew I could make my work better, but why waste my effort if I already like my grade? So, I just made a few changes here and there and figured that it was already an improvement from before." Obviously, students were relaxed when they received a high grade, but in most cases chose not to invest a lot of effort into their seemingly satisfactory essays.

In terms of students' reactions to praise, those participants who were presented with a laudatory statement reported an overall positive attitude toward it. One of the high-scoring students indicated that receiving praise made her feel happy and enthused: "[The professor] addressed me by my first name, so I thought like he really knew me and really thought I did great. I wanted to do even better." Praise also served as a buffer for students' self-efficacy in those cases when a poor grade

was received. "I liked the praise comment. I first got mad at myself and him [the professor] for the grade, but then I thought he [the professor] gave me 68 just to push me more. I wasn't too devastated because he said I could do it."

Additionally, when discussing students' perceptions on the ideal feedback, the participants unanimously agreed that presentation of comments was the most important component. Students reasoned that individualized comments "tell you what to do," grades "tell you how much you need to do," and praise "makes you feel happy." Rephrasing the question about the ideal feedback to "What kind of feedback is ideal for your *learning*?" led to different responses. Some students noted that grades did not contribute in any way to their improvement, and therefore should be omitted. They saw a potential negative effect of both low and high grades, with the former leading to negative emotional reactions, and the latter reducing the effort they are willing to expend. Praise, on the other hand, was deemed beneficial to students' mood and attitude toward the task at hand. The participants agreed, however, that it was not crucial for their improvement.

3.2 Instructor feedback without a grade

There were nine participants in the second focus group. These students were presented with detailed descriptive feedback which they believed was provided by the course instructor. In addition to personalized comments, four of the participants were praised for their performance. Grades were not presented to anyone in this group.

This group had the highest degree of agreement in terms of their reported reactions to the feedback they received as compared to other focus groups conducted in the study. The participants were enthusiastic about the detailed comments they received, and were appreciative of the clear guidelines on how to approach their revisions. A student noted: "It was like going through a checklist. Fix this and this and you'll be fine. I knew my essay was good, and by fixing what I was told to fix I will only make it better. It was the best exam ever!" Other participants agreed, stating that the comments prompted them to stay focused on the task and helped them to feel productive at all times during the revision process.

Some students were initially alarmed by the abundance of comments, but they reported shifting into a more constructive mode soon after looking through the instructor's suggestions, as the following comment illustrates: "I had a lot of comments, so I thought, wow, I messed it up. But then I thought that whatever wasn't underlined must have been good, so as long as I do what you tell me to do, I'll do well on this exam." Other students reported feeling worried when they saw their essays with numerous corrections, but it was a transient feeling which was immediately substituted by positive emotional reactions: "I kind of, I was really confident at first about my essay, but then once I saw it, my morale dropped because I'm like, wow, this guy really chopped up my essay! But as I went along and as I finished it I was, like, now it's really-really good."

Students were pleased with the professor's commitment to their learning and were grateful for having instructions on how to proceed with their revisions. The following quotes illustrate this point: "I was amazed that he [the professor] gave me so much information… I thought, he's the best;" "I was, like, cool, he took time to

help me better my essay!" and "I never got so much feedback, it's so useful but nobody ever does it. I couldn't believe that he cared so much to do it for us." Obviously, the students responded positively to the feedback they received. They perceived the comments as the evidence of the instructor's commitment to their progress and as clear directions to what they needed to accomplish to improve their performance on the exam. These two themes emerged in the utterances of every participant of this group.

When asked about their reactions to praise, the four participants who had been presented with a laudatory statement reported feeling encouraged by it: "I thought, cool, he likes what I did and he thinks I can do better!" Students felt that praise was a professor's way to let them know that he believed they could do well on the exam. Interestingly, one student remarked: "I thought, maybe I already did great—after all he [the professor] said it—and now he just wants to push me. Maybe I didn't need to do all of it [the revisions] to get a good grade." Apparently, in this particular case praise led the student to conclude that his work may have been already good enough to receive a great score. Speculatively, praise may have depressed student's motivation to invest a lot of effort into his work on the essay.

The focus group discussed the idea of the ideal feedback. The students agreed that detailed comments were crucial for their improvement: "Tell me, like, specifically what you want, because if you tell me what you want I can give it to you." The pivotal role of comments was clear to all. Additionally, the participants noted that praise may enhance the beneficial effect of detailed feedback. Students reasoned that praise would make them feel good about themselves and therefore, would lead to even better improvement. However, they concluded by saying that praise was not nearly as important as comments: "Praise without comments is not worth much." In regards to grades, students acknowledged the grades' potential to hurt their performance and motivation. They suggested that high grades would inevitably lead to reduced effort, and low grades would be very discouraging. At the same time, they noted that there could be a potential benefit in grades. Grades may inform students of how much work needs to be done. However, if the goal is to advance understanding (in this case, to learn to write the best possible essay), grades should not be provided.

3.3 Computer feedback with a grade

The third group comprised eight students who received a grade along with detailed feedback. Students were led to believe that both the grade and comments were computer-generated. In addition to the grade and comments, four participants received a general statement praising them on their performance and encouraging them to work to improve their essay.

In regard to specific feedback, students agreed that it was very helpful. The participants noted that when they first learned that their work had been evaluated by specialized software, rather than the instructor, they were cautious about the quality of the comments. However, their opinion changed as they began incorporating the suggestions into their work. A student mentioned: "I was a bit iffy at first. But then I saw that the comments were so great and to the point. I didn't know that computers could be this sophisticated." Other students echoed by saying how amazed they were

by the level of detail provided by the machine and the relevance of feedback to their essay.

Several participants mentioned that they were relieved when they realized that the computer graded their work. A student who received a low score remarked "I thought, thank God it wasn't the professor who read it [the essay]—it was so bad! I would have been mortified if he was the one who graded it." Apparently, students perceived the computer as being non-judgmental and impersonal which helped them focus on their work without worrying about their tainted reputation in the eyes of the course instructor. Similarly, some students felt that computer could have been fairer than the instructor when assigning grades and providing comments:

> I like the computer because I know there wouldn't be any bias. You can perceive things based on someone's name or how they use their words, and with a computer I feel like I'm pretty sure it would be programmed to not be biased.

Despite the undisputed quality of the comments, seven of the eight participants felt that some suggestions provided by the computer did not apply to their work. A student remarked: "I thought the computer didn't understand what I was trying to say. It told me to change things but it didn't make any sense. It would require major rewriting. So I still kept the same idea." Another group participant continued: "Some suggestions were weird so I figured I'd ignore them." Apparently, when in doubt, students chose to ignore the computer's comments, justifying their decision with potential flaws in the software.

Students' reactions to the grade they received were consistent with the aforementioned pattern. Some participants were skeptical of the fairness of their grade. Students reported thinking that their grade did not correspond to reality and was too low. They reasoned that the software may be inappropriate for assessing the complex task of writing, so it cannot gauge more advanced structures and their underlying meaning. "I didn't take it personally. It [the grade] was too low to be true. I just kept revising and using my own reasoning to do it." Three of the participants, however, reported having different thoughts after receiving their grade. These students felt that the grade was fair because the machine was impartial when conducting assessment and generating the score. Obviously, students' views on the computer-provided grades and comments differed. Some students trusted the grade and the comments, reasoning that the machine was unbiased, whereas other students were very skeptical about the relevance of the feedback and the fairness of a grade, rationalizing that computers are not suitable for evaluating complex tasks. Skepticism of the quality of computer-generated comments in fact was apparent to varying degrees in the remarks of most participants.

Students who received praise did not report thinking seriously about it. The participants judged the encouraging comment to be "too generic," "dry," and "meaningless." They felt that the same comment was provided to every student in class, and therefore, carried no special meaning. The participants agreed that although praise may not help with their performance, it does not hurt either.

The kind of feedback that was identified by the participants as the most effective for learning was "specific comments with a grade and praise." The functions of each

were identified as follows: Detailed feedback would provide information about specific mistakes and ways to correct them. Grades would tell students how much work needs to be done. Praise was deemed as having no particular purpose, but was judged as having a potential to elicit positive emotion. Some participant noted that they would only want to receive a grade if it is very high. In that case, they would not be discouraged, but at the same time would not be as motivated to exert additional effort and work on improving their work. Finally, most students stated that they would prefer receiving feedback from the instructor, as opposed to the computer. The instructor's feedback would naturally be more personal, accurate, and trustworthy, and therefore would be most effective in ensuring progress of learning.

3.4 Computer feedback without a grade

The fourth focus group comprised eight students who received detailed comments with the understanding that the feedback was computer-generated. In addition to comments, four participants had been praised on their performance. Grades were not presented to anyone in this group.

Group participants reported feeling pleased with the comments they received. Students agreed that clear guidelines provided by the computer helped them during the review process. Not only were the comments instrumental in helping students make corrections, they also helped students concentrate on the task at hand, making them feel more confident that they could succeed. Students mentioned feeling relieved that there was a chance to rework their essay to ultimately get a higher grade. "I thought it was great. It gave me a chance to review. I basically looked them [the comments] over and knew I needed to revise my paper more and get a good grade. The comments were so to-the-point!" In regards to the credibility of the computer comments, many students remarked having original doubts but later finding the feedback to be useful and relevant. However, students admitted encountering comments that they felt did not apply, as the following statement illustrates: "The comments were great. It was so interesting that the machine is so smart. It was very useful, but some comments didn't really work, so I figured, oh, well, it's a machine after all. I should decide what to use." Other participants confirmed feeling that in some cases the computer was not quite "understanding" what they were trying to convey. Hence, they ignored some comments but incorporated those that they felt were relevant.

The participants who were praised on their performance were in accord in terms of their reactions to the praise comment. Students perceived praise as a machine-generated generic statement, as shown by the following remark: "I thought it was a comment that everyone got. The computer isn't warm and fuzzy, so I didn't think my essay was special. I just thought it was a general statement." Those students who were not presented with praise said that they would not have wanted to receive it. The participants felt that praise has a minimal, if any, value in their achievement. At the same time, students unanimously agreed that if presented by a person, praise may be useful for enhancing positive affect and increasing motivation. Still, they noted that personalized comments were far more important for ensuring success, with praise serving as a "feel-good factor."

One of the students stated categorically, with universal head-nodding in agreement, that grades are unnecessary if a chance to review work is offered.

Students explained that a grade is perceived as a "final step" and therefore is not conducive to motivating additional work on the assignment. "If I got a grade, I would be, like, OK, that's what I got. If you gave me an 85, I would look at the comments, but with my busy life, I would probably be, alright, I am satisfied with the grade, why stress about it?" Other student reaffirmed and expanded: "If it [the computer] gave me a 65, I would have panicked, but I probably would have put a lot in there. But if you want to give me suggestions, don't make me freak out with the low grade. I don't see the point." Obviously, students felt that personalized comments were sufficient and necessary for improvement. Grades, however, were perceived as having a potential to deplete effort and elevate anxiety, and were judged to be undesirable.

Students' views on the ideal feedback were similar to those expressed by participants of other focus groups. Students emphasized the pivotal role that detailed comments play for their improvement. Computer comments were deemed to be very useful. However, students believed that the instructor's comments would be more trusted and, as a result, would lead to higher improvement. In regard to grades, students pronounced in unison that the only value of a grade is to reduce the amount of work in case of a high score. Praise was said to be valuable for elevating mood, but not particularly effective for enhancing performance. The following comment summarizes the general attitude toward the ideal feedback:

> A chance to revise is so fair. I think I would like to get comments on every paper I write. Even when I saw that my paper was obliterated with marks on it and stuff, I still thought it was great. I could improve it. A grade would have freaked me out because I saw that much info… Praise is nice, but I don't care. I just want you to tell me what I need to do. This is a learning experience, after all.

3.5 No detailed feedback, grade

The eight students who comprised this group had not been given any specific feedback on their work. All of the participants were presented with a grade, and four students were also provided with a laudatory statement.

The students' reactions to the feedback were not nearly as optimistic as those of the participants in the previous groups. "It was really bad," "I panicked," "I was shocked" were among the common remarks that participants shared. Students who received unsatisfactory grades without any guidance on how to improve their work reported feeing frustrated, as the following comment illustrates: "I worked really hard, and I got a 75. I was completely freaked out. You know it's low, you want to do something better but you didn't know what, so you didn't. I just moved some stuff around and left." Students felt helpless and craved any form of specific instructions. The majority of students reported not being able to considerably revise their essays. Rather, they claimed to have made minor adjustments to their work, due to a lack of information on their errors and ways to improve them. "I just corrected spelling. I didn't know what else I could possibly say," noted one of the participants.

Two of the participants received a high grade on their essay. They had a different reaction to the feedback as compared to their peers who got a lower grade. One

student explained that "it was great to get an 85. I was satisfied so I resubmitted my essay without doing pretty much anything." The other high-scorer echoed, "I was glad to have been shown my grade. I was like, great, I did well, I don't need to revise." The students admitted, however, that if their grade had not been high enough to be considered satisfactory they would have been discouraged by it.

Four of the group participants received praise in addition to a grade, and their reactions to it varied. Two participants who had been praised on their performance did not find the comment to be useful. A student thought that praise was "a way of sugar coating what's coming your way. I got a 77 and then I got good praise, yeah, I didn't like that, it was weird." Another student agreed, noting, "I had a 73 and I didn't think he [the professor] was serious. I was just surprised to hear him say "you had a good start." It didn't really seem to reflect reality." Other participants had a different view on the matter: "I felt like, oh, I did a great job! I'm on the right track! So I didn't totally fail this." Among those students who did not receive praise, the opinions had a similar split. Some students wished they had gotten an encouragement, reasoning that it would have made them feel more confident about their performance, whereas other students said that praise would not have been instrumental. A student mentioned, "I didn't even want positive comments so much, just some sort of pointing out where the weak points are would be really helpful." This comment reflects this group's view on the feedback they consider being the most effective.

The participants agreed that having detailed comments would be of a great assistance, as "telling me exactly what to do is the best thing a professor can do." Students repeated that a grade on its own is not helpful. However, some of the participants suggested that a combination of a grade with the comments would constitute the ideal feedback:

> I think if it was comments and a grade it would have been the best. If I got a low grade, I would have appreciated it kind of, because I would be, all right, this is the grade I got, I need to improve it, definitely focus on the things they said I should improve upon.

However, the student mentioned that if he had gotten a higher grade, he "would be more hesitant to change anything," afraid to "make it worse." With the further exploration, the group derived certain contingencies associated with the presentation of a grade: "I think it would just depend, like, if someone got an 80 they'd probably know all the stuff they just don't have it all together. And then if they got that grade they can say "Ok, I just got to go back and put it all together." If someone gets a 60, that just means they didn't know what they were talking about, so the grade would make them get mad or panicked. If someone gets a 90, they're going to say "I'm not going to change anything." If it was me I wouldn't." Thus, students felt that in most situations a grade will tend to have a harmful effect on performance. The participants further elaborated by saying that for those courses that they were interested in and wanted to gain competence, grades would be discouraging. Conversely, for those courses that they dislike, a grade would be desired because it "just saves your effort." A student proposed that "if it's a passing grade, I would just resubmit it [the essay]. I wouldn't care even if I could improve." Students explained that "not

knowing a grade in this case would be annoying, because I would be afraid that I failed, and would be forced to revise."

Students' comments indicate that they consider the usefulness of grades from two perspectives. On the one hand, if their goal is to get through the class with the minimal effort, a grade is judged to be helpful for telling them how much work they need to do. On the other hand, if the goal is to enhance understanding and improve learning, grades either make them nervous or prevent from investing a lot of effort. Thus, even if students are motivated to learn, they feel that grades may add an unnecessary constraint which may take their mind away from learning, and focus on energy and effort conservation: "Even if you want to learn more, you'd be like, why do it if you already have a 90." One of the participants voiced an opinion supported by all of the others, saying that when there is a choice between receiving no feedback at all, and receiving a grade, the latter alternative is definitely more appealing.

In regard to praise, students felt that praise would be a pleasant addition to the comments, but not a very useful form of feedback when presented alone. There was general agreement that more personalized and work-specific praise may be beneficial for students' motivation and a general sense of well-being. It was also noted that praise may soften the negative effect of a low grade by canceling out or mitigating the negative emotions which usually follow an unsatisfactory performance. However, when compared to the importance of comments, its value was deemed to be quite limited.

3.6 No detailed feedback, no grade

The final group was conducted with eight students, four of whom received praise in the course of the experiment, and four who received no feedback of any kind. Interestingly, students' reaction to the absence of feedback was not overwhelmingly negative. One of the focus group participants reported feeling grateful to receive a chance to revise his essay, as illustrated by the following remark: "I wish I could have gotten some feedback. Still, the second time I regrouped and remembered everything a little better. Revising really helps." Other students agreed, but added that it was frustrating not to have any guidance on what to change. They reported working on "whatever came to mind," "correcting stylistic and spelling errors," but not working so much on the content of their essay. Students felt that changing the content may lower their final score, so they chose not to make considerable revisions.

One of the participants interpreted the lack of any feedback as a clear indicator of the high quality of her work. The student commented: "I thought mine was really good. I had almost everything, like all the information that he [the professor] gave us in the lectures. I only deleted one sentence and added another, and left." Quite surprisingly, those students who were praised on their performance had stronger negative reactions than those who did not receive anything. A student noted: "It was a joke, like, "you made a great start, let's try to make it better," and nothing else! I didn't know how to make it better!" Another student echoed: "I thought it was just totally generic. It didn't help. It made me get more nervous, because I had no idea

what to change." Obviously, the lack of clear instructions on how to improve work was very frustrating to these students. They knew they were expected to revise their work but were not provided with any guidance on how to proceed with this task. Despite their frustration, students who received laudatory comments reported that they spent more time on revisions than their counterparts who received no feedback.

The participants of this group unanimously concluded that the ideal feedback should include specific comments. They felt that the chance to revise was useful, but that they lacked one important component—feedback. A participant commented: "If I'm going to get a chance to revise and get it back, I like to have some sort of idea of what I did wrong and what I should do to fix it." Students stated that if the detailed comments were impossible to be compiled, any form of feedback would be conducive to improvement. "Any information is better than no information," one of the participants shrewdly noted.

In regard to a grade, students felt that neither numeric scores nor letter grades were effective in promoting improvement, reasoning that "it makes you too anxious to actually know your grade." The participants agreed that "it's better just to critique your ideas, like, this is what you need to fix. If they give you a low grade it'll make you nervous. If you get a high grade, it's like I don't want to fix it." Similarly to the previous groups, the participants of this group were clear about their view on the potential negative effect of grades.

Additionally, students were in agreement that personalized comments represented the ideal form of feedback and were believed to lead to the best improvement. They also hypothesized that when praise was added to detailed feedback it enhanced the beneficial effect of the feedback. However, as one of the participants noted, "I don't care about praise. I want directions. Tell me what to do, and I will." This statement reflects the significance students assigned to constructive feedback as opposed to evaluative forms of feedback.

4 Synthesis of general themes that emerged in the focus group discussions

The analysis of the focus group discussions allowed for construction of a general picture of the personal experiences of the participants who had received different forms of feedback in the course of the experiment. Not surprisingly, students' responses differed both within each group and among the six groups.

The participants who received detailed feedback from both the instructor and the computer emphasized the usefulness of comments for their performance. Students agreed that information about errors they committed and suggestions on how to correct them was necessary for improvement. For the instructor group, students felt that in addition to assisting them in the revision process, feedback was an indicator of the instructor's commitment to their progress. As a result, they were eager to make the adjustments and improve their work. Computer-generated feedback was appreciated by the focus group participants and was generally deemed to be relevant and helpful. Interestingly, however, students in the computer group unanimously agreed that some of the comments they received did not apply

to their work or were too outlandish to consider. Those in the instructor condition received essentially the same comments, but did not feel this way. This indicates that students judged the quality of the feedback based at least partially on its source.

Students who were not presented with detailed feedback reported feeling discouraged when they received a grade that they were not satisfied with, whether or not they received praise as well. They noted that having no guidance as to how to approach the task of making revisions was frustrating and often led to elevated negative affect. Conversely, students who received only a high grade were content and did not spend time trying to improve their work. Some of the students who did not receive any feedback at all interpreted the lack of comments as an indication of the high quality of their work. As a result, they chose not to revise their essays. Praise statements were deemed to be more useful than grades, for students who received an encouragement to improve their work knew they were expected to revise their essay. The general conclusion inferred by the students was that any feedback is better than no feedback.

Students' reactions to grades were different for students across the six focus groups. Participants who received a grade from the instructor reported feeling angry when their score was low, but said that they still worked hard to incorporate the instructor's comments, thus improving their essays. Students whose grade was high enough to match their own standards admitted spending very little time considering the instructor's comments. In the computer condition, low grades were received with skepticism by the majority of the participants. They were far less upset by their low score when it came from a computer, reasoning that the software was not capable of understanding the logic of their arguments, so the grade did not reflect reality. High grades, however, were thought to be fair so significant revisions were not made. The strongest negative reaction to the grade was reported by the students who did not receive anything in addition to the grade. In this case, they felt angry and helpless, not knowing what went wrong or how to improve it.

Praise elicited the most diverse responses from students. Under the instructor condition, the majority of students reacted positively to praise. They reported feeling happy and were encouraged to perform well on the exam. However, computer praise was dismissed by the majority of the focus group participants, though they said if it were made more personal, there could be a potential benefit to praise from a computer as a vehicle to increase their mood and motivation. When combined with the grade, praise appeared to have mitigated some of the negative effects of the grade. Students who received praise in addition to a grade felt less frustrated than did their counterparts who were presented with only a grade. The latter group wished they had received a laudatory comment. General themes that emerged in the focus group discussions are listed in Table 1.

Students in all six focus groups concurred on their definitions of the ideal feedback. The participants stressed the importance of detailed comments for their improvement and said that it was the form of feedback they preferred. A grade was deemed to be unnecessary when the goal is to ensure progress in learning. In every group, students said that receiving a high grade inevitably leads to decreased effort, since there is little room for improvement, as is receiving a low grade because it is

Table 1 Prevalent themes of focus group discussions

Theme	Instructor grade	Instructor no grade	Computer grade	Computer no grade	Grade only	No feedback
Detailed feedback is useful, shows exactly what to do	√	√	√	√		
Some comments do not apply			√	√		
Low grades cause frustration	√				√	
High grades lead to reduced effort	√		√		√	
Personal praise is encouraging, elevates mood	√	√	√	√	√	
Praise softens the effect of grade	√				√	
Praise is useless			√	√		√

discouraging. In regard to praise, students commented that it is a pleasant addition to constructive feedback, elevating morale and confidence. However, its role in students' performance was not considered to be critical.

5 Conclusions and practical implications

The purpose of the study was to uncover potential reasons behind the effects of differential feedback messages and to obtain students' perspectives on what they construe as the most valuable information. The main findings were as follows: Students unanimously stressed that detailed comments were the most effective form of feedback. Grades were perceived as potential obstacles to student improvement, especially by those in the instructor condition. Computer feedback was viewed as unbiased. Unfavorable comments and grades provided by the computer were dismissed as irrelevant. Praise was considered the least influential type of feedback, useful only to soften the demotivating effect of grades and to provide a general sense of well-being.

It seems apparent that providing specific personalized information about individuals' work and allowing them to make changes based upon this information leads to improvement of their performance and makes them motivated to work on perfecting their work. In the educational system, knowing techniques that consistently work and lead to the best progress is a luxury. If detailed neutral comments represent one of the most powerful of these techniques, educators should put substantial effort into providing them, and cultivate the resources to make it possible (Schute 2007). Developing such comments may be an onerous task. However, if the goal of education is to advance students' progress, this practice appears to clearly be worthwhile.

Furthermore, it appears that as long as the feedback message encourages "mindfulness" in students' responses (Bangert-Drowns et al. 1991), students will treat computers as equals to humans and will use computer feedback to improve their work. This conclusion is consistent with the "computers as neutral tools"

(CASA) perspective currently popular in the literature that states that people would treat computers as unbiased sources of information (Ferdig and Mishra 2004; Mishra 2006; Nass et al. 1999). However, skepticism articulated during the focus group discussions along with the different patterns of responses for computer and instructor conditions indicated that students do not treat human and machine generated feedback as the same. In fact, quite contrary to the CASA perspective, students felt that in some cases, computer-provided suggestions for improvement were faulty and irrelevant.

Although it is hard to disagree with the convenience and effectiveness of grades when used for summative purposes (Airasian 1994; Marzano 2000; Smith and Gorard 2005), the formative function of grades as tools that lead to progress in learning is quite dubious (Black and Wiliam 1998; Shute 2007). Students unanimously stated that grades are ineffective when mastery of learning is needed. In some educational settings, however, presenting a grade is a requirement. As a result, figuring out ways to do so with the least damage to students' achievement and the least impact on their self-efficacy and motivation is crucial for educators across all academic environments (Schute 2007; Tolli and Schmidt 2008).

Similarly, praise was deemed to be a pleasant but useless addition to the feedback. If educators have an option to choose, personalized comments without praise or grade appear to be presented as an optimal form of feedback leading to the highest increase in students' learning when a task is going to receive additional work by a student (Black and Wiliam 1998; Henderlong and Lepper 2002). Consider this illustrative comment that a focus group participant made: "Just comments, tell me what I did wrong, where I could change it. Just comments and error marks."

A quote of one of the focus group participants summarizes the discussion:

> I wish I had all tests in high school and college like this [referring to the essay exam] because they didn't give you any feedback. I always hated grades because they don't tell you anything. I'm sitting there [at the exam] thinking this is great; I can fix anything I messed up on. I can make it [the essay] better… So I don't think grades or praise would be good. Just comments, tell me what I did wrong, where I could change it. Just comments and error marks.

This statement captures the findings of the study with uncanny precision.

References

Airasian, P. W. (1994). *Classroom assessment*. New York: McGraw-Hill.

Bangert-Drowns, R. L., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*, 213–238.

Baumeister, R. F., Hutton, D. G., & Cairns, K. J. (1990). Negative effects of praise on skilled performance. *Basic and Applied Social Psychology*, *11*, 131–148.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*, 7–68.

Butler, R. (1988). Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, *58*, 1–14.

Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, *78*, 210–216.

Cameron, J., & Pierce, D. P. (1994). Reinforcement, reward, and intrinsic motivation: a meta-analysis. *Review of Educational Research*, *64*, 363–423.

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

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*, 627–668.

Dev, P. C. (1997). Intrinsic motivation and academic achievement: what does their relationship imply for the classroom teacher. *Remedial and Special Education*, *18*, 12–19.

Earley, P. C. (1988). Computer-generated performance feedback in the Subscription-processing industry. *Organizational Behavior and Human Decision Processes*, *41*, 50–64.

Elawar, M. C., & Corno, L. (1985). A factorial experiment in teachers' written feedback on student homework: changing teacher behavior a little rather than a lot. *Journal of Educational Psychology*, *77*, 162–173.

Ferdig, R. E., & Mishra, P. (2004). Emotional responses to computers: experiences in unfairness, anger and spite. *Journal of Educational Multimedia and Hypertext*, *13*, 143–161.

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: an experimental and individual difference investigation. *Journal of Personality and Social Psychology*, *52*, 890–898.

Guskey, T., & Bailey, J. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks: Crown.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81–113.

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: a review and synthesis. *Psychological Bulletin*, *128*, 774–795.

Kanouse, D. E., Gumpert, P., & Canavan-Gumpert, D. (1981). The semantics of praise. In J. H. Harvey, W. Ickes & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 3, pp. 97–115). Hillsdale: Erlbaum.

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

Lajoie, S., & Derry, S. (1993). *Computers as cognitive tools*. Hillsdale: Earlbaum.

Landauer, T. K., Latham, D., & Foltz, P. (2003). Automatic essay assessment. Assessment in Education, 10.

Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 75–106). Hillsdale: Earlbaum.

Lepper, M. R., Henderlong, J., & Gingras, I. (1999). Understanding the effects of extrinsic rewards on intrinsic motivation-uses and abuses of meta-analysis: comment on Deci, Koestner, and Ryan (1999). *Psychological Bulletin*, *125*, 669–676.

Lipnevich, A. A., & Smith, J. K. (2008). *The effects of differential feedback on students' performance*. Princeton, NJ: Educational Testing Service.

Marzano, R. (2000). *Transforming classroom grading*. Alexandria: Association for Supervision and Curriculum and Development.

Mishra, P. (2006). Affective feedback from computers and its effect on perceived ability and affect: a test of the computers as social actors hypothesis. *Journal of Educational Multimedia and Hypermedia*, *15*, 107–131.

Mishra, P., Nicholson, M., & Wojcikiewicz, S. (2001). Does my wordprocessor have a personality? Topffer's law and educational technology. *Journal of Adolescent and Adult Literacy*, *44*, 634–641.

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, *45*, 669–678.

Nass, C., Moon, Y., & Carney, P. (1999). Are respondents polite to computers? Social desirability and direct responses to computers. *Journal of Applied Social Psychology*, *29*, 1093–1110.

Oosterhof, A. (2001). *Classroom applications of educational measurement*. Upper Saddle River: Merrill Prentice Hall.

Orrell, J. (2006). Feedback on learning achievement: rhetoric and reality. *Teaching in Higher Education*, *11*, 441–456.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research and applications*. Upper Saddle River: Pearson Prentice Hall.

Schute, V. J. (2007). *Focus on formative feedback*. Princeton: Educational Testing Service. RR-07-11.

Shanab, M. E., Peterson, D., Dargahi, S., & Deroian, P. (1981). The effects of positive and negative verbal feedback on the intrinsic motivation of male and female subjects. *Journal of Social Psychology*, *115*, 195–205.

Shute, V. J. (2007). Tensions, trends, tools, and technologies: Time for an educational sea change. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning*. New York: Lawrence Erlbaum Associates, Taylor & Francis Group.

Smith, E., & Gorard, S. (2005). They don't give us our marks': the role of formative feedback in student progress. *Assessment in Education Principles Policy & Practice*, *12*, 21–38.

Tolli, A. P., & Schmidt, A. M. (2008). The role of feedback, causal attributions, and self-efficacy in goal revision. *Journal of Applied Psychology*, *93*, 692–701.

Torrance, H. (1993). Formative assessment: Some theoretical problems and empirical questions. *Cambridge Journal of Education*, *23*.

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