ExamBank: a Pedagogic and Administrative System to Provide Effective Student Feedback and Stable Assessment Across Disciplines

Deborah O'Mara[a], Rosanne Quinnell[b], Imogene Rothnie[a], Leo Davies[a], and Matthew Pye[b]

Corresponding author: rosanne.quinnell@sydney.edu.au
[a]Sydney Medical School, The University of Sydney, Sydney NSW 2006, Australia
[b]School of Biological Sciences, Faculty of Science, The University of Sydney, Sydney NSW 2006, Australia

Keywords: evidence-based assessment, assessment feedback, technology-enhanced assessment processes, learning in medicine, learning in biology

International Journal of Innovation in Science and Mathematics Education *22*(3), 62-73, 2014.

Abstract

Engaging students in the effective use of assessment feedback to meet learning objectives is critical. ExamBank is a software tool developed by the Sydney Medical School (SMS) to manage the assessment process for high-stakes and formative examinations from item and examination creation to statistical reporting and the delivery of student feedback. ExamBank has been implemented in four medical schools in Australia and overseas and in other faculties at The University of Sydney, including The School of Biological Sciences. ExamBank tracks the assessment lifecycle from creation of draft items through peer review and approval to performance in multiple examinations over time. The web-based interface means ExamBank can be accessed by academics remotely via a secure login system, which allows flexible role-based access for individual assessors. Questions can be meta-tagged with key curriculum information (e.g. learning objective, subject area, unit of study, year). Statistical performance indicators for each question can be stored in the database and used to audit assessments. The implementation of ExamBank in two faculties at The University of Sydney is described to illustrate how a technology-enabled reporting system enables academics to improve the quality of assessments and the resulting improvements in curriculum design, implementation and administration and in feedback to students. This study is the first to describe the use of an item banking system for improvement of tertiary academic assessments in medical and biological science degrees in Australia.

Introduction

Constructive alignment is a curriculum design approach that seeks to optimise the learning environment in order to encourage students, through assessment and learning activities, to adopt effective learning strategies that are aligned with learning objectives (Biggs 2003). These learning objectives are aimed at developing the desired attributes of graduates. Once curricula have been designed and implemented, the assessment process is critical to determining whether the learning activities have been effective in enabling students to develop and demonstrate mastery, not only of the course content, but in those skills and professional attributes considered critical to practice in the discipline and/or profession. Assessments and feedback to students are essential components of communication between teachers and students and direct learning.

A substantial number of higher education courses use multiple choice assessment formats for all or part of their assessment of students (DiBattista & Kurzawa, 2011). Science and high-

stakes accredited courses such as medicine use multiple choice questions (MCQ) to a greater degree than other faculties and in the social sciences, the multiple choice assessment format is not favoured (Heron & Lerpiniere, 2012). Recent studies have called on universities to replace extended prose essays in final examinations with online MCQ assessments where there is focus on student feedback (Muldoon, 2012; University of Edinburgh Assessment Futures Task Group, 2011).

The quality of multiple choice assessments is of prime concern to educators, particularly in high stakes courses such as medicine (Norcini, Anderson, Bollela, Burch, Costa, Duvivier, Galbraith, Hays, Kent, Perrott & Roberts, 2011). Because of the level of dependency on this assessment mode, the training of assessment writers is critical. This includes using clinical vignettes, writing only plausible options (Naeem, van der Vleuten & Alfaris, 2012; van der Vleuten, Schuwirth, Muijtjens, Thoben, Cohen-Schotanus & van Boven, 2004), blueprinting assessment items to the curriculum to ensure close alignment, and expert review of items prior to use to improve the quality of assessment items (Tarrant & Ware, 2012). Collating item statistics such as the point-biserial correlation, facility and Rasch difficulty levels together with longitudinal analysis of an item's performance over time are psychometric methods that can be used to evaluate the quality of items (DiBattista & Kurzawa, 2011; Hohensinn & Kubinger, 2011).

The impact on learning of providing students with feedback from assessments has been evaluated in many studies across multiple disciplines including accountancy (O'Connell Ferguson, De Lange, Howieson, Watty, Carr, Jacobson, Campitelli, Gora & Milton, 2010), history (Sendziuk, 2010) and law (Butler, 2011) as well as medicine (Coombes, Ricketts, Freeman & Stratford, 2010). However, most previous research concerns written open-ended assessment tasks or assessment in practical sessions such as clinical assessments for medical students (Garry & Stirling, 2012; Moss, Derman & Clement, 2012). Some research does exist on the effect of feedback in multiple choice assessments conducted in the latter years of medical programs (Coombes et al., 2010; Bekkink, Donders, van Muijen & Ruiter, 2012).

The Sydney Medical School (SMS) has developed an assessment software system "ExamBank" which addresses many of the concerns raised about multiple choice assessments. ExamBank has been designed to facilitate the monitoring and improvement of assessment processes and practices. Although it was originally designed for a medical school, it has been successfully modified for the School of Biological Sciences at The University of Sydney, is currently being considered for adoption by other faculties and has been implemented in three overseas medical schools.

Design of ExamBank for the Sydney Medical School

ExamBank was originally developed by the SMS as an item banking system for the Royal Australasian College of Surgeons. The assessment team at the SMS used ExamBank for the first time in 2007 and the system was substantially redeveloped during 2009-2010. As it stands now, ExamBank is a web-based tool that allows the creation and editing of assessment items, creation and export of examinations and psychometric reporting on items, examinations and student performance. The main alternatives to ExamBank are commercially available databases intended for creating, delivering, analysing and reporting on assessments. *ExamSoft* is a US product that holds all information about questions and student scores on the proprietor's own server. Speedwell is a UK based product for exam delivery whose clients include medical schools and post-graduate medical colleges. This software is purchased in

modules, one of which is *QuestionBank*. Another US product, *Questionmark* is used by a range of commercial organisations, academic institutions and faculties for surveys, quizzes, tests and exams. These alternative products have been developed for large-scale school-based assessments and have been adapted for use at universities. There is little flexibility in the format of the products and the focus of many companies is in promoting their add-on bureau services. The International Consortium for Assessment Networks (ICAN), developed out of the University of Heidelberg for medical education, has the most commonality with ExamBank. This is a not for profit organisation that offers examination management software for medical schools and other health faculties and currently has 50 partners. It is specifically designed for medical and health education and includes an item management system that involves some sharing of items across partners as well as other modules. Again, they host the database for all users. All the available products have a search engine display that is worksheet based with limited hierarchical tagging. ExamBank has been designed with capabilities for multiple tagging with logical hierarchies relevant to a university-based curriculum. Non-competitive products have a facility to include Rasch statistics in the reporting of items. The main disadvantages of these competitors are that most host the data on proprietary servers outside of the client institutions control and/or require the user to fit their system and include sharing of questions with other institutions.

Question Items

ExamBank supports multiple question types; Single Best Answer (SBA) Type A questions (also known as MCQs), Extended Matching Type R questions, Short answer questions (MEQs) and Observed Structured Clinical Examinations (OSCEs). Questions are entered directly by academics from anywhere they have web access. Once entered, the author and other ExamBank users with sufficient levels of access can edit and improve them. At the lowest level of access authors have the right to browse the questions they have entered and see performance statistics (e.g. Rasch difficulty indices) for their own items as they accrue. Figure 1 shows a screen image of a created item. Each item in the bank stores the question itself, together with a variety of associated metadata including:

- Author information
- Systems and Disciplines linked to the question
- Copyright owner
- Unit of study
- Related learning objectives
- Related resources (images, text, etc.)
- Tracking information about who entered the question, who has edited and reviewed it
- Exam usage statistics
- Item statistics (once the question has been used in an assessment)

ExamBank can be programmed to make certain metadata fields mandatory such as the learning objective, thus improving blueprinting of assessments. In SMS, assessments are mapped to the online curriculum to ensure curriculum alignment is made evident. Existing banks of items can be batch imported from a word document with minimal programming. As the question is reviewed its status will change from ‘draft’ to ‘active’, and then to ‘locked’, once it has been used in an assessment. Through the process of question evolution, ExamBank has an inbuilt system of quality control. A review committee resets a menu box when a question’s status is changed from draft to active. It has been the experience of SMS that the question review process is a positive and collegial activity that not only improves item quality but also educates faculty about curriculum interactions.

In addition to the question types described above, the School of Biological Sciences has begun uploading and meta-tagging short and extended answer questions (i.e. open ended response questions) used in final examinations along with the associated marking schemes of these questions. This practice will extend the use of ExamBank to be a central archive of examinations across the School from first year, where the examinations are comprised largely of multiple choice questions, to second and third year where the examinations are comprised largely of short answer and long answer open ended response questions.

Security

ExamBank has a layered security model, with four levels of access which are controlled by the designated administrative manager for each installation

- Author level gives the user access only to questions they have created.
- Discipline level gives the user access to all questions within a given discipline. This can be granted as read only or read/write.
- Global level gives access to all questions and exams. This can be granted as read only or read/write.
- Administrator level gives access to all questions and exams, as well as to additional administrative functions.

The different levels of access control minimize the amount of leakage if a user account is compromised. An institution or faculty can limit the number of users with access to the whole database to the smallest possible number of assessment coordinators. ExamBank enables each faculty to define the security limits of questions and exams and is integrated with the University's lightweight directory access protocol authentication system (or LDAP) and requires users located off-site to log in under the protection of a virtual private network (or VPN).

Figure 1. School Medical Science ExamBank: Sample question view.

Examinations

ExamBank allows examination committees to work together online to build an exam. The system allows the academics responsible for the curriculum to create an examination blueprint, which is a template for exam creation. The blueprint ensures that the exam contains the appropriate number of questions and covers the appropriate course and discipline areas.

The blueprint defines how the selection of questions for the exam will be allocated. ExamBank creates a collection of 'Shopping Carts', corresponding to the constraints set in

the blueprint, that allow the exam creators to fill them with an appropriate set of questions in their area of expertise. Once all the carts are filled, they are combined to create an exam. Figure 2 illustrates part of the Examination creation process.

Exams can be created with several relevant parameters incorporated into the blueprinting. For example, you can stipulate that an exam must include 12 questions at Stage 2 level on cardiology that have not been used in an exam for at least 12 months. Exam reports can be generated prior to finalisation to ensure that there is a balance of questions in terms of difficulty level and prior usage. Exams can be exported from ExamBank in a variety of digital formats including MS Word or PDF and html. This enables the exported exam to be formatted to meet specific university examination centre requirements or to be exported to other platforms such as Blackboard for use as an online formative assessment.

ExamBank has a 'Check Exam' facility that enables secure online checking of an examination by the responsible academics. This final process ensures that the answer key stored within the system is correct, typographic errors are corrected, there is an adequate and balanced representation of the course learning objectives and that the items have a spread of difficulty levels and measure higher order thinking rather than factual recall. There is an option to randomise the order of questions or to force a specific pre-specified question order.

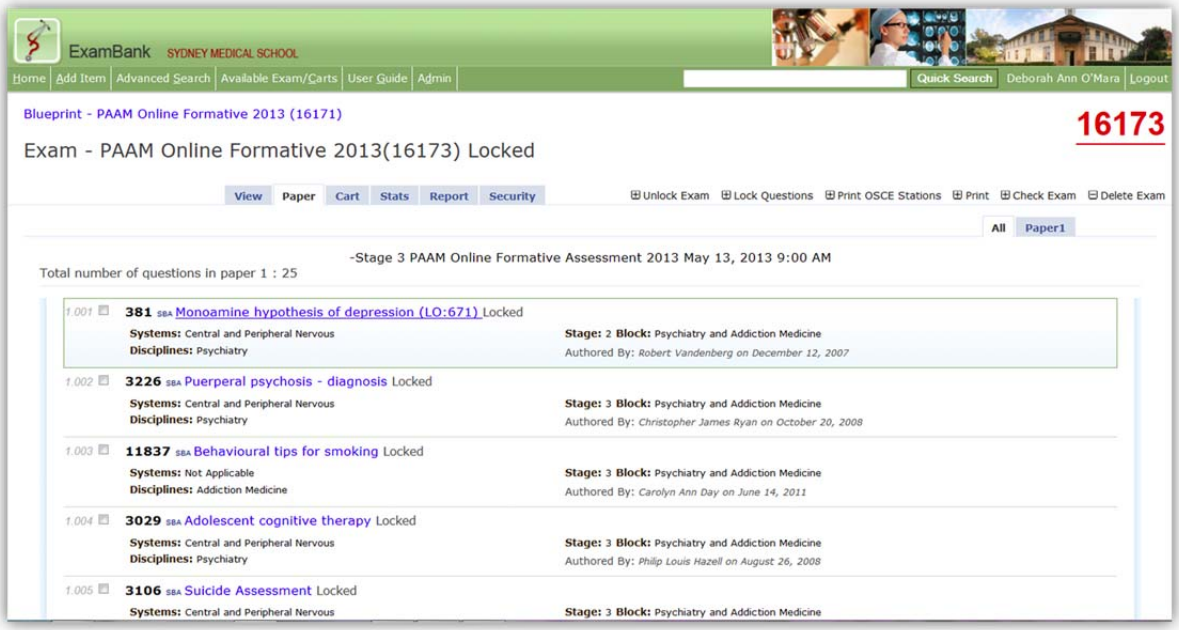


Figure 2. SMS ExamBank: Sample exam view

ExamBank Functionality

Search

ExamBank has a powerful search capability that allows searching for text or via the relational meta-data in the database. This enables exam creators to find appropriate questions in the bank. There are three levels of searches; simple, advanced and extended. This high-level function allows specific search parameters such as the date of prior usage of a question and the difficulty level as well as curriculum content parameters. Almost any combination of meta data may be combined to create complex searches. Any search can be exported as a comma separated variable file (CSV), which can be used to fine tune question selection for expert review or inclusion in an examination.

Media storage

ExamBank integrates with MediaBank, a media repository system developed at Sydney Medical School. This allows the upload of a variety of media including images, PDF format, documents, videos and audio files. When media are attached the system allows entry of associated meta data so that the media files can be categorised, searched and shared between questions. ExamBank can also integrate with existing media collections, allowing question creators to search for the best images in their own libraries and link these to a question.

Integrity

To determine whether a question should be used in an exam, it is crucial to know if and when it has been used before, and how it has performed in previous exams. Without this information, institutions may fall into the trap of overusing questions, which encourages bad student behaviour such as sharing remembered items. ExamBank allows question performance statistics to be uploaded after the exam results have been analysed, and this builds a picture of how a question has been used over time, and how effective it has been. Sudden drift in a question's performance suggests that either the curriculum has changed or the item's security has been compromised. Prior usage statistics enable exam questions to be used as anchors (or 'internal standards') to evaluate changes in overall examination difficulty and/or cohort performance over time.

Reports

ExamBank has several levels of reporting. As noted earlier, the output of a search can be saved as a CSV report. Reports can be produced on past or developing examinations, with an option to report on the prior statistics of a question for the most recent usage or for all prior usages.

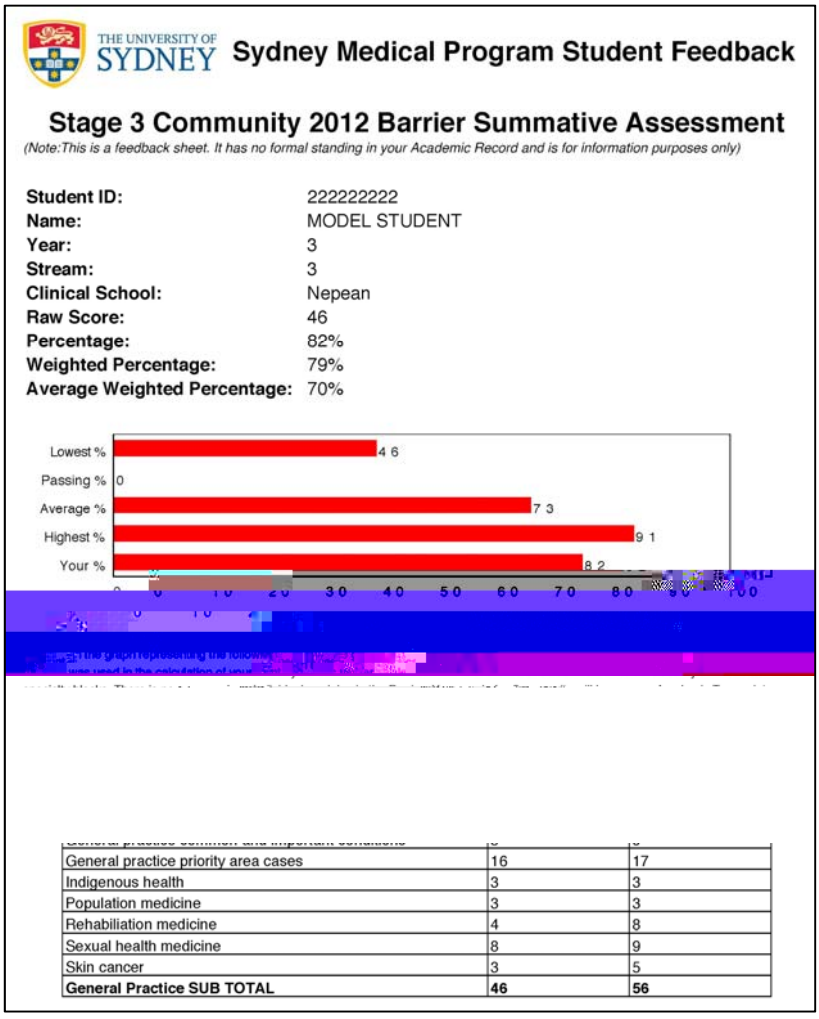
With respect to providing students with feedback, individual student reports are created by uploading a CSV file with student responses for each item in an assessment. A student feedback report is then created that summarises the number of items the student answered correctly categorised by the curriculum blueprint for that examination. The report output is driven by the meta-tagging of the questions included in the examination as well as the constraints of the examination. The student report can include summary statistics such as the average score, the passing score, the lowest score and the highest score as well as the individual student score (Figure 3). There is no passing score on the example shown as it was one component of a larger examination that had an overall passing score, illustrating ExamBank's flexibility in reporting.

Using ExamBank in a generalist degree in Biology

Alignment with Threshold Learning Outcomes

Although Biology and Medicine have a considerable overlap with respect to content, the learning objectives of a science degree with a major in Biology are very different from the learning objectives of a medical degree. This difference is evident in the documentation from the Office of Learning and Teaching (OLT, formerly the Australian Learning and Teaching Council (ALTC)), which has recently defined the threshold learning outcomes (TLO) for discipline clusters. Medicine is clustered with other health professions, including veterinary sciences and here the TLO are defined around clear capabilities and professional behaviours that a student must demonstrate during their studies (Henderson, O'Keefe & Pitt, 2011).

The ALTC funded Biology educators network (Vision and Innovation in Biology Education (VIBE.net)) has adapted these TLO to contextualise them for biology educators and biology graduates (Figure 4). In Biology, the capabilities of a graduate are defined along the lines of demonstrating an integrated understanding of the knowledge content and being able to design, conduct and critique experimental protocols. It is not surprising that the Science and Biology TLO (Figure 4) frequently use the verbs, 'analyse', 'explain' and 'apply' which are ascribed to the relational and extended abstract levels of Biggs and Coulis' Structure of the Observed Learning Outcome (or SOLO) taxonomy (Biggs & Coulis, 1982). These are the higher order thinking levels appropriate for a graduate. However, this begs the question of how best to design an assessment structure for a generalist degree that enables assessments to be mapped onto the TLO and for student to be able to monitor their progress toward meeting these academic standards.



THE UNIVERSITY OF SYDNEY **Sydney Medical Program Student Feedback**

Stage 3 Community 2012 Barrier Summative Assessment

(Note:This is a feedback sheet. It has no formal standing in your Academic Record and is for information purposes only)

Student ID:	222222222
Name:	MODEL STUDENT
Year:	3
Stream:	3
Clinical School:	Nepean
Raw Score:	46
Percentage:	82%
Weighted Percentage:	79%
Average Weighted Percentage:	70%

General practice priority area cases	16	17
Indigenous health	3	3
Population medicine	3	3
Rehabiliation medicine	4	8
Sexual health medicine	8	9
Skin cancer	3	5
General Practice SUB TOTAL	**46**	**56**

Figure 3. Sample student report from Sydney Medical School ExamBank

Meta tagging and curriculum alignment

The School of Biological Science at the University of Sydney adopted ExamBank in 2011 and piloted it in the two 2012 Summer School units of study, Concepts in Biology and Human Biology. The School's student feedback process uses the 'Fast Personalised Feedback' system (Bridgeman & Rutledge, 2010) and this speaks to the ability of the ExamBank system to integrate into other automated feedback systems.

Currently the School of Biological Sciences' implementation of ExamBank is constructed around a curriculum largely defined by content. Meta-tags describe year of study, unit of study, content area (knowledge topics and sub-topics; Figure 5) and skill topics and sub-topics (The Practice of Science; Figure 6). To accommodate the integrative nature of Biology, multiple meta-tags can be selected. These meta tags were established prior to the publication of the OLT Academic Standards report for Science. There is, however, a level of alignment between the existing meta-tags and the TLO of a Biology graduate as both emphasis knowledge and skills.

Upon completion of a bachelor degree or major in biology, graduates will:	
Understanding biology	1.1 Demonstrate a coherent understanding of biology by articulating the methods of biology and explaining why current biological knowledge is both contestable and testable through further inquiry.
	1.2 Demonstrate a coherent understanding of biology by explaining the role and relevance of biology in society.
	1.3 Recognise that biological knowledge has been acquired by curiosity and creativity and demonstrate creativity in thinking and problem solving.
	1.4 Recognise and appreciate the significant role of biodiversity in sustaining life on our planet.
Biological knowledge	2.1 Exhibit depth and breadth of biological knowledge by demonstrating a well-developed understanding of identified core concepts in biology.
	2.2 Exhibit depth and breadth of biological knowledge by demonstrating that these 'core concepts' have interdisciplinary connections both within science and other disciplines.
Inquiry and problem solving	3.1 Gather, synthesise and critically evaluate information about biological phenomena from a range of sources.
	3.2 Critically analyse observations of biological phenomena by creating and developing models and proposing and testing hypotheses.
	3.3 Design and conduct field, laboratory-based or virtual biological experiments.
	3.4 Select and apply practical and/or theoretical techniques.
	3.5 Collect, accurately record, interpret, analyse, and draw conclusions from biological data.
Communication	4 Effectively synthesise and communicate biological results using a range of modes (oral, written, and visual) for a variety of purposes and audiences.
Personal and professional responsibility	5.1 Be accountable for their own learning and biological work by being independent and self-directed learners.
	5.2 Work effectively, responsibly and safely in individual and peer or team contexts.
	5.3 Demonstrate knowledge of the regulatory frameworks and ethical principles relevant to their sub-disciplinary area within biology and apply these in practice.

Figure 4: Biology Threshold Learning Outcomes (draft) (VIBEnet, 2013) describing the key concepts and competencies (knowledge and skills) of Biology graduates. These BTLOs are based on the Science TLOs (Jones et al., 2011).

Pilot project in Biological Science

The Summer School units of study where ExamBank was piloted are offered in 'intensive mode' over a 6-week period rather than a 13-week semester and the enrolment is a subset of the main cohorts in semester 1 (n = 28 and 38, Concepts in Biology and Human Biology, respectively). Student engagement is critical as is timely feedback on assessments. As a proportion of these students are taking these units of study for the second time it is useful to have a means to precipitate a conversation about study strategies. In the Summer School pilot, students were given aggregated feedback on topic and/or skill areas. For example, the mid-session test, comprised of MCQs, covered three content topic areas and laboratory or technical skills.

After the mid-session tests were marked, the students received a test feedback summary that included their total score, the class average and their scores for each topic area covered in the test. The test scores were sent to students by email the same day as the tests. The feedback to students also included a statement as to how they could use their topic scores as indicators of their understanding of each topic. Topics where they scored poorly were flagged and students were advised to revise that topic. A critical part of this assessment strategy was allocating time in the practical class for a follow up discussion about the feedback process and the

feedback itself. Positive comments made by students in these discussion sessions suggest that most students were pleased with the feedback offered. This finding is consistent with the findings of Bridgeman and Rutledge (2010). Important to note is that this particular assessment strategy did not divulge the 'right answers', rather students are offered their individual scores on 'domains' that let them know the areas where they are doing well and where they have deficits in the knowledge. Teaching staff saw that this feedback strategy focused students on reflecting on their learning and provided direction to their individual study program. Some students were not happy that they were not given the answers to the questions in the test papers, but, because these students were quite vocal and identified themselves, this provided the unit coordinator the opportunity to discuss assessment practices and the link to the expected learning outcomes and to generic skills. A few students were not convinced that a test where all the questions were MCQ could truly reflect their understanding of the discipline. Asking first year students to see assessments as a sampling process rather than a comprehensive syllabus summary can be difficult, but it is a crucial component of creating adult learners. Indeed, the TLO for science states that a science graduate 'be accountable for their own learning by being independent and self-directed learning' (Jones, Yates & Kelder, 2011).

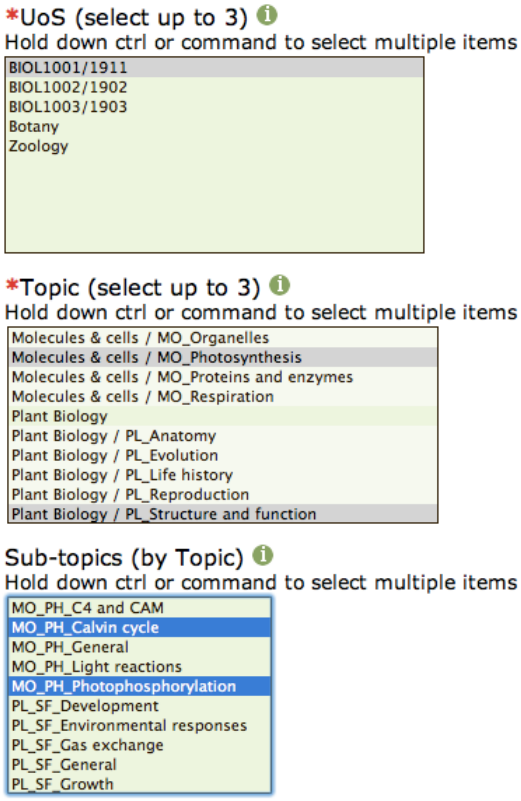


Figure 5. School of Biological Sciences question meta tags for knowledge domain and subdomains

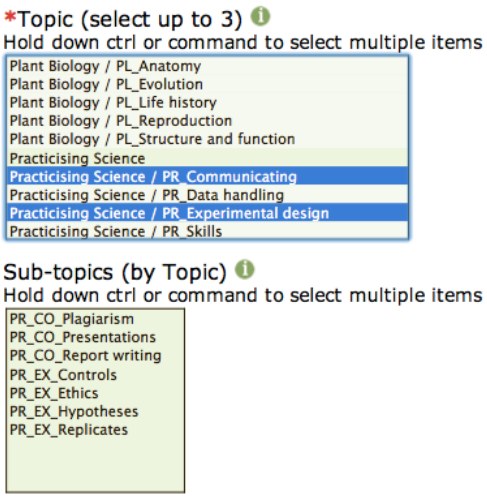


Figure 6: Science Skill meta tags in Biology

The School of Biological Sciences ExamBank is being used for managing the assessment of the School's large (n >1500) first year units of study and as an assessment archive, including marking matrices, for some intermediate and senior units of study. Our next step is to give students automated feedback that explicitly, rather than implicitly, focus on the science TLO that speak to outcomes such as *Personal and professional responsibility* as defined by the OLT (Jones et al., 2011); and, more specifically, those of a Biology graduate (Figure 4;

VIBEnet, 2013). It has to be noted that the Biology TLO are in draft form only and this is why meta-tagging according to a learning taxonomy (Blooms or SOLO) may afford a better solution rather than tagging with specific TLO that may change in the future. In Medicine, where the curriculum is somewhat less flexible than Biology, where the stakes are high and there is formal accreditation, learning outcomes are monitored at each phase and statistics are kept for each year. Currently the focus in Biology is to ensure that the use of ExamBank is supported by sound administrative and academic practices and to this end staff using the Biology ExamBank have a user wiki to record the developments, decisions and justifications of their usage of the ExamBank assessment system.

Discussion

It is critical to delivering a life sciences curriculum that students engage with the curriculum as a whole and that feedback invites students to develop effective learning strategies. The Higher Education sector insists on evidence of 'Best Practice' and our assessment practices need to be supported using both the best available statistical evidence and on pedagogical grounds.

ExamBank is an example of a software system that improves the quality of tertiary education assessments. The online accessibility allows individual academics to add assessment questions at any time or location encouraging greater participation by teachers in the assessment process. The structure of item input is such that it facilitates improvements to the quality of items through meta-tagging to the curriculum and through review by academic peers prior to exam usage. It is flexible enough to be modified for any faculty. The ability to generate reports and attach statistical information to items together with the question and examination search facilities greatly assist the creation of examinations. Reports on past exams are readily obtained. We believe that systems of this type foster academic collaboration and we encourage the wider use of ExamBank or equivalent software packages.

In part due to the high level of security provided by the bank, the SMS has increased the number of secure assessment items it holds by at least two thousand each year. It is now possible to focus on question quality and provide practice questions for students in formative online exams through Blackboard. The SMS has recently introduced a quota for new and anchor questions which enables the statistical equating of the difficulty of examinations from year to year while facilitating renewal of the question bank in line with the evolution of the curriculum.

Having a system that provides timely student feedback tied to the curriculum blueprint, the next step is to investigate how students use this feedback to improve their learning habits. Although research suggests there is variable use of feedback and it is mostly by better performing students (Sinclair & Cleland, 2007), there is widespread support for the provision of student feedback from assessments. The importance of feedback to medical students from examinations was highlighted in the most recent edition of *Academic Medicine* (Sharma, 2013). Since the introduction of the examination feedback summary sheet produced by ExamBank, medical students expectation for feedback have grown. While the information in the ExamBank student feedback is of immediate diagnostic relevance to students sitting supplementary examinations, other students report using it for applications for scholarships and future internships.

By linking assessment items to curriculum, institutions are able to ensure that their assessment program has adequate curriculum coverage and ExamBank has proved invaluable

in helping to identify gaps in assessment; In the post exam review of questions those that performed poorly reflect gaps in the syllabus while those that have a good prior track record nearly always reflect an unrecognized change in the syllabus. There are few published studies about item banking for tertiary education. We identified a feasibility study for an item bank relating to technical skill training in the USA (Derner, Klein & Hilber, 2008). Internet searches for item banks most commonly produce results limited to commercial software companies, even if the search is limited to medical education. Our study is the first report to describe item banking for medical education and the biological sciences in Australia.

ExamBank has enabled the School of Biological Sciences to implement an innovative assessment process that will improve student learning and develop the capacity of the faculty to provide accountable and improved assessments. The current meta-tags in the Biology ExamBank, and the TLO of Biology graduates, focus on demonstrating a coherent understanding in Biology, Biological knowledge and the practical skills of Biology. We need to make explicit to students that part of the learning and teaching agenda is to address 'personal and professional responsibility' (Science TLO 5.1 - 5.3). This explicit nature is considered critical for effective online feedback (Hepplestone, Parkin, Irwin, Holden, Thorpe & Burn, 2010).

As stated by Gibbs and Simpson (2004), 'Standards will be raised by improving student learning rather than by better measurement of limited learning.' Using item banks it is possible to improve how we craft exams and improve the feedback we give to students by including meta-tags that explicitly align to the TLO or, better still, a learning taxonomy (as has been described by Stranger-Hall, 2012) that allows students see progress towards the learning outcomes of a graduate. It is possible that ExamBank will allow us to map assessments with respect to the TLO by meta-tagging items with not only the TLO but with the learning taxonomy level (e.g. SOLO). It will be interesting to see how the student perception of difficulty aligns with our expert views defined by statistical difficulty indices and whether this alignment gets closer over the course of a degree. Certainly, systems such as ExamBank will allow us to be more rigorous and accountable for our assessment practices.

Acknowledgements

We acknowledge ICT support at the level of the University and at the level of the School (both School of Biological Sciences and Sydney Medical School).

References

Bekkink, M.O., Donders R., van Muijen, G.N.P., & Ruiter D.J. (2012). Challenging medical students with an interim assessment: a positive effect on formal examination score in a randomized controlled study. *Advances in Health Sciences Education, 17*(1), 27-37.

Biggs, J.B. & Coulis K.F. (1982). *Evaluating the quality of learning: the SOLO taxonomy (structure of the observed learning outcome).* Academic Press, New York.

Biggs J. (2003). Aligning teaching and assessing to course objectives. Teaching and Learning in Higher Education: New Trends and Innovations Apr 13-17; University of Aveiro (2003. Retrieved August 2, 2013, from: http://e-bit.lk/wiki/images/2/2c/CA1.pdf.

Bridgeman, A. & Rutledge, P. (2010). Getting personal: feedback for the masses. *Synergy* 30, 61-68. Retrieved August 2, 2013, from: http://www.itl.usyd.edu.au/synergy/30/index.html?pageNumber=61

Butler, D. (2011). Closing the loop 21st century style: providing feedback on written assessment via MP3 recordings. *Journal of the Australasian Law Teachers Association, 4*(1), 99-107.

Coombes, L., Ricketts, C., Freeman, A. & Stratford, J., (2010). Beyond assessment: feedback for individuals and institutions based on the progress test. *Medical Teacher, 32*(6), 486-90.

Derner, S., S. Klein, & D. Hilber. (2008). Assessing the Feasibility of a Test Item Bank and Assessment Clearinghouse: Strategies to Measure Technical Skill Attainment of Career and Technical Education Participants. MPR Associates, Inc.

DiBattista, D. & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *Canadian Journal for the Scholarship of Teaching and Learning, 2*(2), 23.

Garry, A. & Stirling, K. (2012. Achieving 360 student feedback using SPaCE. The clinical teacher: 9(4), 222-7.

Gibbs, G. & Simpson, C. (2004). Conditions under which Assessment supports Student Learning. *Learning and teaching in higher education, 1*, 3-31

Henderson, A., O'Keefe, M. & Pitt, R. (2011). Learning and Teaching Academic Standards Project Health, Medicine and Veterinary Sciences Academic Standards Statement. Australian Learning and Teaching Council, Strawberry Hills, NSW. Retrieved August 6, 2013, from http://www.olt.gov.au/resources.

Hepplestone, S., Parkin, H., Irwin B., Holden, G., Thorpe L. & Burn, C. (2010). Using technology to help students engage with their feedback. A best practice guide for academic staff. Sheffield Hallam University. Retrieved August 6, 2013, from http://www.enhancingfeedback.ed.ac.uk/documents/guide+for+academic+staff+FINAL.pdf

Heron, G. & Lerpiniere J. (2012). Re-engineering the multiple choice question exam for social work. *European Journal of Social Work.* [Electronic version]. DOI:10.1080/13691457.2012.691873

Hohensinn, C. & Kubinger, K.D. (2011). Applying Item Response Theory Methods to Examine the Impact of Different Response Formats. *Educational and Psychological Measurement, 71*(4), 732-746.

Jones S., Yates B. & Kelder J. (2011). Science Learning and Teaching Academic Standards Statement. Australian Learning and Teaching Council. Strawberry Hills, NSW. Retrieved August 6, 2013, from http://www.olt.gov.au/resources.

Moss, H.A., Derman P.B., & Clement, R.C. (2012). Medical student perspective: working toward specific and actionable clinical clerkship feedback. *Medical Teacher, 34*(8), 665-7.

Muldoon, R. (2012). Is It Time to Ditch the Traditional University Exam? *Higher Education Research and Development, 31*(2), 263-265.

Naeem, N., van der Vleuten C., & Alfaris E.A. (2012). Faculty Development on Item Writing Substantially Improves Item Quality. *Advances in Health Sciences Education, 17*(3), 369-376.

Norcini, J., Anderson B., Bollela, V., Burch, V., Costa, M.J., Duvivier, R., Galbraith, R., Hays, R., Kent, A., Perrott, V. & Roberts, T. (2011). Criteria for good assessment: consensus statement and recommendations from the Ottawa, 2010 Conference. *Medical Teacher, 33*(3), 206-14.

O'Connell, B., Ferguson C., De Lange P., Howieson, B., Watty, K., Carr, R., Jacobson, B., Campitelli, L., Gora, Y. & Milton A. (2010). Enhancing Assessment Feedback Practices in Accounting Education: Issues, Obstacles and Reforms. Australian Learning and Teaching Council Ltd, Strawberry Hills: Australia. Retrieved August 2, 2013, from http://www.olt.gov.au/system/files/resources

Sendziuk, P. (2010). Sink or swim? Improving student learning through feedback and self- assessment. *International Journal of Teaching and Learning in Higher Education, 22*(3), 320-330.

Sharma, N. (2013). The Importance of Real-Time Feedback in Undergraduate Assessments, *Academic Medicine,* Letter to the Editor, 88(11), November 2013, 1593.

Sinclair, H.K. & Cleland, J.A. (2007). Undergraduate medical students: who seeks formative feedback? *Medical Education, 41*(6), 580-2.

Stranger-Hall, K.F. (2012). Multiple-choice exams: an obstacle for higher-level thinking in introductory science classes. *CBE Life Science Education, 11*, 294-306.

Tarrant, M. & Ware, J. (2012). A framework for improving the quality of multiple-choice assessments. *Nurse Educator, 37*(3), 98-104.

University of Edinburgh Assessment Futures Task Group. 2011. Extended prose examinations and word processing: What's to be done? Report. Retrieved August 7, 2013, from http://www.ltc.vet.ed.ac.uk/docs/open/ExtenddProse_WP_consultn_(2).pdf

van der Vleuten, C.P.M., Schuwirth, L.W., Muijtjens, A.M., Thoben, A.J., Cohen-Schotanus, J. & van Boven, C.P. (2004). Cross institutional collaboration in assessment: a case on progress testing. *Medical Teacher, 26*(8), 719-25.

VIBEnet: Vision and Innovation in Biology Education. (2013). Biology Threshold Learning Outcomes (draft). Retrieved August 7, 2013, from https://sites.google.com/site/vibenet101/home/draft-btlos