



Business student's attitudes to criteria based self-assessment and self-efficacy

Danny Carroll

ASB eLearning, Learning and Teaching Portfolio, Australian School of Business
University of New South Wales

Reducing student passivity and designing reflective skills into tasks contributes to developing student's professional judgment capabilities (Boud, 2000). This study analyses Business student attitudes and practices related to self-efficacy, self-regulation, assessment and self-assessment in two courses where students practiced criteria level self-assessment. A survey instrument was developed and an exploratory factor analysis in both sampled groups showed broad consistency in factor identification and reliability. Both cohorts' evidenced similar presentations related to self-efficacy, positive associations with socially mediated learning and a positive attitude towards developing better professional judgment. Student's confidence in their ability to understand task level requirements and instructions was greater than their confidence in their ability to accurately judge against criteria. A significant number of third year undergraduate students reported they had few opportunities to do self-assessment activities as part of their degree. This highlights the need to better integrate self-assessment practice into our Program design.

Keywords: student attitudes to self-assessment, self-efficacy, professional judgment formation

Introduction

The context for this study is two tasks conducted in a postgraduate and an undergraduate Business School course where students did optional self-assessment against criteria. Review, online-marking software, was used to mark against the task criteria, provide feedback and collect student self-estimates against the criteria. This paper reports on student attitudes, self-regulation, experience and confidence related to assessment and self-assessment in criteria based tasks in two selected courses in the Australian School of Business (ASB), University of New South Wales (UNSW).

A strong sense of self-efficacy and robust self-regulatory mechanisms are synonymous with successful independent learners. Strong self-efficacy in learning contexts predisposes students to persist with difficult tasks and maintain and develop strategies that assist success. Learners with high self-efficacy will have high levels of perceived agency in tasks and are predisposed to cognitively engage with their learning. Cognitive engagement is a reinforcing process that assists learners to know better how to learn and provides critical frameworks for reflecting on what they do and don't know (Zimmerman 2000).

This research is informed by Bandura's social cognitive theory, which describes social, cognitive, and environmental inputs to individuals learning and development (Bandura 1994). Bandura's theories made a profound contribution to learning theory, educational approaches, psychology and organizational management. Central to the understanding of social cognitive theory is the role of self-efficacy, which is people's belief in their ability to succeed in specific situations (Luszczynska 2005), which supports motivation and success in learning. In relation to developing self-efficacy, Bandura describes determinants that promote developing self-efficacy including previous successful (mastery) experiences (which involved challenges), the examples of

significant models (e.g. teachers), by social inputs that encourage success and by affective factors that contribute to or inhibit learning (Ramachaudran 1998).

In this research, a particular focus is drawn on the relationship of self-efficacy and the predisposition towards and formation of accurate professional judgments in a discipline. In this particular sub-field, previous research into learner over-confidence and inaccurate self-assessment, starting with Dunning and Kruger's (1994) studies on novice learners, has been instructive. Studies by Garavalia (2003) and Hacker (2008) on calibration and Pajare's work on self-efficacy beliefs in academic settings (1995 and 2002), among others, were referenced to map out the landscape that informs this research.

Self-efficacy has a profound impact on individual learner's engagement with learning, perseverance, and reflection (cognitive engagement). From the point of view of Educators, a key Graduate Outcome is the learner's development of these critical and reflective learning skills. As students progress through Degree programs, they are at least tacitly expected to be approaching a state of proficiency, or even 'expert' status in their field. A necessary component of expertise is the development of accurate judgment. Recognizing what constitutes quality in domain knowledge, skill or performance is a foundational professional attribute. The ability and predisposition to critically judge one's own work and the work of others should be a key skill that is developed in undergraduate education. However there is evidence that the development of self and peer assessment skills is haphazardly applied throughout the mosaic of course curriculum that aggregate into degrees. (Berzins 2008)

The development of student's critical and reflective skills is best achieved through repeated, explicit and embedded use of a range of student and learner centered authentic assessment tasks through the course of study, starting in first-year (Leamson 1999). Regular use of tasks that scaffold, develop and allow or require students to test their ability to form accurate judgments, are the operational method to develop the challenging success experiences Bandura described (1994). In practice, these can include reflective writing tasks, student developed assessments, self-directed and self-reported tasks, and peer and self-assessment activities. Scaffolding student's understanding of what constitutes quality through embedded reflective activities that employ exemplars, rubrics and criteria that contextualize and make explicit the essential building blocks that 'break down' the key components of what 'quality' looks like, makes comprehensible to students why a D (Distinction grade) is a D.

Sadler (2010) describes key processes in the formation of expert judgment and notes the necessity of moving students from consumers of marks and feedback to active participants and judges in tasks and assessment. This requires systematic assessment design that builds opportunities for students to practice and test out their developing judgment skills. 'A defining characteristic of any profession is that it depends on the wise and effective use of judgment and knowledge...' (Angelo and Cross in Dunn, 2002). Student self-assessment is part of a suite of practices that helps students form a predisposition to rigorous critical and reflective practice. One of the clear developmental roads to achieving accurate holistic judgments is the practice and development of accurate discipline oriented judgments against criteria, although as Regehr (2006) notes, there are inherent challenges in accurate self-estimations. While Lew (2010) finds that there is no relationship between student beliefs about the utility of self-assessment practice and their self-assessment accuracy, there is a counter argument that even where student self-assessment accuracy does not demonstrably improve, that the engagement with reflective processes can encourage cognitive insight and gains (Basnet 2012; Regehr, 2006). Despite the challenges, the development of accurate judgments of oneself and others is arguably a critical graduate learning outcome.

Embedding meaningful, well-constructed and frequent self-assessment activities in assessment tasks is likely to promote both a social learning culture of and individual's predisposition towards practicing critical and reflective thinking skills in a systematic way. In particular, integrating self-assessment activity with online tasks in 24/7 systems allows learners to record their self-assessment judgments at any time, at the learning moment, whenever that is. The ability to make initial judgments and update them (anywhere, anytime) as the student's engagement and critique develops, is a powerful affordance of online self-assessment.

The wider use of embedded reflective activities such as self-assessment can change student's experience of assessment. The potential for reflective approaches including self-assessment to increase student engagement with course learning goals and assessment are clear (Gibbs, 2004). Indeed, the ideal outcome is for our assessment to do double duty: to both serve the needs of certification and of learning (Boud, 2000).

Context of study

The long-term research plan seeks to identify relationships around self-efficacy, student attitudes to assessment and self-assessment accuracy in student Business School cohorts and if there are measurable or self-reported effects of teacher interventions designed to improve student self-assessment accuracy. For the purpose of this paper (which represents an introductory stage of the research project), the focus is constrained to an early examination of student attitudes related to self-efficacy, self-regulatory behavior, self-assessment and confidence in their domain judgments and socially based learning, in a first year postgraduate and third year undergraduate course in the ASB at the University of New South Wales.

Lecturers in Charge (LICs) of two courses (one postgraduate, one undergraduate) hosted an ethics approved study to promote student self-assessment practices, survey students on their attitudes to self-assessment and self-efficacy and analyse the relationships between student course outcomes (task and final course grades) and their self-assessment accuracy and practices. These latter results will be reported on in detail in a separate paper.

Both courses marked two tasks through Review where students were encouraged to do self-assessment. The software allows staff to mark and comment against criteria and for students to self-assess their work against the same criteria. Students can self-assess more than once and were encouraged to use the self-assessment process to check progress to task and criteria fulfillment. The criteria judgment can be placed at any point on a visual slider scale for each criterion, which indicates a point on the University grading scale 0-49 Fail (F), 50-64 Pass (P), 65-74 Credit (C), 75-84 Distinction (D), and 85-100 High Distinction (HD). The staff marks and student self-assessment predictions are stored in a centralised database and can be downloaded for analysis.

Methods

The overall research plan employs a mixed methods design to investigate student attitudes, practice and accuracy in criteria based self-assessment. Components of the design include the opportunity to do embedded self-assessment (in Review), optional participation in surveys, interviews, lecture visits, supportive reinforcement emails and encouragement from the lecturers and a small number of in-class exercises (conducted by tutors) to encourage students to discuss self-assessment and improve their self-assessment accuracy.

This paper focuses on analysis of survey data from the two cohorts relating to student attitudes to self-assessment, their confidence in self-assessment in task and criteria fulfillment and their self-regulation, experience, their self-regulation activity and their experiences with assessment.

Factor Analysis Process:

An exploratory factor analysis was conducted to reduce the data set to a smaller number of correlated variables and identify latent relationships and themes in the surveyed data and population. Data from the surveys was collated in Excel; missing or uninterpretable variables were coded as missing. The coding design was verified prior to importation into SPSS (a statistical analysis software) and the one negatively worded item (item C22 survey1 and C23 survey 2) was reverse coded. Principal axis factoring with orthogonal (Oblimin) rotation was chosen for the first interrogation of the data, as this is a new instrument and no theoretical presupposition was placed on interpretation (Field 2009). The ordinal values in each survey were loaded as variables; missing values were excluded listwise and the analysis was set to sorted by size and small coefficients were suppressed at an absolute value $<.30$.

The regression method was chosen because correlation between scores was both acceptable and expected. When the Factor Analysis was run, a check for Eigen values greater than 1, and a check for excessive collinearity (greater than .9) in the Correlation Matrix were completed. The KMO statistic was above 0.7 in the analysis of both data sets and this is acceptable (Field 2009) for a sample of this size (Kaiser 1974). Bartlett's test of Sphericity was significant at $<.001$. In conclusion, while the samples are small, tests of sampling adequacy showed they were acceptable.

Table 1: Survey sampling adequacy tests

Survey 1: KMO and Bartlett's Test		Survey 2: KMO and Bartlett's Test	
KMO Measure of Sampling Adequacy.	.777	KMO Measure of Sampling Adequacy.	.865
Approx. Chi-Square	718.04	Approx. Chi-Square	1973.51
Bartlett's Test of Sphericity df	136	Bartlett's Test of Sphericity df	171
Sig.	.000	Sig.	.000

The extraction Method was Principal Axis Factoring, Rotation Method was Oblimin with Kaiser Normalization

In both surveys, the Eigen values and scree plot were checked and potential viable factors were identified. The pattern matrix was scanned and items with low communalities and weak or cross factor loadings (under .4) were noted for deletion. The KMO figure was checked on item deletion and strengthened throughout this process. The factor scores in the pattern matrix table were reviewed and the values were deemed acceptable. Deletion of items was stopped when a clean and interpretable factor set was found and a review of all the tables conducted. Sample sizes were noted as was the total variance of each factor after extraction. Reliability analysis was then performed on all the scale items of each factor and the Cronbach alpha scores and significance were checked and found acceptable in each instance, indicating the reliability of the factor and sub-scale items.

Instrument

An original survey modeled on existing standardised survey instruments was developed for this study into student's attitudes and practices related to assessment and self-assessment. Standardized instruments referenced in the development of this scale included Albert Bandura's Guide for constructing self-efficacy scales in Self-Efficacy Beliefs of Adolescents (2006), the 'Motivated Strategies for Learning Questionnaire', Zimmerman's structured interview for assessing student use of self-regulated learning strategies (1987) and Scherbaum's critique of the validity of generalised self-efficacy scales, were referenced in the construction of the survey instrument (2006).

The items on the test were organized into five sections, an introductory demographic section (6 items) and a question (A7) on student's final course mark prediction. The other items (questions 8 to 34 in the postgraduate survey and 8 to 37 in the undergraduate survey) were grouped in four sections titled 'Self-efficacy for self-regulated learning', 'Self-efficacy for socially mediated learning', 'Self-efficacy for learning and judging' and 'Attitudes about learning'. The second survey iteration (with the undergraduate group) saw the deletion of one item, introduction of four new questions and a number of minor rewordings. For this reason, references to item (question) numbers between the surveys do not exactly match in all cases.

Non-demographic questions were positively worded and presented as 5 point likert scale with 1 representing strongly disagree, 2, disagree 3 neither disagree nor agree, 4 agree and 5 strongly agree. The final question in the undergraduate survey was an open-ended response on the experience of doing self-assessment against criteria. The anonymous survey was distributed in paper based format at postgraduate lectures (with completion rates being high) and via an anonymous online survey to the undergraduate cohort (with acceptable completion rates).

Process

Two surveys were conducted with a first-year postgraduate and third-year undergraduate cohorts. The first survey was made available early in semester and the second immediately after the final assignment task and just before the exam period. The conduct and results of each initial survey in each cohort is reported on.

The first survey was conducted after students had a ten minute introduction (at lectures) on the reasons for doing self-assessment in tasks, had had a short (ten minute) socially based class activity conducted by their tutors, and had access to Review to do self-assessment in their first task (details below). The purpose of these inputs was to activate their thinking about the utility of self-assessment as a useful activity related to their learning and as an encouragement to meaningfully do the self-assessment in their assessment tasks.

The information at lectures included practical details on how to access Review, how to enter their self-assessment estimates, conditions of the study (participation in surveys was voluntary and did not affect course grades) and feed-forward of the criteria accuracy experience of previous cohorts at UNSW. The feed-forward also contained references to elements of theory related to optimism biases (Dunning and Kruger 1999, and Dunning, Heath and Suls, 2004,) and observations on local practice. Related to theory, slides with data presenting the optimistic self-assessment of weaker students and the under-estimation estimates of stronger students were presented. Relating to local practice, we showed similar data from previous cohorts with examples of over-optimistic student criteria and task self-assessment and performance gaps (against actual performance).

As part of the feed-forward process, students were advised in the lecture and via a summary email to avoid over-optimistic or aspirational self-assessment. Specifically to control for the misconception that over-optimistic self-assessment might influence markers to inflate their mark, students were informed that their self-assessment estimates were only visible to markers after marking was complete and therefore high self-estimates couldn't influence the markers judgment.

The socially based class activity was a ten to fifteen minute group based discussion in tutorials, analyzing the

criteria and discussing peer's approaches to fulfilling the task. This simple task operationalizes the social constructivist learning theories advanced by Bandura. One of the causes of learning changes and developing self-efficacy in individuals are the positive influences of significant models and social mores which students identify with. These reinforce actions they feel are achievable (Pajares, 2002). In domain specific educational contexts, this can be realized by lecturers and tutors (significant leaders) making time for socially mediated activity where individuals share and focus their thoughts on task deconstruction, criteria interpretation, task regulatory approaches and discussions around quality. Clearly, the goal is both to encourage deeper cognitive engagement with key elements of the task and the potential for affective factors (social and personal identification) to reinforce students' self-efficacy, engagement and persistence. (Bandura, 2001).

Survey one: first year postgraduate cohort

Ninety six of one hundred and thirty one students in a first year postgraduate course (semester 2, 2012) were voluntarily surveyed, 59.4 % were females and 40.6% males. 74% of the students were under 25 years of age and 75 % of the students self-reported that they were internationals. Chinese nationals represented the largest grouping (47.9%) followed by Australians (19.8%), then Indonesians (9.4%). 85.4 % of students reported that English was not their first language.

In relation to expected final mark (question B7) for the course, 32.3% predicted they would achieve a HD, 55.2% expected a D grade, 11.5% expected a C, and only 1% of students reported expecting to get a P. No students reported that they expected to receive an F on the course.

Four factors accounting for 55% of the total variance were presented as interpretable due to the strength of the values indicated in the final Pattern Matrix. The items also exhibit theoretical soundness as factors in relation to existing literature in self-efficacy, self-regulation and confidence in self-assessment in academic performance (Falchikov 1989). Negative values were noted in the column for Factor 2, but as these values were based on strongly positive mean scores in all cases, this was not interpreted as implying directionality, and a negative construct was not inferred; the negative value was seen as a result of the rotation process. (Walker 2012).

Factor 1:

Factor 1 was named 'Student confidence in understanding task and criteria' and item loadings that comprise the factor were D24 'I understand what the teacher wants me to produce in the assessment task', .893, D23 'I understand what the assessment questions means for this task', .847, D25 'I understand the meaning of the criteria I'm being assessed against in this task', .527 and D26 'I am confident I can accurately self-assess my work against these assessment criteria' .424.

Factor loading decreases from general understanding and confidence in understanding the requirements of the task and the task directions, to lower scores in comprehension of the criteria to the lowest score relating to student's confidence in their ability to accurately self-assess against the criteria (D26). Of the four factor items, this item also had the lowest mean score of 3.7 (out of a possible maximum score of 5). While this is well above the mid-point (2.5), it shows less certainty and confidence on this sub-skill. Despite some cross-loading onto Factor 4, D26 was retained as it enhanced factor comprehensibility. Reliability analysis on the sub-scale items returned a Cronbach alpha of .804 indicating the reliability of the factor set and this was significant at $p < .5$.

Factor 2:

Factor 2 was named 'Student identification with the utility of socially based learning'. The scale items were C20 'Discussing how my classmates approach, plan and do their assessments is very useful' -.825, E31 'Discussing a task with others helps me to understand the task better' -.812, E28 'I think studying with others is a good way to learn', -.748, C19 'Discussing assessment tasks with classmates helps clarify the assessment requirements' -.731, and C17 'I like to discuss assessment tasks with classmates' -.587. The items all relate to student identification with the usefulness of learning with others. The first four items demonstrate that students perceive a strategic advantage in using the expertise of their peers to prepare for assessments. The weakest item, C17 is related to their feelings around socially based learning, 'I like to discuss assessment tasks with classmates'. The items of this factor returned a Cronbach alpha of .865 indicating reliability and was significant $p < .5$

Factor 3:

Factor 3 was named 'Student identification with the utility of self-assessment practices' and the items were E29 'I think practicing self-assessment accuracy is useful' .783, D27 'Practicing self-assessment (in assessment tasks) regularly, is likely to lead to improved professional judgment' .726, E34 'I want to develop better self-assessment judgments' .626, E33 'Developing accurate professional judgment is as important as being judged

by teachers / experts' .486. The strongest items relate to the usefulness of practicing self-assessment and the likelihood that practicing self-assessment would likely lead to improving professional judgment. On these items the postgraduate group was more positive about the positive impact of practicing self-assessment than their undergraduate peers (see Table 4). Students surveyed are predisposed to developing better self-assessment judgments (E34). The weakest factor item by average score was E33, 'Developing accurate professional judgment is as important as being judged by teachers / experts'. The items of this factor returned a Cronbach alpha of .756 indicating reliability and was significant $p = <.001$.

Factor 4:

Factor 4 was named 'Student use of feedback and self-regulation'. The items were B16 'When I receive feedback, I usually consider how I can apply it to future tasks' .738, B15 'When I receive feedback, I usually read it carefully' .715, B10 'I check the task instructions and requirements more than once while completing the task' .517, B16A 'When doing a new assessment task, I usually think about the feedback from past assessments'. The strongest items relate to referencing current task feedback to future tasks (B16) and students assertion that feedback is diligently considered (B15). B10 refers to self-regulatory tasks processes and B16A (the weakest item) asserts that students think less about earlier task feedback when doing new tasks. Descriptive statistics for these items returned high means, 4.23, 4.38, 4.28 and 3.97 (out of 5) respectively. Again, incorporating feedback from earlier tasks (B16A) recorded the lowest mean level of agreement of these items. The Cronbach alpha of .743 indicated factor reliability and was significant at $p = <.001$.

Survey two: third year undergraduate cohort

One hundred and nineteen of six hundred and twenty two enrolled students voluntarily responded to the online Blackboard survey in the third year undergraduate course (semester 1, 2013). 61.3 % were females and 38.7% males. 98.3 % of the students were under 25 years of age. 52 % of students reported as locals and 47 % of the students self-reported as internationals. Chinese nationals represented the largest international grouping (33%) Notwithstanding the higher representation of local students, 68.1 % of students reported that English was not their first language, indicating the multi-cultural composition of our local student population.

25 students responded to the open-feedback question. Responses were grouped thematically on attitudes to self-assessment and were marked on-topic, then subdivided into negative, positive, observational (neither positive nor negative) and off-topic (course feedback and comments on the survey). Of the on-topic responses, 11 comments were positive, 1 was negative and there were 5 general observations and comments on self-assessment (e.g. 'I need to develop better self-assessment'). 8 responses related to comments on the course and the survey itself, The 1 negative commentator saw the activity as 'an easy way for uni teacher (sic) to do less work'. Far more students were favorable however, and comments included 'Self-assessment is very important for us to improve ... but hard to do at the beginning. I would like to have some more instructions about how to do it' and 'This is the first time I've completed a self-assessment survey and it has certainly helped me think about what it is I'm doing and how I can better approach not only this assessment but more assessments in the future'. In a nod to the way Review scaffolds embedded student self-assessment and the provision of criteria based feedback was the affirming comment, 'I wish learning technologies like these could be used in other courses as well'.

An examination of the pattern matrix, sums of squared loadings column and the scree plot showed there were three strong factors and one weaker one; the four factors accounted for 69.6 % of the cumulative variance on extraction.

Factor 1:

Factor 1 was named 'Student confidence understanding task and criteria and utility of self-assessment' and included most of the scale items identified in factors one and three of the previous years' postgraduate survey. The items comprising the factor and their factor loadings were D27 'I am confident I can accurately self-assess my work against these assessment criteria', .900, D25 'I understand what the teacher wants me to produce in the assessment task' .884, D24 'I understand what the assessment questions means for this task' .832, D26 'I understand the meaning of the criteria I'm being assessed against in this task' .808, D29 'I have a lot of experience of doing self-assessment of my own work' .724, D28 'Practicing self-assessment (in assessment tasks) regularly, is likely to lead to improved professional judgment' .692, B14 'I am confident I can successfully complete course related tasks' .650, E31 'I think practicing self-assessment accuracy is useful' .606 and E33 'Developing accurate professional judgment is an important skill' .428.

Item D29, 'I have a lot of experience doing self-assessment of my own work' scored the second lowest mean

score (3.18 out of a maximum possible score of 5) in the descriptive frequencies table. Surprisingly for a third year cohort, only 29.3 % of students reported they had had a 'lot' of self-assessment practice. The 66.4 % non-positive response (strongly disagree, disagree and neutral) was the largest negative response for any single survey question. This disturbing response from final-year students reveals that self-assessment activities are not widely practiced in our student's current course of study.

Again, in contrast to this, the factor items evidence strong positive student identification with the usefulness of self-assessment activity and their belief that practicing it can assist in developing sound professional judgments (D28, E31 and E33). Item E33, 'Developing accurate professional judgment is an important skill' recorded the highest mean survey score; 4.58 out of 5. When a reliability analysis was run on the sub-scale items it returned a Cronbach alpha of .948 indicated excellent reliability of the factor set which was significant at $p < .5$.

Factor 2:

Factor 2 was named 'Student identification with the utility of socially based learning' and was the same as factor 2 in the postgraduate survey, including most of the same items. The items comprising the factor were C21 'Discussing how my classmates approach, plan and do their assessments is very useful' .821, C20 'Discussing assessment tasks with classmates helps clarify the assessment requirements' .802, C17 'I like to discuss assessment tasks with classmates' .625, C23 'The opinions or examples of my class mates influences the way I prepare for assessments' .567.

Overall, the items show strong student identification with the utility of socially based learning. The item C20 'Discussing assessment tasks with classmates helps clarify the assessment requirements', scored 66.4 % agreement, the highest positive item score in the survey and C21 'Discussing how my classmates approach, plan and do their assessments is very useful' scored 58.6 % agreement.

The weakest factor item was C23 (see above) student survey raw survey scores showed that 38.8 % answered positively (strongly agree or agree), 43.1 % were neutral, and only 18.1% felt they were not (disagree and strongly disagree) influenced by peer example. So, while the students readily recognize the utility of socially based preparation around assessment, a smaller number of students recognize an influence or change in their own practice as a result of it. The reliability analysis on the sub-scale items returned a Cronbach alpha of .734 indicating the reliability of the factor set which was significant at $p < .5$.

Factor 3:

Factor 3 was named 'Student observations on learning' although a clear interpretation for a unifying theme in this factor is challenging. The items identified were B12 'I am easily discouraged when I do difficult tasks' .710, E36 'I find it difficult to be critical of my own work.' .584 and B19 'Learning is primarily an individual activity' .533.

The student's survey responses of these items show evidence of persistence. Item B12 showed a slightly negative kurtosis in responses to the question 'I am easily discouraged when I do difficult tasks', indicating that students felt they persevered when faced with difficult tasks. 35.3% replied negatively, indicating they considered themselves resilient, 38.8% were neutral and 25.8% indicated they were easily discouraged. On E36, 41.4 % of students either agreed or strongly agreed with the statement 'I find it difficult to be critical of my own work'. Student responses to item B19 'Learning is primarily an individual activity' were evenly divided; 38.8 % of students agreed, 29.3 % were neutral and 31.9 % disagreed. That such a large percentage of students felt that learning was a socially mediated activity was encouraging. Reliability analysis on the sub-scale items returned a Cronbach alpha of .802 indicating the reliability of the factor set which was significant at $p < .5$.

Factor 4:

Factor 4 was named 'Student use of feedback and self-regulation' and was consistent with factor 4 from the postgraduate survey. This factor comprised the items B16 'When I receive feedback, I usually consider how I can apply it to future tasks' -.805, B15 'When I receive feedback, I usually read it carefully' -.767 B17 'When doing a new assessment task, I usually think about the feedback from past assessments' -.698. The strongest factor loadings relate to referencing current task feedback to future tasks (B16) and students assertion that feedback is diligently considered (B15). B10 refers to self-regulatory tasks processes and, consistent with the postgraduate survey, students thought less about earlier feedback when doing new tasks than applying feedback to future tasks (B17).

The descriptive statistics of these factor items show the percentage of students who responded strongly agree or agree for B15, was 87.9 %, B16, 68.1 % and B17, 62 %. The strong negative skew (bias to the positive values)

on B16 and especially B15 is evidence that students do read feedback and report positive behaviors relating feedback to future task accomplishment. The factor's Cronbach alpha was .868 and it was significant at $p < .5$

Results and discussion

Exploratory factor analysis of both populations showed broad consistency in factor identification, the sub-scale items and factor reliability. Further refinement and testing of this instrument into student attitudes to self-assessment and self-efficacy with other cohorts, larger sample sizes and different contexts is required to confirm its validity. Both postgraduate and the undergraduate cohorts evidenced high degrees of self-reported self-efficacy, strong positive associations and identification with socially mediated learning and positive attitudes towards development of critical and self-reflective skills. Confidence in their self-assessment ability was reported in both. Somewhat surprisingly, the third-year undergraduate group provided less realistic course grade outcomes than the postgraduates, many of whom were recent arrivals. Several possible explanations exist, and apart from the cases of rising expectation and harder grading in final year courses, we have observed a general plateauing of student third year mark averages; it seems that final year undergraduates economise their effort and while confident of their ability, more students cruise to some extent in this year.

In relation to expected final mark for the course, the predictions of third year students were less inflated than those of the postgraduate group (more of whom were recent arrivals in Australia). Only 14.3 % (versus 30.1%) predicted they would get a HD and a larger number of postgraduates (57% versus 50.4%) predicted a D. Undergraduates more realistic projections of a C (credit grade, 31.1 % versus 11.8% undergraduate), is likely informed by calibration of recent grade experiences but still retains an overly optimistic bias. Very few students predicted getting a Pass and (unsurprisingly) no students in either cohort predicted they would fail.

Table 2: Question A7 – undergraduate and postgraduate students predicted and actual course grades

Third year undergraduates (surveyed, n = 119)			First year postgraduates (surveyed, n = 93)	
Grade	Percentage grade predicted	Percentage actual grade	Percentage	Percentage actual grade
F	0	5.12	0	1.04
P	3.4	48.20	1.1	25.77
C	31.1	27.90	11.8	48.45
D	50.4	10.22	57.0	20.62
HD	14.3	8.56	30.1	4.12
Invalid	.8	0	0	0
Total	100.0	100	100.0	100

In the mean scores in both surveys, students expressed slightly less confidence in their ability to accurately self-assess themselves against known assessment criteria in a task than other related items (see Table 3 below). Both cohorts recorded higher mean scores for the items D24 to D26, which related to more general understanding the task requirements, task criteria and their general confidence in fulfilling tasks in their Programs. Lower confidence in self-assessment accuracy against criteria could be associated with a number of causes. One cause is the lack of frequent opportunities to practice self-assessment and calibrate (and gain confidence in) their criteria accuracy judgment, which students referred to. Another cause could be an interference factor, from general (over) confidence and their strong sense of generalized self-efficacy, which dilutes as students consider more precise and practice-based task /discipline items. In any case, students' high level of confidence in their self-assessment accuracy was not matched by accuracy in either their course final grade prediction or their criteria self-assessment estimates (discussed in a future paper).

Table 3: Student average response scores on task comprehension and criteria confidence

	D24 I understand what the assessment questions means for this task	D25 I understand what the teacher wants me to produce in the assessment task	D26 I understand the meaning of the criteria I'm being assessed against in this task	D27 I am confident I can accurately self-assess my work against these assessment criteria
postgraduate survey	3.82	3.77	3.89	3.7
undergraduate survey	3.76	3.57	3.7	3.28

Score is the mean score of all respondents out of a possible 5 (mid-point is 2.5)

It's interesting that postgraduate students showed a more negligible difference (possibly evidence of general over-confidence as many were recent arrivals). In contrast the third year undergraduates reported a lower mean confidence score on D27 (ability to accurately self-assess against criteria) and, despite nearing graduation, they reported the second lowest mean statistic response (3.18) of the survey on item D29, 'I have a lot of experience doing self-assessment of my own work'. The low score on D27 exists in sharp contrast to the undergraduate's response on E33, 'Developing accurate professional judgment is an important skill', which at 4.47 agree, was one of the highest averaged scores in the undergraduate survey.

This evidence of how the students value the development of professional judgment skills and the patchy integration of course embedded self-assessment practice in this degree, points to a need to more systemically embed foundational reflective practices into program designs. Nulty (2011) observed the importance of the explicit introduction of these practices from the start of the undergraduate program. The foregrounding of the development of reflective / professional judgment attributes and its realization through frequent, course-embedded use and socially based discussion would create a learning culture predisposed to reflective practice.

Both cohorts returned similar mean scores and rated the aspirational item E35 'I want to develop better self-assessment judgments' and the identification with the construct in item E33, 'Developing accurate professional judgment is an important skill' as the highest average score items in the set of questions on the practice and utility of self-assessment (see Table 4). Both groups reported E34 'Developing accurate professional judgment is as important as being judged by teachers / experts' as the third highest item score. This indicates a strong positive orientation in the third year and first year students to learner independence and self-regulation.

Table 4: Student average response scores on self-assessment practices and utility

	D28 Practicing self-assessment regularly, is likely to lead to improved professional judgment	E31 I think practicing self-assessment accuracy is useful	E33 Developing accurate professional judgment is an important skill	E34 Developing accurate professional judgment is as important as being judged by teachers / experts	E35 I want to develop better self-assessment judgments
post-graduate survey	3.94	3.98	4.26	4.06	4.37
under-graduate survey	3.57	3.51	4.47	4.03	4.26

Score is the mean score of all respondents out of a possible 5 (mid-point is 2.5)

Both groups evidenced strongly positive responses to socially mediated learning; they report enjoying and being informed by peer contact. Only 18.1% of the undergraduate students felt they were not influenced by peer example while 38.8 % answered strongly agree or agree, indicating that a significant number of students do feel they change their behavior through the influence of peer's practice and examples related to assessment.

Conclusions

The consistency of the factors and factor items identified in results with the two student groups suggests that further confirmatory tests should be conducted with other cohorts to confirm the validity of this instrument to assess student attitudes to confidence in self-assessment and the utility of socially mediated learning. The evidence in this study reveals that large numbers of final year undergraduate students in the surveyed course feel they haven't had frequent opportunities to practice self-assessment. As the practice of self-assessment is a foundational professional skill in the development of critical and reflective approaches, this situation is illuminating for our institution and relevant for accreditation and regulatory bodies and Program Directors. Higher Education leaders and employer groups are increasingly focused on embedding the formation and development of graduate capabilities, such as critical and reflective thinking and professional judgment development, in Degree structures. The demonstration of student's professional judgment can only be achieved through regular, calibrated embedded practice, preferably through a systemic, complementary range of types of reflective activity. However, the most immediate, personal and practical reflective activity for students is a personal reflexive predisposition towards objective and informed self-critique and self-assessment. Achieving this would be a worthy learning outcome of a University program and positively contribute to life-long learning.

The students in this study positively identify with self-assessment and desire to improve and develop their

professional judgment skills. However, students not only self-reported a lack of practice, but their lower scores on confidence in their ability to accurately self-assess themselves against assessment criteria than task confidence, points to a lack of practice in this specific skill area. Additionally, their self-assessment estimates on their course grade pointed to endemic over-confidence, even in the final year of the undergraduate program. A wider, multi-cohort or program-based longitudinal study into practices that assist the development of self-assessment accuracy with an increased qualitative focus is proposed.

Acknowledgements

Peter McGuinn, Learning and Teaching Portfolio, Australian School of Business, University of New South Wales

Yiping Lin, Learning and Teaching Portfolio, Australian School of Business, University of New South Wales

Richard Morris, Victoria Clout and Young Un Kim, Australian School of Business, University of New South Wales

Chris Davison, Head of School, School of Education Studies, University of New South Wales

William Ashraff, Senior Manager Learning and Teaching Assurance, University of New South Wales

References

- Bandura, A. (1994). Self-efficacy, [Wiley Online Library](#).
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1-26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-Efficacy Beliefs of Adolescents*. Information Age Publishing, Charlotte North Carolina, <http://www.uky.edu/~eushe2/Bandura/BanduraGuide2006.pdf>
- Basnet, B., Basson, M., Hobohm, C., & Cochrane, S., (2012) Students' self-assessment of assignments - is it worth it? Proceedings of the 2012 AAEE Conference, Melbourne, Victoria. http://eprints.usq.edu.au/22981/2/Basnet_et_al_AAEE2012_AV.pdf
- Berzins, M., & Sofo, F. (2008). Developing critical thinking skills in first year Australian university students. *International Journal of Learning*, 15(1), 261-272. <https://doi.org/10.18848/1447-9494/CGP/v15i01/45584>
- Boud, D., Lawson, R., & Thompson, D., Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, accessed on 12/05/2013 <http://www.tandfonline.com/doi/abs/10.1080/02602938.2013.769198#preview>
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399-413. <https://doi.org/10.1080/02602930600679050>
- Dunn, L., Morgan, C., & Parry, S. (2002). Seeking quality in criterion referenced assessment. At the Learning Communities and Assessment Cultures Conference, University of Northumbria, 2002
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment. implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106.
- Falchikov, N., Boud, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research*, Vol 59, No 4. <https://doi.org/10.3102/00346543059004395>
- Field, A. (2009). *Discovering statistics using SPSS*. London, Sage Publications
- Garavalia, L., & Ray, M. (2003). Distinctions among subgroups of developmental students: Differences in task value, self-regulated learning, and grade expectations. *Research & Teaching in Developmental Education*, 19(2), 29-40.
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1(1), 3-31.
- Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and Learning*, 3(2), 101-121.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Leamson, R. N. (1999). *Thinking about teaching and learning: Developing habits of learning with first-year college and university students*. Stylus Publishing, Cork
- Lew, M., Alwis, W., & Schmidt, H., (2010). Accuracy of students' self-assessment and their beliefs about its utility, *Assessment & Evaluation in Higher Education*, Volume 35, Issue 2 <https://doi.org/10.1080/02602930802687737>

- Luszczynska, A., & Schwarzer, R. (2005). Social cognitive theory. *Predicting Health Behaviour*, 2, 127-169.
- Nulty, D., (2011). Peer and self-assessment in the first year of university. *Assessment & Evaluation in Higher Education*, Volume 36, Issue 5. <https://doi.org/10.1080/02602930903540983>
- Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice*, 41(2), 116-125. https://doi.org/10.1207/s15430421tip4102_8
- Pajares, F., & Kranzler, J. (1995). Role of self-efficacy and general mental ability in mathematical problem-solving: [A path analysis](#).
http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED387342&ERICExtSearch_SearchType_0=no&accno=ED387342
- Regehr, G., & Eva, K. (2006). Self-assessment, self-direction, and the self-regulating professional. *Clinical Orthopaedics and Related Research*, 449, 34-38. <https://doi.org/10.1097/01.blo.0000224027.85732.b2>
- Scherbaum, C. A., Cohen-Charash, Y., & Kern, M. J. (2006). Measuring general self-efficacy A comparison of three measures using item response theory. *Educational and Psychological Measurement*, 66(6), 1047-1063.
- V. S. Ramachandran (Ed.). (1998). *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press.
- Walker, T.T. (2012). *Statistics in Criminology and Criminal Justice: Analysis and Interpretation*, Chapter 14, Jones & Bartlett Learning; Burlington
- Zimmerman, B. J. (1999). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82-91. <https://doi.org/10.1006/ceps.1999.1016>
- Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4), 614-628. <https://doi.org/10.3102/00028312023004614>

Author contact details: Danny Carroll, d.carroll@unsw.edu.au

Please cite as: Carroll, D. (2013). Business student's attitudes to criteria based self-assessment and self-efficacy. In H. Carter, M. Gosper and J. Hedberg (Eds.), *Electric Dreams. Proceedings ascilite 2013 Sydney*. (pp.118-128). <https://doi.org/10.14742/apubs.2013.1379>

Copyright © 2013 Danny Carroll.

The author(s) assign to ascilite and educational non-profit institutions, a non-exclusive licence to use this document for personal use and in courses of instruction, provided that the article is used in full and this copyright statement is reproduced. The author(s) also grant a non-exclusive licence to ascilite to publish this document on the ascilite website and in other formats for the *Proceedings ascilite Sydney 2013*. Any other use is prohibited without the express permission of the author(s).