FROM CONSTRAINTS TO POTENTIALITIES: SUSTAINABLE ASSESSMENT AS LEARNING GENERIC SKILLS THROUGH THE CREATIVITY COMMUNICATION FRAMEWORK (CCF)

Wing C. Lau[1]

[1]School for Higher and Professional Education (Design), Hong Kong Design Institute, Tseung Kwan O, Hong Kong,
wing_lau@vtc.edu.hk

ABSTRACT:

The paper aims to invigorate the discussions of assessment reforms in creative domains. Educational researchers need to understand not only what types of assessment reform work best, but also what work best when it is implemented at an institutional level. Some institutional barriers are shown to be difficult to tackle. The study provides a way of achieving productive student learning by focusing on what students do instead of what students are or what tutors do, following the principles of Outcome-Based Learning (OBL). Although it may not address the institutional barriers directly, holistic reforms implemented in a large-scale may not be as effective or efficient as incremental reforms and there seems be an inherent limitation of the time required for a large-scale reform regardless of manpower. Educational researchers need to play special attentions to this challenge when creating their proposal for changes.

1. INTRODUCTION

Increasing student diversity as a result of widen participation of higher education (HE) in the past two decades has led us to the use of a problem-based and an outcome-based approach (OBA) to promote student active learning (Biggs & Tang, 2011). The OBA introduces the concept of constructive alignment providing a visible pathway from intended learning outcomes (ILOs), to learning activities and to assessments. The systemic view of active learning enables a vertical integration from the programme level, to the curriculum level and finally down to the individual course level. The assessment of outcomes (both formative and summative) are especially important playing the gatekeeper role of OBA learning and has been an area of intensive research (Brown & Glasner, 1999). The changing contexts also call for innovative assessment methods including self-assessment and peer-assessment and many successful cases of innovative assessments have been reported (Bryan & Clegg, 2006; Rust, Price, & O'Donovan, 2003).

Nonetheless, there has been a long history of tensions and accountability issues in assessment well before the introduction of outcome-based learning (OBL). The debate is linked to economic and political changes in society calling for greater accountability of institutions and more democratic modes of assessment (Somervell, 1993). The more recent debate is directed toward OBL criticizing its shortsightedness of catering only for intended learning outcomes while neglecting emergent but desirable learning outcomes (Carless, 2009a). Boud (2000, p. 151) suggests that there is a dual role of assessment that "assessment acts need to meet the specific and immediate goals of a course as well as establish a basis for students to undertake their own assessment activities in the future." His framework for sustainable assessment emphasizes the importance of formative and self-assessment. He proposes that part of the learning outcomes for the learning society should include "self-assessment strategies, understanding and setting criteria, ability in identifying cues and clues from the context of learning, making appropriate judgements, giving and receiving feedback (Boud, 2000, p. 163)." These guiding principles also are advocated and adopted by other educational researchers (Brew, 1999; Carless, 2007; Rust et al., 2003; Sadler, 2010) when piloting their assessment reforms or interventions.

2. EFFECTIVENESS OF ASSESSMENT REFORMS

Since there are many types of educational reforms or interventions that are shown to be effective in promoting student learning (performance), Hattie (2009) synthesized over 800 meta-analyses and proposed a more objective indicator, an effect size, should be used to compare and determine what type of intervention works best, at least, in an experimental setting. An effect size (d) indicates the mean difference of student performances between the experimental group with the intervention and the control group without the intervention. The larger the effect size, the bigger the impact of an intervention on student performance. He suggested a cut-off point of 0.4 (in terms of a standard deviation) in order to judicially select the more effective interventions for implementation. The top 10 educational interventions related to assessment and feedback include student self-report grades (Rank=1, d=1.44), formative evaluation to lecturers/tutors (Rank = 3, d=0.90) and feedback (Rank=10, d=0.73). Many pilot reforms have been reported to be successful. For instance, Rust et al. (2003) reported that there are significant immediate and long-term (1-yr after) effects (effect size = 0.6 and 0.55) on student performances by developing student understanding of assessment criteria and processes in the business domain. The intervention includes the use of a 3-hr workshop that exposes students to the experience of grading and discussing of sample assignments with assessment criteria and grade definitions and the subsequent use of self-assessment with student actual coursework. These results further support our case for implementing sustainable assessment reforms.

2.1 ASSESSMENT REFORMS AT AN INSTITUTIONAL LEVEL

Ideally, assessment reforms should focus on these proven aspects in order to create the biggest impact on student learning. The knowledge gained from these pilot interventions and innovative assessment methods may tempt researchers to logically conclude that the time for a holistic change of assessment cultures in HE has finally arrived. The only missing component seems to be the additional resources required to fully implement these innovations and reforms at every levels articulated by Outcome-Based Learning (OBL). However, the analyses of several cases of assessment reforms at an institutional level seem to suggest that there are potential human and institutional barriers including the issues of communications, additional trainings, skepticism among staff members and organization structure that are unyielding to change. Several relevant cases are discussed below.

For instance, the Assessment Compact initiative (Rust, Price, Handley, O'Donovan, & Millar, 2013) that calls for staff and students to become assessment literate. The initiative was well designed with a strong theoretical and practical evidence base for change and high stakeholder (senior management and student union) involvement and commitment, and a long-term process of monitoring and evaluation. After a 2-year implementation period, the results suggested that the initiative was partially successful in raising awareness among staff (78%) but had a limited effect on students (23%). The authors concluded that the initiative might have sought to achieve too much too fast and "there has been a tension between precision of language and communication throughout the process (2013, p. 158)." Significant changes in staff attitudes to assessment and pedagogic culture were difficult to achieve due to the tensions resulted from a wide range of different views concerning assessment (purposes, relationship to learning, etc.), and a steep learning curve for staff whom are unaware of the complexity of assessment and its relevant research. There was also an issue of the number of real buy-in by stakeholders. In retrospect, the authors would have planned a longer, staged process to raise awareness and achieve grassroots buy-in … educating both staff and students to understand the terminology and underlying concepts in order to reshape their practice more confidently (Rust et al., 2013).

Another multiyear university-wide quality enhancement project (Holden & Glover, 2013) was initiated aiming to improve the low ratings related to assessment and feedback in the UK National Student Survey (NSS). The initiative was designed to engage a large numbers of staff across the institution involving formal and informal activities through disseminating research-informed principles, developing resources and local case studies of good practices, and using technology to improve the efficiency and effectiveness of assessment and feedback practices. The initiative was complemented by different related activities at each level of the organization (i.e. student, module, course/programme, faculty level, institutional, external) and adopted an unusual strategy similar to a marketing campaign to increase stakeholders' awareness for the proposed changes. Nonetheless, the authors concluded that the initiative "is a cautionary lesson for those responsible for instigating and delivering institutional change programmes (Holden & Glover, 2013, p. 170)" since student dissatisfaction with feedback remains a significant issue. The challenges encountered were

summarized including the changing priority of the annually-elected members of the student union as well as the voluntary and varied participation of the initiative concerning students and staff in different subject areas. Surprisingly, their recommendation was going back to basics to act at the course/module level by creating a direct impact on student experience.

2.2 A LOCAL ASSESSMENT REFORM TO PROMOTE INSTITUTIONAL CHANGE

In the local context, the Hong Kong University Grants Committee (UGC, 2008) has started to advocate the adoption of the outcome-based approach to student learning during the first audits of local universities between 2008 and 2011. Carless (2007) initiated a high-profile UGC-funded project (2002-2005) using learning-oriented assessment and feedback to enhance student learning. Being aware of the potential difficulties of motivating staff to become involved due to competing demands on staff (other projects, own teaching, and research) and the general negative connotations of assessments, a combination of bottom-up strategies and top-down support from senior management was used to tackle the challenge. The project was designed to utilize prestigious overseas consultants, traditional academic and "scholarship of teaching" products and collaborate with the educational development unit in the university. Despite these efforts, active staff engagement was reported to be around 10% (based on 40 submissions from roughly 400 academic staff at the initial stage). He concluded that the project has been less successful in trying to promote wider institutional change and listed four institutional barriers. The barriers include the authority of change resided in the committee structure, the varying priority of senior management, the failure to mobilize middle management (such as heads of department or programme leaders (cf. Knight, 2000), and finally, the limiting impact on staff who were already relatively receptive to the ideas being promoted. Based on the experience, Carless (2009b) observed that "accountability and the lack of trust created an atmosphere constraining the use of innovative assessment methods." "The rewards for risk-taking in assessment are relatively low, and the challenges relatively high, contributing to conservatism in the technology of assessment. Any changes tend to affect those whom are more receptive." How to stimulate staff on the periphery is an ongoing challenge meriting further attention.

2.3 EFFECT SIZES REDUCED TO ACCOUNT FOR INSTITUTIONAL BARRIERS

Based on the above cases, the impact of educational interventions from pilot studies could not be replicated easily in large-scale implementations. A holistic change and innovative practices of assessment in promoting OBL seem to require tremendous efforts at an organizational and institutional level. The observation also implies that the effect size of a large-scale implementation will likely be reduced accordingly if resources (manpower, time, etc.) are diverted to tackle these barriers leaving fewer resources committed to the reform. In other words, educational researchers not only need to know what intervention works best, but also need to know what works best when implemented at an institutional level. Since both communication issues across different levels at an institution and low staff engagement are often cited as the potential obstacles, the reduced effect sizes from these factors can be roughly estimated. For instance, if only a portion of the staff members are genuinely adopting the reform (60% of real buy-in, 40% of low engagement/disengagement for reasons stated above, the ratio is used for illustrative purposes only), the effect size of the intervention will be reduced proportionally while the resources committed in terms of manpower may have been planned to the full extent.

The amount of efforts needed to communicate the values, meanings, and beliefs associated with a systemic assessment reform in addition to the principles, procedures and methods that may alter the social, interpersonal, disciplinary and cultural norms of assessment seems to increase exponentially with scale. Besides, a systemic reform most likely will involve multiple stakeholders from diverse backgrounds and positions including course teams, programme teams, academic committees and boards making the communication environment no less complex than that of a large software development project. The potential large reduction in effect sizes of educational interventions facing low staff engagement and communication overhead is astonishing and may prompt researchers to rethink whether pursuing a holistic reform is always effective or efficient. The ubiquitous understanding of OBL also seems to create a kind of hindsight emphasizing that a grandeur reform is always better, in other words, a holistic change in assessment is preferred to a few small changes. This may be the reason for some educational researchers (Rust et al., 2013) to oppose an incremental or piecemeal approach to assessment reforms. A major principle of OBL seems to be neglected that everything should focus on what students "do" instead of what students "are" or what teachers "do". Building capacity on the student side may be more cost-effective than

tackling changes at an institutional level. Some authors (Holden & Glover, 2013) actually recommended focusing changes at the course/module level in order to create a direct impact on student learning.

By using a simple yet effective technique to sustain changes similar to students learning generic skills or thinking skills (e.g. the technique of six-thinking hats (De Bono, 1999)), the proposed change should not generate any types of accountability or institutional issues as opposed to a reform of an existing assessment method. Using the issue of creativity assessment in the HE context as the focus of my discussion, a pilot study is used to demonstrate how the Creativity Communication Framework (CCF) can be considered to be a low-hanging fruit version of implementing sustainable self-assessment for life-long learning. The creativity communication framework lies at the technique level that does not require a heavy involvement nor commitment by a facilitator or assessor although such a commitment most likely will bring additional benefits.

3. PRINCIPLES OF THE CREATIVITY COMMUNICATION FRAMEWORK (CCF)

The technique follows many well-documented approaches (Cowan, 2006; O'Donovan, Price, & Rust, 2004) to improve student performance based on cultivating student self-evaluative skills, understanding of assessment criteria and standards, and self-report grading with justifications. The major difference is that the criteria are generic enough to be used in creative domains. The CCF contains 3 core principles. It aims to cultivate student understanding of creativity-related theories, criteria, and standards. It utilizes student self-assessment using self-grading with justifications. Students are expected to discuss their concepts using the criteria with others including but not limited to tutors. Finally, there is an audited assessment by tutors. Creativity grading and reflection of the CCF criteria can reveal hidden assumptions between students and tutors by prompting concrete feedback from tutors to address student strengths and weaknesses perceived in the report.

Key components of the CFF are as follows:
- Rubrics are developed based on reliable criteria related to measuring creative outcomes (Besemer & O'Quin, 1987; Sternberg, Pretz, & Kaufman, 2003), see Appendix for details.

- A Creativity Reflection Report (CRR) contains a student self-grade based on a subset of the above criteria and justifications with evidence for the concept generated. Only a subset of the criteria is used in order to balance student workload and the efforts for documentation using the CRR. The CRRs are then audited by tutors.

4. RESEARCH DESIGN AND RESEARCH METHODS

It is often suggested that the engagement of both tutors and students are equally important in promoting student learning. No empirical study has shown the relative effect of tutor engagement versus the effect of student engagement with an educational intervention. The experimental study posits to investigate the relative importance of the two factors and suggests that a technique-based reform may be more efficient and as effective by circumventing some of the inherent institutional barriers discussed above. The empirical study involves 32 undergraduate design students (Male: 12, Female: 20) in visual communication design undertaking their final year projects (FYP) under 7 design tutors' supervisions. Both statistical analysis of student performance (final project grade) and qualitative analysis of semi-structured student interviews are used to investigate the relationship between student engagement and tutor engagement with the CCF. Only the quantitative results are presented in this paper (cf. Horan, 2010 for details).

4.1 THE CCF IMPLEMENTATIONS: TRAININGS AND DATA COLLECTION

Both students and their tutors attended a three-hour presentation introducing the CCF criteria on two screens, one providing general examples, and the other, specific examples drawn from the field of visual communication. Additional handouts and examples elaborating the criteria were also provided. Students submitted interim and final reports along with their project work. The report contains a creativity self-grade and the relevant CCF criteria for assessing the ideas generated or developed in the given period. Ideally, tutor-student pairs should be randomly assigned into the experimental group and the control group for the study. Nonetheless, there were mixed reactions

from both students and tutors after the initial training of the CCF that include fairness of grade, complexity of the CCF, and the possibility of CCF's criteria constraining design concept creativity. The program leader chose to implement the model. As a compromise, students had the option to release their data, or not. All supervision was conducted at regular weekly intervals. Tutors and students naturally formed into experimental and control groups in this quasi-experimental study.

After the FMP ended, students representing each tutor were asked to volunteer for semi-structured video-interviews. Students responded to questions about their experience (or lack thereof) with CCF during both the design process and assessment as well as about their working relationship with the tutor. Interviews were time-coded, conducted in Cantonese, then translated into English by language studies research assistants. Another assistant edited and categorized responses, filtering out extraneous dialogues. Tutors were unaware of the video interviews thereby reducing bias stemming from tutors' opinions. Based on interviewee responses, tutors were divided into two groups: engaged with the CFF or lowly/disengaged with the CFF. Samples of dialogues extracted from the student interviews are listed below (Table 1) while student engagements are independently assessed by the degree of criteria comprehension expressed in the student CRRs. Because substantial project feedback was provided, students exhibiting either weak, or no, understandings of criteria were considered lowly/disengaged with the CFF. A 4-quadrant matrix was formed representing all possible combinations of tutor/student engagement (Fig. 1).
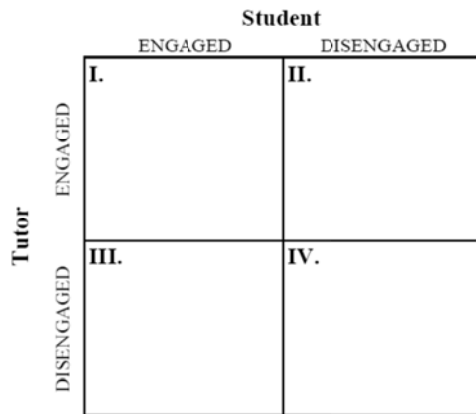

Fig. 1: The four quadrants of tutor/student engagement with the CFF.

Tutor No.	Sample dialogues extracted from the student interviews	Tutor CCF engagement (E/DE)*
	Tutor Support [Did tutor support the student in using CCF?]	
T-2	I have talked to (name) in terms of understanding the general factors and tried to get some hands-on understanding of what exactly I should write about the "reflection" thing.	E
T-4	We didn't discuss these terminologies. I think they are not related to the project.	DE
	CCF Understanding [Did the student understand CCF?]	
T-1	At first I only understood half of it, but when I was working on the final year project, I began to grasp more of the concepts. It helped me to explain my creative work. I didn't know creativity could be rationalized back then.	E
	CCF Usefulness [Did the student find CCF useful?]	
T-7	I think I seldom use this model. I worked on my design in a more alternative way. I was walking on a different road.	DE

Table 1: Sample student transcripts used to categorize tutors' CCF engagement.
*E: Engaged with the CCF, DE: lowly/disengaged with the CCF

5. RESULTS

5.1 DESCRIPTIVE STATISTICS

The project final grade (from 4.0 to 0.0) is consisted of three components: student creativity self-grade, tutor creativity grade, and tutor grade on the student CRR for the interim report and the final report. Means and standard deviations of the final grades corresponding to the four quadrants are presented in Table 2. The average final grade in Q1 is the highest (M = 3.31, SD = .66) with student-tutor pairs both engaged with the CFF and it is followed by the grade in Q3 (M = 3.20, SD = .50) with only students engaged with the CFF. The third ranked average grade is in Q2 (M = 2.70, SD = .36) with only tutors engaged with the CFF. The lowest average grade is in Q4 (M = 2.49, SD = .78) with student-tutor pairs both lowly/disengaged with the CFF. Most participants fell into either quadrant I or quadrant IV representing student-tutor pairs are either both engaged with the CFF or both lowly/disengaged with the CFF.

Final grade (Mean / S.D.)		Student		Total (Tutor)
		Engaged	Lowly/Disengaged	
Tutor	Engaged	Q1 (N=10) 3.31 (.66)	Q2 (N=3) 2.70 (.36)	N=13 3.17 (.65)
	Lowly/Disengaged	Q3 (N=5) 3.20 (.50)	Q4 (N=14) 2.49 (.78)	N=19 2.68 (.77)
Total (Student)		N=15 3.27 (.59)	N=17 2.53 (.72)	Total (N=32) 2.88 (.75)

Table 2: Student final grades organized by quadrant.

5.2 TWO-WAY ANALYSIS OF VARIANCE (ANOVA)

A two-way analysis of variance (ANOVA) was conducted to test whether there are significant differences concerning student engagement and tutor engagement with the CFF on student performance (final grade). The two independent variables are student engagement with the CFF and tutor engagement with the CFF while the dependent variable is the student final grade. All the p-values reported from the Shapiro-Wilk normality test and the Kolmogorov-Smirnova Z test were more than 0.05 indicating the final grade data are normally distributed. The Levene's test of Homogeneity of Variances did not show any violations.

There was a significant main effect for student engagement ($F(1, 28) = 5.30$, p =.029, partial η^2 =.159). The main effect of tutor engagement ($F(1, 28) = .307$, p =.584, partial η^2 =.011) was nonsignificant. The interaction effect between student engagement and tutor engagement was nonsignificant ($F(1, 28) = .029$, p =.866, partial η^2 =.001). Partial η^2 (eta squared) is a comparable measure for effect sizes (d). Cohen (1988) suggested that values of 0.0099, 0.0588 and 0.1379 would represent small, medium, and large effect sizes respectively. The results suggest a large effect size (an equivalent value of d = 0.88) for student engagement with the CFF on their performance. Fig. 2 shows the average final grades plotted against the four quadrants.

5.3 THREATS TO VALIDITY

Although the statistical results show a significant benefit for student engaging with the CCF, students' performances could be influenced by their motivations and expectations of the new system. However, all students should be equally motivated due to the importance of portfolios from the final year project for future job placement. Regarding tutor's expectations, tutors should provide a similar nurturing environment and support for students regardless of their beliefs of the new system since the tutor-student pair is arranged by mutual agreement rather than by chance. On the contrary, disengaged tutors might have an even greater incentive to perform in order to disprove the new process since they tended to believe in and value the existing process. There are certain limitations associated with the reported study. Shayer (1992) suggests that the lack of normative data or pre-test data may interfere with the interpretation of effect sizes for intervention

studies since participants are assumed to be equal in abilities. These limitations are acknowledged and are ameliorated when designing subsequent studies.
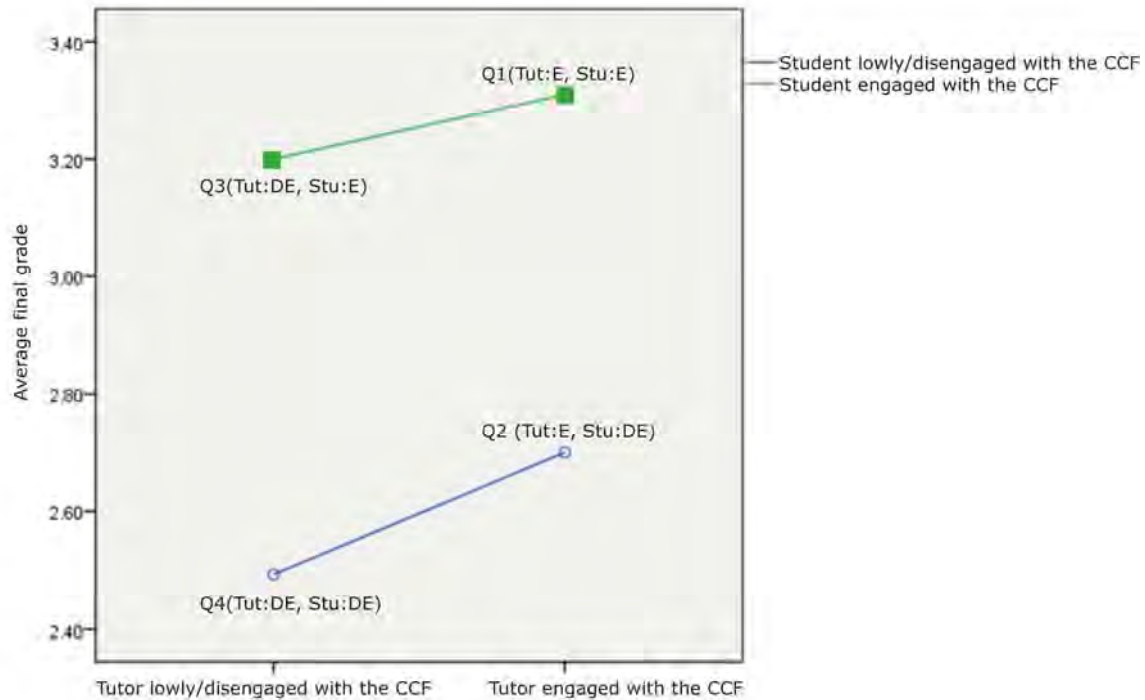

Fig. 2: Average final grades plotted against the four quadrants.
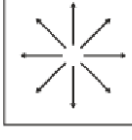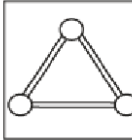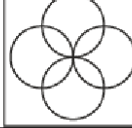
6. DISCUSSION
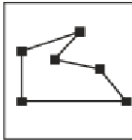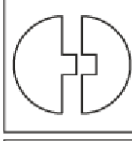
A major finding is that student engagement is the most important factor of improving student performance and seems to confirm that "a learning environment in which the student's learning and not the teacher's teaching is central (Dochy, Segers, & Sluijsmans, 1999, p. 337)." Nevertheless, the results should not be misinterpreted as discounting the impact of tutors on student learning and performance. On the contrary, based on the student interview data, most tutors are committed to student achievements albeit in their own ways. The reasons for low staff engagement with the CFF can be attributed to multiple factors rather than to individuals and are greatly influenced by the specific context. The negative connotation of assessment in the Asian context and the conflicting demands of assessment may play some parts in contributing the low engagement. Moreover, individual tutors may have developed their teaching philosophy over a long period of time and any holistic reforms implemented in a short notice may produce an unintended effect of defensive teaching. Although a holistic assessment reform seems to be preferred in theory, an incremental intervention similar to the CCF should not be repudiated solely because the intervention is a technique requiring mainly student engagement.

To be considered as a technique, the CFF can be used not only for the final year project, but preferably also for courses delivered in the first two years of college since the CFF enables students to understand standards and quality in a particular domain as well as develop self-evaluative capacities at their own pace. These knowledge and skills are pre-requisite for self-assessment, peer-assessment and collaborative assessment that are shown to be a cheap and effective way to change the way students learn (Dochy et al., 1999; Gibbs, 1999). In a knowledge-based economy, information is widely accessible making the old learning metaphor of teacher transmitting knowledge to student obsolete. The teacher role should be activating students to realize their potentials and have the confidence and self-efficacy to engage in learning in new and unfamiliar contexts by developing the practices of sustainable assessment for lifelong learning.

REFERENCE:

Besemer, S. P., & O'Quin, K. (1987). Creative product analysis: Testing a model by developing a judging instrument. In Frontiers of creativity research: Beyond the basics (pp. 367–389).

Biggs, J., & Tang, C. (2011). Teaching for Quality Learning at University. McGraw-Hill International.

Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. Studies in Continuing Education, 22(2), 151–167.

Brew, A. (1999). Towards autonomous assessment: using self-assessment and peer assessment. In S. Brown & A. Glasner (Eds.), Assessment matters in higher education: Choosing and using diverse approaches (pp. 159–171).

Brown, S., & Glasner, A. (1999). Assessment Matters in Higher Education: Choosing and Using Diverse Approaches. Open University Press.

Bryan, C., & Clegg, K. (2006). Innovative assessment in higher education. Routledge.

Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. Innovations in Education and Teaching International, 44(1), 57–66.

Carless, D. (2009a). Learning-oriented assessment: Principles, practice and a project. In L. Meyer, S. Davidson, H. Anderson, R. Fletcher, P. Johnston, & M. Rees (Eds.), Tertiary Assessment and Higher Education Student Outcomes: Policy, Practice, and Research (pp. 79–90). Wellington, New Zealand: Ako Aotearoa.

Carless, D. (2009b). Trust, distrust and their impact on assessment reform. Assessment & Evaluation in Higher Education, 34(1), 79–89.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Routledge Academic.

Cowan, J. (2006). How should I assess creativity? Developing Creativity in Higher Education, 156–173.

De Bono, E. (1999). Six Thinking Hats (2nd ed.). Back Bay Books.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. Studies in Higher Education, 24(3), 331–350.

Gibbs, G. (1999). Using assessment strategically to change the way students learn. In Assessment matters in higher education (Vol. 41–53, p. 53). The Society for Research into Higher Education & Open University.

Hattie, J. (2009). The black box of tertiary assessment: An impending revolution. In Tertiary assessment & higher education student outcomes: Policy, practice & research (pp. 259–275).

Holden, G., & Glover, C. (2013). Fostering institutional change in feedback practice through partnership. In Reconceptualising Feedback in Higher Education: Developing dialogue with students (pp. 160–171). Routledge.

Knight, P. T. (2000). The Value of a Programme-wide Approach to Assessment. Assessment & Evaluation in Higher Education, 25(3), 237–251.

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. Teaching in Higher Education, 9(3), 325–335.

Horan, R. (2010). The Creativity Assessment Project: Research Report (OBA Project Code: 2007-08/OBA/SD4), School of Design, The Hong Kong Polytechnic University, Hong Kong.

Rust, C., Price, M., Handley, K., O'Donovan, B., & Millar, J. (2013). An assessment compact: changing the way an institution thinks about assessment and feedback. In Reconceptualising Feedback in Higher Education: Developing dialogue with students (pp. 147–159). Routledge.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes. Assessment & Evaluation in Higher Education, 28(2), 147–164.

Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. Assessment & Evaluation in Higher Education, 35(5), 535–550.

Shayer, M. (1992). Problems and issues in intervention studies. In Neo-Piagetian theories of cognitive development: implications and applications for education (pp. 107–121).

Somervell, H. (1993). Issues in Assessment, Enterprise and Higher Education: the case for self-peer and collaborative assessment. Assessment & Evaluation in Higher Education, 18(3), 221–233.

Sternberg, R. J., Pretz, J. E., & Kaufman, J. C. (2003). Types of Innovations (pp. 158–169).

University Grants Committee. (2008). Quality Assurance Council : audit manual. Hong Kong SAR, China: Quality Assurance Council.

APPENDIX: A SAMPLE RUBRIC USED IN THE STUDY, CRITERIA WITH CHINESE TRANSLATIONS THAT WERE BACK-TRANSLATED (HORAN, 2010).

Main Criteria	Sub-Criteria	Icon	Description / Grade
B. SUBSTANCE OF SOLUTION			**Grade A**
1. Types of Novelty (創新的種類)	a. Original (原創)		Solution is unusual or infrequently seen in a universe of products made by people with similar experience and training
	b. Germinal (創萌性)		Solution will likely suggest, or influence, future creative products and ideas in the same field or in other areas
	c. Transformative (轉革性)		Solution is so revolutionary that it forces a shift in the way that reality is perceived by others, introduces a significant change or modification. *[Note: this type of solution is rare even in the professional context]*
2. Problem Resolution (解決方案)	a. Adequate (中規中矩)		Solution answers majority or problem needs, or performs a desired function for problems that are very complex involving many variables or little available information. *[Note: Tutor must clearly establish the nature of the problem]*
	b. Logical (合邏輯)		Solution is logical, that is, it accounts for the rules, or rationale, of the domain (e.g. scientifically true or valid, consistent with the facts)
	c. Appropriate (完全合適)		Solution is suitable, makes sense, relevant or aptly applies to the problem *[Note: If this criteria is selected, it overrides 'Adequate' as a criteria]*
	d. Useful (有用, 多功能)		Solution has clear practical applications and the number of applications allows it to be considered very useful or operable.
	e. Valuable (有價值)		Others value solution because it fills an immediate or long-range intrinsic, informational, economic, physical, scientific, social or psychological need *[Note: [Note: If this criterion is selected, it overrides 'Adequate' and 'Appropriate' as a criterion. Also, the specific value(s) must be clearly stated]*
3. Elaboration & Synthesis (細緻及融會程度)	a. Coherent (緊扣性)		Solution has a sense of organizational unity, wholeness, comprehensiveness, balance, proportion and completeness about it.
	b. Complex (複合, 多元多層次)		Solution contains many elements at one, or more levels (e.g. technical, ideational, phenomenal complexity)