

Gender Considerations and Innovative Learning-centred Assessment Practices

Linda Lanyon PhD and Harry Hubball PhD

1. Abstract:

The desire to improve the quality of student learning and respond to growing student diversity in higher education, combined with increasing demands for accountability of learning outcomes have led instructors to seek innovative pedagogies and alternative forms of assessment, beyond that of the traditional lecture and examination format (Angelo & Cross 1993; Frye 1999; Mowl, 1996; Shavelson & Huang, 2003). A growing body of literature supports the use of a wide range of learning-centred assessment methods. However, very few studies have investigated the impact of these innovative assessment practices, particularly with regard to student diversity. In this paper we examine potential gender bias considerations in relation to learning-centred assessment methods. Evidence from practical classroom experiences, as well as the higher education literature suggests that instructors should address potential gender biases during interactive assessment methods in order to enhance inclusivity in self- and peer-assessment learning environments. Recommendations are suggested for implementing innovative learning-centred assessment methods in the college and university classroom. We conclude by highlighting the need for more SoTL research around issues of assessment and student diversity, particularly with respect to gender, age, ethnicity, cultural, sexual orientation and socio-economic status.

Key Words:

assessment; innovative assessment; learning-centred assessment; self-assessment; peer-assessment; gender; diversity; inclusion

2. Introduction

It is currently a challenging time for student assessment¹ in higher education. Increasing numbers and diversity of students have placed considerable pressures on burgeoning faculty workloads. At the same time there has been a perceived need for more accountability in higher education and, hence, learning and assessment must be reconciled with accountability requirements (Frye 1999; Shavelson & Huang, 2003). More than ever, students need to be equipped with lifelong learning skills when they leave university (e.g., critical thinking, effective communication, ethical principles, research skills, etc.). As a result of these factors, and the desire to improve the quality of student learning, teachers have sought innovative methods of assessment beyond that of the traditional lecture and final exam formats.

Traditional exam methods of assessment tend to provide only summative feedback. Assumptions about learning are rooted in a "one-size fits all" process of knowledge transmission and information memory recall. This often leads to surface learning and factual memorization, with little higher levels of learning, and fosters a disregard for the value of further learning in higher education (Angelo & Cross, 1993). In contrast, innovative assessment methods tend to emphasise formative feedback, which aims to facilitate on-going learning, motivation and student leadership throughout the teaching and learning process. According to Mowl et al. (1996; see also Boyd & Cowan, 1985; Brown & Dove, 1993), use of a wide range of learning-centred assessment methods (involving, for example, self, peer, group, and instructor feedback) empowers students to take more responsibility for their own learning.

In learning-centred environments, authentic assessment practices are integral to the learning process. Learning is driven by the method of assessment in both traditional and innovative environments (Race, 1995; Wergin, 1988). The use of a variety of learning-centred assessment methods allows a wide range of skills to be developed and the course content to become more tailored to individual learning aims. A deeper

¹ For the purpose of this paper, assessment refers to the data gathering process about student learning whereas evaluation refers to judgements that are made (e.g., using explicit criteria) based on the quality of student learning. Furthermore, the term formative assessment and/or evaluation refers to on-going data gathering and/or feedback on student learning whereas summative assessment and/or evaluation refers to a final compilation of students work and the judgement that is placed on the quality of that work (Chappuis & Chappuis, 2008).

approach to learning is encouraged (Boyd & Cowan, 1985; Brown & Dove, 1993; Mowl et al., 1996).

Key challenges facing any form of assessment practice centre around issues of reliability and validity, particularly since people are inherently biased in their judgements due to preconceptions about themselves and others. For example, research suggests that students generally tend to give themselves a lower score compared to the tutor's mark, and students with higher grades tend to underestimate their performance, whilst poorer performing students tend to overestimate their performance (Bryan et al., 2005; Edwards et al., 2003; Mowl, 1996). Individual personality traits can affect the estimation of one's own intelligence (Myers & McCaulley, 1985; Furnham & Chamorro-Premuzic, 2004) and, therefore, play a critical role in student-led assessment. Of course, any form of assessment might tend to favour some students over others, but it is necessary to identify potential biases in order to seek pedagogical strategies to enhance inclusivity in the university and college classroom. Investigation of gender, or other group, issues in connection with the use of innovative learning-centred assessment methods currently lacks adequate attention in the SoTL literature.

3. Gender differences and self-assessment practices

A substantial body of literature suggests that male students tend to overestimate their performance, whilst female students tend to underestimate their performance, despite female students actually being equal or outperforming male students (Bryan et al., 2005; Lind et al., 2002; Rees, 2003; Rees & Shepherd, 2005). Across disciplines and regardless of performance, women tend to have greater belief in their abilities in humanitarian subjects and interpersonal skills, whereas men tend to have greater belief in their abilities in science (Marsh 1990; Marsh & Yeaug, 1998).

Rees (2003) provides an interesting comparison of student versus tutor/mentor grades in the context of personal and professional portfolios in a medical school. Approximately equal numbers of students over and under assessed their portfolios compared to their mentor's grade. However, 75% of the cases in which the mentor assessed the portfolio more highly than the student, involved female students. Conversely, the majority of students (72.7%) who assessed their portfolios more highly than their mentor were male. Upon admission to medical school, Lumsden et al. (2005) found no gender differences in students' cognitive abilities, although female applicants tended to be more empathetic, with greater communitarian orientation, than male applicants. As their training progressed, Lurie et al. (2007) found that, compared to males, female medical students became more likely to give themselves lower ratings related to professional work habits. However, over the same period, female students became more likely than males to rate their own interpersonal skills highly. Similarly, Whittle & Murdoch Eaton (2001) found that skills requiring information handling, managing self-learning and understanding technical aspects were rated higher by male students than their female counterparts. These differences also tend to persist after graduation. For example, Clack and Head (1999) found that male medical graduates rated themselves more highly than did female graduates in attributes related to biomedical investigation, leadership and autonomy, whereas the females rated themselves more highly in interpersonal attributes and teamwork.

In the field of counselling, Fitzpatrick (1999) found that female students were less confident than males in rating their abilities and skills, and suggested that females perceived that a higher level of competence must be demonstrated in order to reach a particular standard of competence. This also applied to peer-assessment learning environments whereby females tended to view peers as less competent than their male counterparts rating the same peers. The University of Cambridge Faculty of History's Gender Working Party (1994) suggested that female students have a strong personal investment in "getting it right" and tend to make a cautious and detailed examination of the subject matter, examining different points of view; whereas males more easily adopt a particular point of view, dismiss others and take risks. This approach might also explain higher levels of anxiety found in female students (Abouserie, 1994; Masson et al., 2004; McKean & Misra, 2000; Pierceall & Keim, 2007).

In general then, the empirical evidence suggests that females in higher education apply more stringent criteria to their self-assessments and are less confident in their skills, except those relating to interpersonal attributes. Interestingly, in general cognitive testing, males are consistently found to be more confident than females in the accuracy of their responses although there tends to be a greater mis-calibration between accuracy and confidence in male students (Pallier, 2003).

4. Gender considerations and learning-centred assessment practices

It is not clear why gender differences in perception occur during self-assessment, despite generally equal and better student learning outcome performances by female students (Masson et al., 2004, Richardson & Woodley, 2003). Research suggests that females typically demonstrate lower levels of self-confidence than males with respect to advanced academic achievement (Dix, 1987; Glance & O'Toole, 1988, King & Cooley, 1995). Various theories propose that gender differences are innate (Myers & McCaulley, 1985), or due to early socialization which may lessen over time as children's upbringings place less emphasis on stereotypical gender roles and values. Aronson, Steel and colleagues (Aronson et al., 1999; Steele, 1997, 1998) have demonstrated that "stereotype threat" can influence student performance, such that extra pressure is felt by members of stereotyped groups not to behave in ways that confirm the perceived lack of ability in their group. This adversely affects assessment outcomes for females in disciplines such as mathematics, science and technology (Spencer, Steele & Quinn, 1999). The negative effect cannot be mitigated simply by students trying to suppress negative thoughts but can be overcome by using an alternative positive stereotype message (McGlone & Aronson, 2007).

Gender bias assessment issues become amplified when one considers the natural progression from self-assessment to peer-assessment as an integral part of learning-centred assessment practices. Gender issues in peer-assessment thus pose additional challenges (Boud, 1990; Brown & Knight, 1994; Topping, 1998). In addition to the inherent gender bias with self-assessment methods already discussed, the quality of peer relationships (e.g., issues of equality, respect and appropriate power relations) tends to have a greater bearing on the psychological well-being of female college students than male students (Frey et al., 2006). This difference could influence not only the quality of the learning experience but also performance assessment. Studies also

suggest that significant differences in subjective assessment of cognitive ability level exist dependent on the sex of the person being rated. For example, males and females consistently rate male family members as more intelligent than female family members (Rammstedt & Rammsayer, 2000; Furnham & Chamorro-Premuzic, 2005; Furnham & Valgeirsson, 2007). When this attitude is manifested within the higher education classroom it can create the existence of a stereotypical notion that men will perform better than women on intellectual tasks. Thus, peer-assessment practices might not provide the ideal complementary learning strategy to rectify gender inequality found in self-assessment practices unless proactive pedagogical steps are taken. However, the risk of "gendered feedback" (Murphy & Elwood, 1998), whereby academic staff might reward student work and offer feedback in line with their own gendered beliefs, exists in both traditional and learning-centred assessment practices.

5. Recommended strategies to address gender bias in learning-centred assessment practices

To empower all students to perform self- and peer-assessment more effectively instructors should proactively facilitate inclusivity in the university and college classroom. Essentially students should receive education to counteract biased belief systems in relation to assessing self and fellow students. The following recommendations to mitigate inequality in learning-centred assessment practices are not exhaustive:

- ❖ Metacognitive activities and reflection should form an early and integral part of students' courses to facilitate both transition into higher education and to develop students' higher order learning skills (e.g. Kell, 2006). As part of these activities, students need to receive awareness training to enable them to examine their own gendered, racial, ageist or other biased belief systems and, thus, counteract any biased feedback during peer-assessment. Priming students with positive messages about their group identity provides an effective way of alleviating the "stereotype threat" (McGlone & Aronson, 2007).
- ❖ Self-evaluation is difficult for students (Rees & Shepherd, 2005). They need support and training, which should include an explanation of research evidence on self-assessment. Given the importance of self-assessment to academic and future professional careers, students should be given ample opportunity to practise their self-assessment techniques and be trained to evaluate their performance more objectively. This will help them develop a realistic perception of their own performance and perhaps counteract group-related differences in self-assessment. Similarly, objective and fair performance assessment is required for peer-assessment in order that gender, race and age biases do not affect results. Objective measures of performance should be created either by the tutor or in collaboration with the class, so that students are encouraged to focus on actual performance without stereotypical bias. Scoring rubrics (Angelo & Cross, 1993; Brualdi, 1998; Moskal, 2000) provide one method of describing performance standards in detail. In class practise performing peer- and self-evaluation using rubrics provides feedback that may redress inequalities. Rubric case studies could specifically address issues of stereotyping so that students are encouraged to examine their own pre-conceptions.

- ❖ Gender bias influences that can result from self and peer-assessments reinforce the notion that tutor-led assessment should continue to form the key part of the student's overall assessment. When students have underestimated their own work, feedback from tutors can enhance self-confidence (Race, 2001). Conversely, when students have overvalued their work, feedback from tutors can be given in an explanatory manner that supports further learning. As an integral component of learning-centred assessment practices, formative feedback is not only an essential component to support learning but can also provide critical feedback to enhance and regulate self-concept and self-confidence for both male and female students, while also promoting a collegial sense of how to assess oneself and peers.

6. Conclusion

The impact of innovative assessment techniques in relation to diversity issues has received little attention in the SoTL literature. On the basis of the emerging evidence and practical classroom experiences, it appears that gender (among a broad range of other inclusion issues such as age, ethnicity, socio-economic class, sexual orientation, etc) must be considered as a moderator of the accuracy of learning-centred assessments in a wide range of disciplines (Blascovich et al., 2001; Crawford & Stankov, 1996; Pallier, 2003; Sorensen & Robinson, 1992; Stevenson & Stigler, 1992; Yates et al., 1997, 1998).

Across the range of skills required of students using these new techniques, gender and other group or individual biases might be equalized because different components of learning-centred assessment methods are likely to benefit different groups. For example, although potentially disadvantaged in self- and peer-assessment, female students may be advantaged compared to males in relation to the ability to be an autonomous learner and readiness for self-directed learning (Kell, 2006), as well as confidence in interpersonal communication (Clack & Head, 1999; Lurie, 2007; Omigbodun & Omigbodun, 2003). Further empirical investigation of these issues, across disciplines, is needed so that instructors can ensure they use a complementary and inclusive range of assessment methods.

In summary, this paper provides a critical review of gender bias considerations when employing learning-centred assessment practices in the university and college classroom. Ideally, experiences in higher education provide a positive influence on self-concept and the adaptation of perceptions of skill levels. Instructors have a vital role in this process by ensuring equity when introducing methods of assessment, particularly where those methods rely on judgements of, and confidence in, students' own abilities. While learning-centred assessment practices become more widely implemented, instructors should be aware of differences in attitudes, confidence and approaches. Further empirical SoTL research around issues of assessment and student diversity, particularly with respect to gender, age, ethnicity, cultural, sexual orientation and socio-economic status would yield a more inclusive higher education learning environment.

Authors' contact Information

Linda Lanyon¹ and Harry Hubball²

1. Departments of Medicine and Ophthalmology & Visual Sciences

2. Department of Curriculum Studies,

University of British Columbia

Address: Human Vision & Eye Movement Laboratory, University of British Columbia, Room 365, 3rd Floor Research Labs, VGH Eye Care Centre, 2550 Willow Street, Vancouver, B.C. V5Z 3N9 Canada

tel: +1 604-875-4111 x69003 fax: +1 604-875-4302

email: llanyon@eyecarecentre.org

References

- Abouserie, R. (1994). Sources and levels of stress in relation to locus of control and self-esteem in university students. *Educational Psychology*, 14, 323-30.
- Angelo, T.A., & Cross, K.P. (1993). *Classroom assessment techniques: A handbook for college teachers*. San Francisco, CA: Jossey-Bass, 58-59.
- Aronson, J., Lustina, M.J., Good, C., Keough, K., Steele, C.M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35, 29-46.
- Blascovich, J., Spencer, S.J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: the role of stereotype threat. *Psychological Science*, 12(3), 225-9.
- Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education*, 15(1): 101-111.
- Boyd, H., & Cowan, J. (1985) A Case for Self-Assessment Based on Recent Studies of Student Learning. *Assessment and Evaluation in Higher Education*, 10(3): 225-235.
- Brown, S., & Dove, P. (1993) Self and peer assessment: learning from experience. *Natfhe Journal*.
- Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.
- Brualdi, A. (1998). Implementing performance assessment in the classroom. *Practical Assessment, Research & Evaluation*, 6(2). Retrieved October 5, 2006 from <http://PAREonline.net/getvn.asp?v=6&n=2>
- Bryan, R.E., Krych, A.J., Carmichael, S.W., Viggiano, T.R., & Pawlina, W. (2005). Assessing professionalism in early medical education: experience with peer evaluation and self-evaluation in the gross anatomy course. *Annals of the Academy of Medicine Singapore*, 34(8), 486-91.
- Chappuis, S. & Chappuis, J. (2008). The best value in formative assessment. *Educational Leadership*, 65(4), 14-19.
- Clack, G.B. & Head, J.O. (1999). Gender differences in medical graduates' assessment of their personal attributes. *Medical Education*, 33(2), 101-5.
- Crawford, J.D., & Stankov, L. (1996). Age differences in the realism of confidence judgements: A calibration study using tests of fluid and crystallized intelligence. *Learning and Individual Differences*, 8(2), 83-103.
- Dix, L.S. (1987). Women: Their under-representation and career differentials in science and engineering. *Proceedings of a workshop at the Natonal Academy of Sciences*. Washington DC: National Academy Press.
- Edwards, R.K., Kellner, K.R., Siström, C.L., & Magyari, E.J. (2003). Medical student self-assessment of performance on an obstetrics and gynecology clerkship. *American Journal of Obstetrics & Gynecology*, 188, 1078-82.

- Fitzpatrick, C. (1999). Students as evaluator in practicum: Examining peer-self assessment and self-efficacy. *ACES Conference*. New Orleans.
- Frye, R. (1999). *Assessment, Accountability & Student Learning Outcomes*. Office of Institutional Assessment and Testing, Western Washington University. Retrieved October, 2007, from <http://www.ac.wvu.edu/~dialogue>
- Frey, L.L., Beesley, D. & Miller, M.R. (2006). Relational Health, Attachment, and Psychological Distress in College Women and Men. *Psychology of Women Quarterly*, 30(3), 303-11
- Furnham, A., & Chamorro-Premuzic, T. (2004). Estimating one's own personality and intelligence scores. *British Journal of Psychology*, 95, 149–160.
- Furnham, A., & Chamorro-Premuzic, T. (2005). Estimating one's own and one's relatives' multiple intelligence: a study from Argentina. *Spanish Journal of Psychology*, 8(1), 12-20
- Furnham, A. & Valgeirsson, H. (2007). Parents' estimations of their own intelligence and that of their children: A comparison between English and Icelandic parents. *Scandinavian Journal of Psychology*, 48, 289–298.
- Gender Working Party (1994). *Mens' and women's performance in Tripos examinations 1980-1993*. Unpublished paper, Faculty of History, Cambridge University: Cited by Murphy, P. & Elwood, J. (1998). Gendered learning outside and inside school: influences on achievement. In Epstein, D., Elwood, J., Hey, V. & Maw, J. (eds), *Failing Boys? Issues in Gender Achievement* (pp. 162–181). Buckingham, Philadelphia: Open University Press.
- Glance, P.R., & O'Toole, M.A. (1988). The impostor phenomenon: an internal barrier to empowerment and achievement. *Women & Therapy*, 6(3), 51-64.
- Kell, C. (2006). Undergraduates' learning profile development: what is happening to the men? *Medical Teacher*, 28(1), e16-24.
- King, J.E., & Cooley, E.L. (1995). Achievement orientation and the impostor phenomenon among college students. *Contemporary Educational Psychology*, 20(3), 304-12.
- Lind, D.S., Rekkas, S., Bui, V., Lam, T., Beierle, E., & Copeland, E.M. (2002). Competency-based student self-assessment on a surgery rotation. *Journal of Surgical Research*, 105, 31–4.
- Lumsden, M.A., Bore, M., Millar, K., Jack, R., & Powis, D. (2005). Assessment of personal qualities in relation to admission to medical school. *Medical Education*. 39(3), 258-65.
- Lurie, S.J., Meldrum, S., Nofziger, A.C., Sillin, L.F. 3rd, Mooney, C.J., & Epstein, R.M. (2007). Changes in self-perceived abilities among male and female medical students after the first year of clinical training. *Medical Teacher*, 29(9), 921-6.
- Marsh, H.W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, 2, 77-172.

- Marsh, H.W., & Yeaug, A.S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in development of math and English constructs. *American Educational Research Journal*, 35, 705-38.
- Masson, A.M., Hoyois, P., Cadot, M., Nahama, V., Petit, F., & Ansseau, M. (2004). Girls are more successful than boys at the university. Gender group differences in models integrating motivational and aggressive components correlated with Test-Anxiety. *Encephale*, 30(1), 1-15.
- McKean, M., & Misra, R. (2000). College students' academic stress and its relation to their anxiety, time management, and leisure satisfaction. *American Journal of Health Studies*, Winter 2000. Retrieved January 1st, 2008, from http://findarticles.com/p/articles/mi_m0CTG/is_1_16/ai_65640245
- McGlone, M.S., & Aronson, J. (2007). Forewarning and Forearming Stereotype-Threatened Students. *Communication Education*, 56(2), 119-133.
- Moskal, B.M. (2000). Scoring rubrics: what, when and how? *Practical Assessment, Research & Evaluation*, 7(3). Retrieved October 4, 2006 from <http://PAREonline.net/getvn.asp?v=7&n=3>
- Mowl, G. (1996). Innovative Student Assessment. *Deliberations*, London Metropolitan University. Retrieved January 2008 from <http://www.londonmet.ac.uk/deliberations/assessment/mowl.cfm>
- Murphy, P., & Elwood, J. (1998). Gendered learning outside and inside school: influences on achievement. In D. Epstein, J. Elwood, V. Hey, & J. Maw (Eds), *Failing Boys? Issues in Gender Achievement* (pp. 162–181). Buckingham, Philadelphia: Open University Press
- Myers, I.B., & McCaulley, M.H. (1985). *Manual: A guide to the development and use of the Myers–Briggs type indicator*, Consulting Psychologists Press, Palo Alto, CA
- Omigbodun, O.O., & Omigbodun, A.O. (2003). Influence of gender on undergraduate performance in psychiatry at Ibadan, Nigeria. *Medical Education*, 37(12), 1091-3.
- Pallier, G. (2003). Gender differences in the self-assessment of accuracy on cognitive tasks. *Sex Roles*, 48(5-6), 265-76.
- Pierceall, E.A., & Keim, M.C. (2007). Stress and Coping Strategies among Community College Students. *Community College Journal of Research and Practice*, 31(9), 703-712.
- Race, P. (2001). *A Briefing on Self, Peer and Groups Assessment*. York, UK: Learning and Teaching Support Network. <http://www.ltsn.ac.uk>. Cited by Rees & Shepherd (2005)
- Race, P. (1995). The art of assessing. *Deliberations*. Retrieved May, 2008, from <http://www.londonmet.ac.uk/deliberations/assessment/art-of-assessing.cfm>
- Rammstedt, B., & Rammsayer, T.H. (2000). Sex differences in self-estimates of different aspects of intelligence. *Personality and Individual Differences*, 29(5), 869-80.

- Rees, C. (2003). Self-assessment scores and gender. *Medical Education*, 37, 571-573.
- Rees, C., & Shepherd, M. (2005). Students' and assessors' attitudes towards students' self-assessment of their personal and professional behaviours. *Medical Education*, 39, 30-39.
- Richardson, J.T.E., & Woodley, A. (2003). Another Look at the Role of Age, Gender and Subject as Predictors of Academic Attainment in Higher Education. *Studies in Higher Education*, 28(4), 475-93.
- Sorensen, C.K., & Robinson, D.C. (1992). Gender and Psychological Type: Implications for Serving Nontraditional Students. *Continuing Higher Education Review*, 56(1-2), 35-47.
- Shavelson, R.J., & Huang, L. (2003). Responding responsibly to the frenzy to assess learning in higher education, *Change*, 35(1), 10-19.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-26.
- Steele, C.M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-29.
- Steele, C.M. (1998). Stereotyping and its threat are real. *American Psychologist*, 53(6), 680-1.
- Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York: Summit. Cited by Yates et al. (1998).
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Education Research*, 68(3), 249-76.
- Wergin, J.F. (1988) Basic issues and principles in classroom assessment. In J.H. McMillan (Ed) *Assessing Student's Learning*. New Directions for Teaching and Learning. San Francisco: Jossey-Bass
- Whittle, S.R., & Murdoch Eaton, D.G. (2001). Attitudes towards transferable skills in medical undergraduates. *Medical Education*, 35, 148-53.
- Yates, J.F., Lee, J.W., & Bush, J.G. (1997). General knowledge overconfidence: Cross-national variations, response style, and "reality". *Organizational Behaviour and Human Decision Processes*, 70(2), 87-94.
- Yates, J.F., Lee, J.W., Shinotsuka, H., Patalano, A.L., & Sieck, W.R. (1998). Cross-cultural variations in probability judgment accuracy: Beyond general knowledge overconfidence? *Organizational Behaviour and Human Decision Processes*, 74(2), 89-117.