RESEARCH ONLINE

University of Wollongong
Research Online

Faculty of Creative Arts - Papers (Archive)

Faculty of Law, Humanities and the Arts

2010

Peer assessment in tertiary level singing: changing and shaping culture through social interaction

Lotte Latukefu
University of Wollongong, latukefu@uow.edu.au

Publication Details

Latukefu, L. (2010). Peer assessment in tertiary level singing: changing and shaping culture through social interaction. Research Studies in Music Education, 32(2), 61-73.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

RESEARCH ONLINE

Article published as: Latukefu, L. (2010). Peer assessment in tertiary level singing: changing and shaping culture through social interaction. Research Studies in Music Education, 32(2) pp.61-73

Peer assessment in tertiary level singing: changing and shaping culture through social interaction

Lotte Latukefu

Abstract

In 2008, peer assessment was introduced into the singing component of a tertiary level undergraduate creative arts performance course within an Australian regional university. The study investigated what effect changing the role of the actor/singer in an assessment has on the culture of the course as well as individual development of graduate qualities, such as critical thinking and responsibility. It also looked at what process was involved in order to integrate peer assessment into the subject, and what kind of support was needed to achieve this. Results suggested that students saw themselves as agents of their own assessment activities by taking control of assessment, and that having to think critically about other student performances made them reflect on how effective their own performances were.

KEYWORDS: peer assessment; vocal pedagogy; reflection; self-regulation

Introduction

In 2008, peer assessment was introduced into the singing component of the undergraduate performance course within an Australian regional university. The purpose of this exercise was to encourage students to become better self-regulated learners, who would be capable of continuing with their learning after graduation. In an article that gathered together the main concerns being addressed in studies on self-regulated learning, Montalvo and Torres (2004) found that self-regulated learners "see themselves as agents of their own behaviour, they believe learning is a proactive process, they are self-motivated and they use strategies that enable them to achieve desired academic results" (p. 4). Montalvo and Torres also point out that these characteristics coincide with those attributed to high-performance, high capacity students, but that with adequate training all students can improve their control over learning and performance. In general, students who self-regulate their learning "show greater effort to participate in the control and regulation of academic tasks, classroom climate and structure" (Montalvo & Torres, 2004, p. 3).

The present study on peer assessment was part of an ongoing study to develop a model of learning singing underpinned by socio-cultural theories. In undertaking a Vygotskian approach, the singing class environment was specifically designed to encourage self-regulated learners who learn from social interaction with each other (Latukefu, 2009). Falchikov (2007) has argued that peer involvement in assessment has the potential to encourage learning and develop assessment skills that will last a lifetime. She also states however that peer assessment without modeling or scaffolding has no value added to the student learning, and that if students are merely completing an exercise without understanding the standards or criteria which will help them acquire skills in judgment they are no better off than in the framework of traditional assessment.

Literature review

The present study began with the pedagogical goal of developing in singing students, the ability to critically discern quality (Sadler, 2008). The university that provided the context for the study requires that students be given criteria for all assessment tasks. The breaking down of holistic judgments into

components is supposed to make the process of assessment more transparent for the students. Sadler points out that holistic rubrics use extended verbal descriptions to set out characteristics rather than breaking them into components. He made several observations on the use of analytic versus holistic rubrics, one of which included the discrepancy between global and analytical appraisals where a work or performance judged as brilliant may not necessarily rate outstandingly on each criterion.

A study in music assessment (Stanley, Brooker, & Gilbert, 2002) also supported an argument for developing descriptors of quality with students rather than a conventional analytic rubric. Stanley et al. (2002) reviewed assessment procedures carried out at the Sydney Conservatorium of Music in 1995 and reported that most assessors initially adopted a holistic rather than an analytic approach to assessment of performances. They relied on an initial gut response as to whether or not they were enjoying the performance. They then went through a process of justification in order to identify what characteristics justified the feeling they had initially. They also reported filling out specific criteria at the end of the performance because over the course of a few pieces this might have to be radically adjusted or because they used it as a way to justify their gut feeling that the performance was worth a certain grade. One respondent commented that "you can get a kid that plays out of tune and out of time but you are crying because it is so expressive or so wonderful. You can (also) get a kid that plays dead in tune or dead in time and absolutely immaculate dynamics that leaves you totally cold" (Stanley, Brooker, & Gilbert, 2002, p. 52). This notion of judging performances on their total impact rather than components was important in the development of the research questions developed for this present study.

Another pedagogical goal of the present study was to encourage students to develop graduate qualities such as critical thinking, reflection, and responsibility. The contribution of peer assessment in developing these qualities is supported in the literature. Searby and Ewers (1997) concluded that having to work out what criteria of assessment were to be employed and then having to apply them to real work focused the students' minds on what made the work good or bad. Blom and Poole (2004) concluded that "peer assessment in a tertiary performance programme offers a relevant

and meaningful context for deep learning about performance assessment and performing to occur" (p. 125).

A review of the literature, and analysis of current practice, found that there were relatively few music institutions that had formal peer assessment as part of their programs. Those that did (Blom & Poole, 2004; Daniel, 2004; Hunter & Russ, 1996; Searby & Ewers, 1997) were positive about the learning outcomes for students, but held reservations about the process. These reservations included: over-marking by students, extra workloads for both students and teachers, problems arising when different instruments and genres are involved, and lack of readiness on the part of students to take part in the exercise. Searby and Ewers (1997) discussed the aspects that contributed towards the overall effectiveness of the system and issues and experiences that arose from the implementation of peer assessment at Kingston University. Arriving at a correct mark was a source of worry for students who were new to peer assessment, who felt they were not qualified to make judgments on other students' work, and were reluctant to get involved. Students were expected to provide a detailed report as feedback on the performance. A minority of lecturers complained about the quality of these reports. An aspect that the researchers at Kingston University felt to be important was to make sure that students learnt as much as possible from the process.

Engaging students in the process of integrating peer assessment into the singing subject was considered central to the present research. The importance of student participation in the process of developing assessment criteria was a consistent theme in the literature on peer assessment in music courses (Blom & Poole, 2004; Daniel, 2004; Hunter & Russ, 1996; Searby & Ewers, 1997).

Researchers at Kingston University created a set of generic guidelines in the initial implementation of peer assessment. These included:

- The establishment of a training scheme in peer assessing for students;
- The establishment of a set of criteria of assessment, which would be negotiated with the students concerned (a vital part of the process);

- The establishment of clear and effective administrative systems (Searby & Ewers, 1997, p. 372)

This set of guidelines was used by the author/researcher as the starting point for the design and implementation of peer assessment into the singing subject.

The present project investigated the following questions:

- What effect does changing the role of the singer in an assessment have on the culture of the course and graduate qualities,[1] such as reflection, critical thinking, and responsibility?
- What is the process of integrating peer assessment in a tertiary level singing subject, and what kind of support is needed to achieve this?
- Does constant interaction with descriptors of quality lead to what Sadler (2008) describes as "the creation of environments in which the critical discernment of quality becomes a key aspect of learning, drawing on what is known about connoisseurship in other contexts" (p. 18)?

Method

To capture the developmental nature of this pedagogical project and the context in which it was carried out, a design-based research methodology was employed. There are many different permutations of design-based research methodology (Bannan-Ritland, 2003; Barab & Squire, 2004; Design-Based Research Collective, 2003; Gravemeijer & Cobb, 2006; Lenski, 2001; McKenny, Nieveen, & van den Akker, 2006; Reeves, 2000; van den Akker, 1999). Wang and Hannafin (2005) define design-based research methodology as "a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation,

[1] Graduate qualities are developed by Australian universities to describe the distinctive qualities of a graduate of that university. They are used to guide staff who are engaged in curriculum development and help students to develop personally and professionally. (University of Wollongong Graduate Qualities Policy, available from: http://www.uow.edu.au/about/policy/UOW058682.html#P73_1243)

based on collaboration among researchers and practitioners in real-world settings and leading to contextually-sensitive design principles and theories" (pp. 6–7). Design-based research methodology leads to a development of knowledge that can be used not only in practice, but to inform other practitioners.

Research site and participants

The singing program is currently taught as part of actor skills training in a creative arts degree program at an Australian regional university. The participants are undergraduate acting students aged around 18–21 years of age. Originally singing had been part of a music degree and singers were trained classically, however the music and drama departments were amalgamated in 2002 and the focus changed to contemporary vocal practice. At the end of first year in the degree, all singing students are given an opportunity to audition for a specialist singing class. The focus of this class is on classical singing training combined with contemporary performance practice.

Collaboration between researcher and students

There were two phases in the project. The purpose of the first phase was to gather together a focus group of six students using purposive sampling to ensure that diversity of the student population was represented in the sample. Gender equality was a factor in the purposive sampling and I also wanted to ensure there was representation from Indigenous and African students who were in the course at the time. Finally I tried to get an equal mix of specialist and non-specialist singers in the focus group. This reflected the stage of "collaboration among researchers and practitioners in real-world settings" in development research methodology (Wang & Hannafin, 2005). The aim of the focus groups was to gather as much information as possible from the students about criteria they thought important for high quality singing. The six students attended two focus groups in which they discussed the best process for implementing peer assessment into singing classes and how to solve possible problems that might arise during the exercise. The purpose of the second phase was to implement the design into the singing course in order to include the rest of the student population in the project. The design of the peer assessment task was implemented in singing classes and data about the student experience

of the exercise collected through a survey sent out to students, reflective journals that students are required to keep, and the in-class notes from lecturers. Finally, a systematic documentation, analysis, and reflection on the research process and outcomes were carried out (Wang & Hannafin, 2005).

Implementation

Some of the issues raised by the students in the focus group in relation to peer assessment were as follows:

- How to deal with the different standards of skill levels among students.
- How to account for different tastes or reactions to an individual's style and genre of music.
- Should assessment be relative to an individual's improvement or compare students to see who is best?
- Over marking of friends.
- How many should be on a panel and the practicalities of writing up reports.
- Privacy and confidentiality issues: how to respond if the student wants information about the panel's decision.
- Over professionalism: being too critical to demonstrate how professional the panel member is.

All these points were taken into account when working with the students on peer assessment.

In the second focus group, the students discussed what they thought would make the peer assessment a good process. They decided on the following points when developing descriptors of quality, which were then implemented by the lecturers in class.

- Have criteria clearly laid out and printed and ensure each student has a copy.
- Give a good introduction and encourage discussion when criteria are handed out.

- Explain the reasons for introducing peer assessment, and, in particular, emphasize the idea of learning to critique as a skill development for future work.
- Reiterate all this in the first 4 weeks of classes.
- Give examples in class of different standards of achievement of the performance qualities.
- Have the lecturer critique the critiques of the first 4 weeks giving particular attention to the critiquing skills of students.
- Reinforce the notion that the qualities assessed using peer assessment are those that students are already being assessed on and taught.

There were 15 students in 3rd Year and 20 students in 2nd Year participating in the peer assessment exercise, and all agreed to participate in the study. In week 5, the 2nd and 3rd Year students formed panels and the three students in the panel assessed another student in the class. The panels discussed the performance and agreed on a mark. The panel also provided written feedback to the performer.

Fading of support (Falchikov, 2007) is where the lecturer gives support to the class through modeling or directing students, but slowly withdraws this level of support and involvement over time. In the case of the present study, the lecturer gave the students lots of prompting in the first 2 weeks in order to help them with critiquing each other, but then began to withdraw from the discussion by week 3. This is consistent with the socio-cultural approach used in the study and with the concept of scaffolding in particular. Scaffolding is "a change of quality of support over a teaching session, in which a more skilled partner adjusts the assistance he or she provides to fit the child's current level of performance. More support is offered when a task is new; less is provided as the child's competence increases, therefore fostering the child's autonomy and independent mastering" (Berk & Winsler, 1995, p. 171).

After the implementation of the peer assessment exercise, all the students who took part were sent an open-ended questionnaire (see Figure 1) inquiring about their experience and perception of peer assessment. This feedback allowed for iterative analysis of the developed peer assessment

process. A majority of students (30 out of the 40 students) who took part in the exercise responded to the questionnaire. Some of the questions were based on questions that had been used to evaluate a peer assessment exercise carried out by Falchikov (1995).

Figure 1: Peer Assessment Questionnaire

1. **What did you like best about the peer assessment exercise? Why?**
2. **What did you like least about the peer assessment exercise? Why?**
3. **Peer assessment makes me:**
    a. **Think critically- Strongly agree/agree/strongly disagree**
    b. **Feel a sense of responsibility- Strongly agree/agree/strongly disagree**
4. **Could you comment on how your personal knowledge about the student you were assessing affected your judgment?**
5. **Could you comment on how this peer assessment exercise may have affected your own learning?**

Descriptors of quality

The descriptors of quality developed in collaboration with the students in the present study (see Table 1) used the concept of a holistic rubric to adapt a model developed by staff at the Queensland University of Technology (Thomas & Millard, 2006). Different levels of technical and musical interpretation were expected in the different years, and this meant that two different descriptors of quality were constructed. In the pilot study, students were provided with these descriptors of quality and the instruction to add any others they saw emerge from the performance and mark the performance on the overall impression. This second approach was recommended by Sadler (2008).

Table 1: Descriptors of quality adapted from model developed by staff at the Queensland University of Technology (Thomas & Millard, 2006) 2nd and 3rd Year peer assessment criteria

Technical achievement	Interpretative skills	Professional skills	Qualitative judgments to think about
Anchoring- ability to anchor in shoulders and back constantly while singing **Good Posture** **Silent intake of breath and good airflow** **Energised**- posture is dynamic, gestures are made energetically and vocal tone is vibrant. **Intelligibility**- vowels are well formed and resonant **Sob-** attempt at sob **Twang**- able to incorporate some twang into sound in middle register (in all registers for 3rd Year) **Release of constriction-** ability to release constriction on short phrases and in low to middle register (on long phrases and throughout register for 3rd Year) **Accuracy**- ability to generally sing in tune and rhythmically	**Communicate with audience-** ability to communicate through body language and vocal colour **Ability to affect audience** through imagination and thought process **Appropriate stylistic choices** **Extra criteria for 3rd Year included** Deep involvement with music and commitment to communication with audience	**Memorization** is complete and reliable **Establishes a relationship** with audience and accompanist **Performer is physically and musically prepared** for the performance	Performance is **compelling and forceful** Performance is **sophisticated and commanding in presentation** Performance is **thoughtful and engaging** Performance is **technically well executed, but bland and unimaginative** Performance is **under-prepared and lacks skill** Performance is **unsatisfying and musically unconvincing**

The need to assure that all the students were familiar with all the descriptors was important, as there was not much time in each assessment for students to be re-familiarising themselves with the

descriptors. Findings by Stanley et al. (2002) suggested emphasis needed to be placed in examiner training and in order to do this the descriptors were introduced into all classes from week 1 and students were encouraged to frame their informal critiques using the descriptors as a guide. They had been using the same technical language throughout 1st Year, but this was the first time they were asked to constantly interact with the criteria, with the objective of preparing to assess another student's work. There were two singing teachers, one of whom was also the researcher, and it was important for me as researcher and teacher to discuss all aspects of the peer assessment exercise and the phases of implementation with the other teacher.

Results

Critical thinking

A survey was sent out as part of the iterative analysis. It was concerned with exploring student experience of the peer assessment exercise. Students were asked what they liked best about the peer assessment exercise and by far the most common answer was that critically assessing someone else in performance actually helped improve their own performance. Some of the responses from the students to this question were:

> "By assessing my classmates I found that during my assessment I was thinking critically and could therefore work to apply the things I had noticed lacking in previous assessments."

> "I liked the ability to be able to discuss as a panel why and how the performance of the singer worked. Playing the assessor gave me an understanding what are the standards and criteria I need to full fill (*sic*) to be able to perform well in my own performance."

> "Talking with the rest of the panel was really good for solidifying ideas of what to observe for technique. Also because of the detailed criteria everyone put a lot more effort into preparing for the assessment because we knew what we would be judged on."

> "I got a chance to put myself on the other side of the table. The judging side and see what it is that judges view as important in a performance which helps me reflect on what I need to work on."

> "The exercise enabled me to critically evaluate my peers, which is something which is not done very often."

The majority of students, 83.3% (n = 25) strongly agreed and 16.7% (n = 5) agreed that peer assessment made them think more critically. When asked whether the peer assessment exercise gave them a sense of responsibility to their classmates, 62.1% (n = 18) strongly agreed, 37.9% (n = 11) agreed that it did.

Students were also asked what they liked least about the exercise. Responses to this question fell into two categories. The difficulty of marking friends was remarked upon most often. The next most common response was that students found it difficult to determine a mark because of the lack of clearly defined weightings given to each descriptor. It meant that, while they felt confident about giving feedback on whether or not the performer had achieved technical or interpretative or professional quality, it was hard to transform that into a mark that would reflect what had been done in terms of university standards. Some of the responses from students were:

> "I felt like allocating the mark was too vague and could be a more structured process, for instance, our group simply thought about what was appropriate whereas another panel I sat on allocated a certain weight to each component of the rubric. Having more specific guidelines could make finding an actual specific mark much easier and clearer."

> "The lack of clearly defined marking criteria. The descriptors of quality were helpful in viewing the piece. But not in assessing it and giving it a final mark."

> "I felt we were given no weighting for each component of the assessment criteria – this made me unsure whether the marks I was giving were fair or whether I was

basing them too much on one area of the criteria while not giving enough importance to another area."

"Working with peers is difficult because we each have wavering standards. Some are lenient and some are strict."

"It was at times awkward discussing the performance of a class mate/friend with the other members on the panel which were also your friends."

"Having to disregard all the process the performers have gone through as we know them as mates and sometimes it was hard to just judge the performance and not the progress."

Responsibility

Students were asked to comment on how they thought their personal knowledge of another student affected their judgment. Some respondents commented that prior realisation that this knowledge might affect their judgment made them not only self regulate, but also make sure the rest of the panel was aware of this issue

"It didn't and I made an effort not to let those who I was judging on the same panel with to take any preconceived dispositions into account."

"This was the hardest part about marking. I felt I had to distance myself from my peers in order to give a fair mark for the performance they delivered."

"I don't really think that I let anything affect my judgment. I tried to mark just on the basis of what I heard and the criteria that I had. I would expect everyone to do the same for me and I think the only way we will improve is by people being honest so that's why I tried not to let anything affect my judgment."

Quite a few of the responses to this question commented on how useful it was having the panel there to help overcome any personal bias or comparisons to previous performances in class by the person being assessed.

> "The only affect it may have had is comparing their previous in class performances to their performance on the day; however discussing with the other panel members helped me to try and separate this occasion from previous ones."

The final open-ended question in the survey was how students felt peer assessment affected their own learning. Most students felt that the peer assessment exercise helped with their own vocal development by firstly giving them criteria that described good quality singing, which they could reflect on and then forcing them to reflect on this by having to take responsibility for critically assessing each other's performance using the criteria. Some responses from students showed that they felt the peer assessment exercise was a good way to prepare them for an industry where they might have to critically assess a peer.

> "I feel like I have a better grasp and am more competent in terms of assessing someone's ability to perform well and now have a set of criteria I can apply . . . to my own practice as I can be careful not to do things that impair performance that I have noticed in others".

> "It will assist me to critically assess performances, which I may be required to do when in the industry."

> "I believe it helped my learning. I really enjoyed taking on the teacher role and being able to assess someone extensively and be able to express my own reflections on the student. I feel it is beneficial for the future where I will need to not only accept criticism but give it as well."

Intuition

In a few instances there was a contradiction between the feedback that was given and the mark awarded. For example, in one case the feedback was critical of a few areas of the singing, but the mark awarded was a high distinction. The lecturers on the other hand awarded a distinction for the same performance. This was the feedback given by the student panel:

> **Technical Achievement:** Breathy at times. Posture could be better. Good sob (could be improved). More twang. More attention to the end of phrases (but overall good phrasing). Panel thought she held her breath on long phrases, and this caused some constriction. **Interpretive Skills** Great interpretation! Good choices. Thought processes great when singing, but could be improved at times when not singing. **Professional Skills** Panel thought that more musical decisions could have been made in the musical introduction. Over all good relationship established with the audience. **Qualitative Judgment** Compelling, sophisticated and thoughtful!

In a follow up interview which was conducted in order to find out how the students on this particular panel derived their mark, it was interesting to note that they had broken the assessment into components and given each component a weighting. Then they gave a mark to each component and added up the total, which became the final mark. This was the only time that a mark was more than 10 points different from the lecturers' mark but what had mostly intrigued the researcher was that the mark was not consistent with the feedback, which was more critical. It reinforced the notion that analytical marking of performance does not necessarily give a true reflection of the performance itself. It also showed that by breaking the assessment down into component parts the students overrode their intrinsic knowledge of whether the performance surpassed the descriptors of quality and instead tallied up points and awarded a mark based on overall points scored. While the feedback from students in the questionnaire was that many felt less confident giving a mark without the help of weightings, in fact they were closer in standards to the more experienced lecturers when they marked holistically than when they broke the mark into components.

Discussion

The students in the focus groups spent many hours discussing issues that could arise from the peer assessment exercise. One of first problems raised was how to deal with the different skills levels among students. While they felt students who had already developed pre-existing skill levels in singing from receiving prior lessons should not be prejudiced against, they also agreed this was a difficult thing to deal with when assessing. Judging in panels was agreed would be the fairest way to deal with this issue.

The students in the focus group were concerned about whether or not assessment should take into account an individual's improvement or compare students. This was a contentious issue in the focus group. Some felt that improvement should be taken into account whereas others felt that this was not a true indication of how good a singer the student being marked was and that in the end that was the main point of the assessment. Again Sadler (2008) points out that criterion-based assessment was developed to give a benchmark for marking students so that not so much emphasis would be put onto a student's development or even comparison with other students in class, but with the criteria. The lecturers stressed comparison with criteria in the preparation for the peer assessment exercise.

Over marking of students was a concern that came up in the literature on peer assessment and the student focus group was especially aware of the possibility of people over marking friends. The focus groups also discussed the possibility of panels being too critical in order to prove how professional the panel members were. The students felt that careful selection of panels was paramount for the peer assessment to be successful. They felt that students should be given a chance to let the lecturer know that they would not feel comfortable marking certain other students because of prior practising relationships or friendships. In the exercise there was a certain amount of over marking from students, which was consistent with the other literature on peer assessment in music (Blom & Poole, 2004; Daniel, 2004; Searby & Ewers, 1997). Some of the students admitted in the survey that they had been influenced by the fact they were friends with the person they were marking. This came up mostly when asked what they liked least about the exercise, so they were aware of the pitfalls of having to peer review. The student focus group had discussed this potential situation in the focus

group meetings and minimised the effect by getting the lecturer to parallel mark with one third of the final mark being given by the student panel and the rest by the lecturer.

The students felt feedback on the assessment from the panel should be available. However, they believed that there should be a discussion in class on the ethics of being in a panel, in which it should be made clear that students should not approach individual members of the panel, and not prematurely disclose what had been discussed.

Falchikov (2007) refers to work done by Trevitt and Pettigrove (1995) which also indicated that peer assessment was thought by students to have relevance to their future careers. One comment by students in the focus group was the need to develop skills that would assist with judging or auditioning peers for companies they might set up or festivals they might be involved in running. All the students in the focus group agreed or strongly agreed that peer assessment was relevant to their personal skills development and also their future career plans.

The main benefit that the students perceived from the exercise was that it helped them to reflect on their own practice by having to make the effort to interact with the criteria given in order to properly assess a peer. This corresponded with findings at Kingston University that having to work out what criteria of assessment were to be employed and then having to apply them to real work focused the students' minds on what made the work good or bad (Searby & Ewers, 1997).

The comments made by students about each other's performances showed a level of sophistication about the way they were listening and critically assessing what they heard. The present study suggests that a great deal of preparation is needed in order to develop explicit knowledge about quality in singing. A series of workshops over a number of weeks is not long enough for students' tacit knowledge to develop sufficiently. A feature of the relationship between sign and behaviour or word and thought is that it undergoes fundamental change (Wertsch, 2007). Vygotsky began with "the assumption that signs first emerge in social and individual action without their users' full understanding of their meaning or functional role" (Wertsch, 2007, p. 186). It takes time for students to develop an understanding of terms and techniques they are learning. The entire first year of the

present singing course is spent preparing students for the peer assessment exercise by introducing them to concepts of singing and a language that they can use to describe what it is they are doing, and also getting them to informally critique each other under the guidance of the teacher. The added responsibility of having to assess another student in 2nd and 3rd Year encouraged the students to interact more carefully with the descriptors of quality so that "discernment of quality becomes a key aspect of learning" (Sadler, 2008, p. 18).

Conclusion

The critical feedback from student panels to each other was one of the most valuable features of the peer assessment exercise. The results show that students became agents of their own assessment activities by having to take more responsibility for marking other students, which led to them taking more responsibility for their own assessment. The process of having to think critically about other people encouraged students to reflect critically on their own practice and how effective it was, which is precisely the kind of graduate quality students need in order to keep learning after graduation.

Instead of being passive participants in assessment they became proactive in the process and this brought about changes in individuals who started to recognise other students' ability to perform well and gave them criteria they could apply to their own singing. Learning through social interaction was described by Vygotsky when he wrote, "*An interpersonal process is transformed into an intrapersonal one.* Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first *between* people (*interpsychological*) and then *inside* the child (*intrapsychological*)" (Vygotsky, 1978, p. 57). Thus, it can be concluded that the students co-constructed the assessment related knowledge, which they were able to appropriate as their own and apply to self-assessment.

The process of collaborative development and implementation of peer assessment, as described in this article, was underpinned by the provision of appropriate levels of support for students in the training which was conducted prior to the peer assessment exercise. Most important in

the process was the encouragement of discussion in class at the time criteria were distributed, and having the lecturer give particular attention to critiquing skills of students as part of the training.

Increased numbers of our graduates begin their performing lives by starting up independent companies or producing short play or music festivals. In such a context, they are often required to assess work done by their peers. Thus, the exercise also fulfilled the need to develop the skills and protocols that are required when graduates are later in a position to professionally judge a peer. Recognition by students of this future need meant that they took the peer assessment exercise very seriously and were very positive about it continuing as part of the course.

References

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32*(1), 21–24.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1–14.

Berk, L., & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and early childhood education*. Washington: NAEYC.

Blom, D., & Poole, K. (2004). Peer assessment of tertiary music performance: Opportunities for understanding performance assessment and performing through experience and self-reflection. *British Journal of Music Education, 21*(1), 111–125.

Daniel, R. (2004). Peer assessment in musical performance: The development, trial and evaluation of a methodology for the Australian tertiary environment. *British Journal of Music Education, 21*(1), 89–110.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Teaching International, 32*(2), 175–187.

Falchikov, N. (2007). The place of peers in learning and assessment. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education learning for the longer term* (pp. 128–143). London & New York: Routledge.

Gravemeijer, K., & Cobb, P. (2006). Design research from a learning perspective. In J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational Design Research* (pp. 17–51). London: Routledge.

Hunter, D., & Russ, M. (1996). Peer assessment in performance studies. *British Journal of Music Education, 13*, 67–78.

Latukefu, L. (2009). Peer learning and reflection: Strategies developed by vocal students in a transforming tertiary setting. *International Journal of Music Education, 27*(2), 128–142.

Lenski, S. (2001). Intertextual Connections during discussions about literature. *Reading Psychology, 22*, 313–335.

McKenny, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational Design Research* (pp. 67–90). London: Routledge.

Montalvo, F., & Torres, M. (2004). Self-regulated learning: Current and future directions. *Electronic Journal of Research in Educational Psychology, 2*(1), 1–34.

Reeves, T. C. (2000, April 27). *Enhancing the worth of instructional technology research through "design experiments" and other development research strategies.* Paper presented at the International Perspectives on Instructional Technology Research for the 21st Century, New Orleans, USA.

Sadler, R. (2008). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34, 159-179.

Searby, M., & Ewers, T. (1997). An evaluation of the use of peer assessment in higher education: A case study in the school of music, Kingston University. *Assessment & Evaluation in Higher Education, 22*(4), 371–383.

Stanley, M., Brooker, R., & Gilbert, R. (2002). Examiner perceptions of using criteria in music performance assessment. *Research Studies in Music Education, 18*, 43–52.

Thomas, A., & Millard, B. (2006). *Towards enhancing student learning and examiner reliability with criterion-referenced assessment in the creative arts: The case of music.* Paper presented at the Evaluations and Assessment Conference, Curtin University.

van den Akker, J. (1999). Principles and methods of development research. In j. van den Akker, R. Branch, K. Gustavson, N. Nieveen & T. Plomp (Eds.), *Design Approaches and Tools in Education and Training* (pp. 1-14). Dordrecht: Kluwer Academic.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), 5–23.

Wertsch, J. V. (2007). Mediation. In H. Daniels, M. Cole & J. V. Wertsch (Eds.), *The Cambridge Companion to Vygotsky*. Cambridge: Cambridge University Press.