Assessment & Support Team
Education Infrastructure Division
Education Bureau
The Government of the Hong Kong Special Administrative Region

Assessment & Learning
評估與學習

Issue 2
2013

Assessment and Learning

評估與學習

Issue 2 November 2013
第二期 二零一三年十一月

Published by

Assessment & Support Team
(Formerly Basic Competency Assessment Team)
Education Infrastructure Division
Education Bureau
The Government of the Hong Kong Special Administrative Region
The People's Republic of China

Room 3602, 36/F, Hopewell Centre
183 Queen's Road East, Wanchai
Hong Kong

ISSN: 2305-7114

Contents 目 錄

English Language

Mathematics

Acknowledgements

We wish to express grateful thanks to the following scholars, researchers, principals and teachers for their kind contributions to this publication:

Professor Maurice Galton

Emeritus Professor, and Senior Research Fellow in the Centre for Commonwealth Education, University of Cambridge

Professor Lorna M. Earl

Professor, The University of Auckland, and Director of Lorna Earl & Associates, Toronto, Canada. Formerly Associate Professor, University of Toronto, and the first Director of Educational Quality Assurance Office (EQAO) in Ontario, Canada

Professor William P. Fisher, Jr.

Research Associate, BEAR Centre, Graduate School of Education, University of California, Berkeley, and Principal, LivingCapitalMetrics.com

Professor Peter Tymms

Director of iPIPS, Centre for Evaluation and Monitoring (CEM), Durham University

Dr Qingping He

Principal Researcher, Office of Qualifications and Examinations Regulation, Government of United Kingdom, and Expert Consultant, CEM Hong Kong, Durham University

Mr See Ming Tsui

Director of Centre for Evaluation and Monitoring (CEM) Hong Kong, Durham University

Professor Magdalena Mo Ching Mok

Chair Professor of Assessment and Evaluation, Department of Psychological Studies, and Co-Director of Assessment Research Centre, The Hong Kong Institute of Education

Ms Jingjing Yao

PhD Candidate, Assessment Research Centre, The Hong Kong Institute of Education, and Lecturer of Department of Psychology, Zhejiang Normal University

Dr Rita Berry

Research Fellow and formerly Associate Professor, The Hong Kong Institute of Education

莊紹文博士

荔景天主教中學 校長

林偉業博士

香港大學 教育學院 助理教授

梁佩雲博士

香港教育學院 中國語言學系 副教授

房湘雲老師

元朗天主教中學 老師；前爲 香港教育局 校本專業支援組 借調老師

Dr Seok Moi Ng
Consultant (English Language), Ministry of Education, Singapore, and formerly Senior Lecturer, The University of Auckland

Dr Arthur McNeill
Director of Center for Language Education, and Associate Dean of School of Humanities and Social Science, The Hong Kong University of Science and Technology

Professor Icy Lee
Professor, Faculty of Education, The Chinese University of Hong Kong

Mr Sean McMinn
Senior Instructor, Center for Language Education, The Hong Kong University of Science and Technology

Dr Zoe So Yee Chan
English Panel Chairperson, Lingnan University Alumni Association DSS Primary School

Mr Arthur Man Sang Lee
Teaching Consultant, Faculty of Education, The University of Hong Kong

In the order of appearance in this publication.

Preface

This is the second volume in the series, *Assessment and Learning,* from the Assessment & Support Team (formerly the Basic Competency Assessment Team). Compared to the previous publication this contains more papers from internationally known authors and it says much for their growing reputation in the field of assessment that team is able to call on such distinguished contributors. Teachers may not find some papers an easy read but if they persevere they will be rewarded by many rich insights into the increasingly complex world of educational assessment.

Nearly half the papers deal with student weaknesses in Languages and Mathematics and the means of identifying and dealing with these. One particularly provocative piece concerning extended writing asks whether it is always necessary to pick out and correct all errors or only the consistent ones. On more general assessment issues, two other strong themes emerge from the collection. In the first instance, the argument is put forward that for Assessment *of* Learning empirically constructed tests (the kind that require initial piloting , checking for internal consistency and item facility etc.) are time-consuming, costly to develop and can be problematic in terms of their predictive validity. It is suggested that in their place 'theory driven' tests, such as those based on the Rasch model, should become more readily available to teachers for use as both a formative and summative tool. Second, the need to extend the notion *of* Assessment *for* Learning (AfL) to embrace Assessment *as* Learning (AaL) is emphasised by several contributors based on the research evidence that feedback which promotes *forward* thinking by students pays rich dividends in future attainment gains.

There is much to be said for utilising a linear measure of ability which is dependent only on the difficulty of the item, in much the same way that in an old fashioned mercury thermometer the temperature of an object is proportional to the height of the liquid column. But there is no corresponding simple visual

effect in the use of Rasch scaling. One paper, for example, includes tables where the ability scales have minus scores and this concept may not be easy to understand when presented, for example, to the typical primary teacher. In an age of computer graphics that can create extraordinary 3-D representations it should be possible to produce pictorial images which simplify the presentation and the analysis for those classroom practitioners, who may lack the necessary mathematical ability to understand and therefore interpret the results in their present form. On the other hand, it could be argued that having teachers base their judgements about a student's ability on procedures they do not fully understand can set a dangerous precedent. A parallel exists in the field of educational research where it is possible for anyone who can construct a data matrix using Excel to carry out sophisticated statistical analyses using a package such as SPSS without any understanding of the underlying assumptions of the procedures and their limitations and sometimes with disastrous consequences. It is my hope that future issues will pursue this debate further.

Another issue where further research on theory driven tests is needed concerns construct validity, particularly if the analysis of data using the Rasch methodology is to be used for diagnostic purposes. At the moment there is a tendency for those who construct such tests to rely heavily on face validity and it is sometimes not easy to discern why some items fit on a uni-dimensional scale while others do not. Many years ago I was involved in a project attempting to measure the effects of different teaching styles in science on student attainment[1]. Our hypothesis was that teachers who adopted an enquiry approach, as opposed to a didactic problem solving one, would enhance students' scientific reasoning. We therefore set about constructing tests designed to measure higher order skills such interpreting data, making inferences from data, and formulating hypotheses. But to employ these skills required students to have an understanding of the theories behind the problems we set in the written tests, so we were unsure when a student gave a wrong answer whether it

[1] Eggleston, J., Galton, M. and Jones, M. (1976). *Processes and products of science teaching*. London: Macmillan (for the Schools Council).

was because s/he lacked a theoretical grasp of the principle or whether they were unable to apply the theory to the particular question. One of us suggested we should therefore include a brief account of the theory in the stem of the question which we did, for example, in one case providing a brief summary of the laws of reflection. But then we discovered that our consistency measure was distinguishing between the candidates' ability to comprehend the content of the question stem and not their problem solving skill.

In such cases the answer lies in interrogating the students further in order to discover how they arrived at their respective answers and in time helping them to conduct their own self-evaluations. This form of questioning lies at heart of both AfL and even more importantly AaL and raises the question of where our efforts to reform current assessment practice should be directed in the immediate future. Is it to be educating teachers to make regular use of procedures such as Rasch for formative testing purposes or should we concentrate on the immense task of changing the public (and official) perceptions as to the key purposes of assessment, particularly its use in AfL and AaL. Only by doing this will we remove the pressure on teachers to continue to use conventional testing methods. One contributor argues, and I agree, that this will require all involved in education at the grass root level to wage a 'quiet' but 'determined' revolution.

This collection of papers and the previous volume in the series are a vital component of this revolutionary process. I look forward to more of the same in future issues, for in preparing the next generations of students to face the demands of the twenty-first century there is no more important task than to convince a somewhat sceptical public that assessment involves more than setting regular written tests.

Maurice Galton

Assessment For Learning; Assessment As Learning: Changing Practices Means Changing Beliefs

Lorna M. Earl
Ontario Institute for Studies in Education, University of Toronto
and The University of Auckland

As in many countries around the world, Hong Kong educational leaders are reviewing and rethinking its assessment practices. They are exploring on-line assessment platforms as well as building teachers' assessment literacy and working to enhance the feedback provided to teachers and students in the teaching, learning and assessment cycle. *Assessment for learning* and *assessment as learning* are new concepts that have entered the conversation. As with any new ideas, there is some confusion and multiple interpretations of what these phrases mean and what they look like in practice. In this short article, I hope to clear some of the confusion and provide some insights about how *assessment for* and *assessment as learning* work in classrooms.

The idea of *assessment for learning* arose out of a 1998 landmark research paper by Black and Wiliam in which they synthesized over 250 studies linking assessment and learning and found that the intentional use of assessment in the classroom to promote learning improved student achievement. [1] This meta-analysis supported previous research showing that classroom assessment had both short- and long-term effects on learning.[2]

In the short term, it showed that classroom assessment could:

- focus attention on important aspects of the subject;

1 Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan,* Vol. 80 (2).

2 Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

- give students opportunities to practice skills and consolidate learning;
- guide further instructional or learning activities.

In the medium and long term, assessment held the possibility of:

- influencing students' motivation as learners and their perceptions of their capabilities;
- communicating and reinforcing teaching goals, including performance criteria and desired standards of performance;
- influencing students' choice of and development of learning strategies, skills, and study patterns;
- influencing students' subsequent choice of courses, activities, and careers.

Since that time, it has become obvious that assessment can be a powerful catalyst for learning.

> *Recent reviews of more than 4,000 research investigations show clearly that when [formative assessment] is well implemented in the classroom, it can essentially double the speed of student learning… it is clear that the process works, it can produce whopping gains in students' achievement, and it is sufficiently robust so that different teachers can use it in diverse ways, yet still get great results with their students.*[3]

Assessment for learning shifts the emphasis from summative to formative assessment, from making judgments to creating descriptions that can be used in the service of the next stage of learning. When they are doing *assessment for learning* teachers collect a wide range of data for a different purpose – so that they can modify the learning work for their students. They craft assessment tasks that open a window on what students know and can do already and use the insights that come from the process to design the next steps in instruction. To do

3 Popham, J. (2011). Formative assessment – A process, not a test. *Education Week*, February 2011. http://www.edweek.org/ew/articles/2011/02/23/21popham.h30.html?t

this, teachers use observation, worksheets, questioning in class, student-teacher conferences or whatever mechanism is likely to give them information that will be useful for their planning and their teaching. Considering student work is not designed to assign marks and make comparative judgments among the students but to highlight each student's understanding and thinking and provide students with feedback that will further their learning. *Assessment for learning* happens in the middle of learning, often more than once, not at the end. It is interactive, with teachers providing assistance as part of the assessment. It helps teachers provide the feedback to scaffold next steps. And it depends on teachers' diagnostic skills to make it work.

When I wrote the book **Assessment As Learning: Using Classroom Assessment to Maximize Student Learning**[4], I introduced the notion of *assessment as learning*, a subset of *assessment for learning* that emphasizes using assessment as a process of developing and supporting meta-cognition for students. *Assessment as learning* focuses on the role of the student as the critical connector between assessment and their own learning. Students, acting as active critical thinkers, make sense of information, relate it to prior knowledge, and use it to construct new learning. This is the regulatory process in meta-cognition. It occurs when students personally monitor what they are learning and use the feedback from this monitoring to make adjustments, adaptations, and even major changes in what they understand (p. 47).

Although many teachers would say that they do *assessment for* and *assessment as learning* there is considerable evidence that their assessment practice does not really reflect the intentions and principles that make *assessment for* and *assessment as learning* powerful[5]. I have come to believe

4 Earl, L. (2003, 2013). *Assessment as learning: Using classroom assessment to maximize student learning.* Thousand Oaks, CA: Corwin Press.

5 James, M., McCormick, R., Black, P., Carmichael, P., Drummond, M.J., Fox, A., MacBeath, J., Marshall, B., Pedder, D., Procter, R., Swaffield, S., Swann, J., and Wiliam, D. (2007). *Improving learning how to learn – Classrooms, schools and networks.* Abingdon: Routledge.

that “conceptual change” is at the heart of professional learning if it is to go beyond tinkering with practice. The theories of *assessment as learning* potentially conflict with teachers’ previous conceptions of the purpose of assessment and even their theories of teaching. These shifts challenge some long-standing and deeply held beliefs about what schooling is for and why teachers should collect information about how students are thinking and learning. Because *assessment as learning* is a different way of thinking about assessment, it is unlikely that many teachers will be able to just assimilate it into their practice. Incorporating *assessment for* and *assessment as learning* into practice requires a fundamental shift in the way teachers think about the nature of learning and the rhythm of interactions in classrooms.

The primary aim of *assessment for* and *assessment as learning* is not summative, for grading or reporting; it is formative, to contribute to students’ learning. That means assessment is an integral part of teaching and learning and teachers have the responsibility for identifying aspects of learning as it is developing, using both informal and formal processes, so that they, and the students can decide what to do next to enhance the learning. *Assessment as learning*, in particular, is founded on a belief that for students to become self-motivating and able to bring their talents and knowledge to bear on the decisions and problems that make up their lives, can't just wait for the teacher (or politicians, or salespeople, or religious leaders) to tell them whether or not the answer is "right". Effective assessment empowers students to ask reflective questions and consider a range of strategies for learning and for acting. Over time, students move forward in their learning when they can use personal knowledge to construct meaning, have skills of self-monitoring to realize that they don't understand something and have ways of deciding what to do to next.

Like their students, teachers need to examine their mental models, rethink their practices and develop new skills so that they are comfortable making moment-by-moment decisions on the basis of considered evidence and argument, with student learning as the goal. They need to become adaptive experts who use many effective strategies for teaching, learning and assessment but also have high levels of flexibility that allow them to innovate when

routines are not enough, in order to ascertain when students are not learning, know where to go next, and adapt resources and strategies to assist students to meet worthwhile learning goals[6].

Understanding the power of *assessment for* and *assessment as learning* and changing assessment practices mean hard new learning for teachers because the shifts require fundamental rethinking of their current beliefs and their existing practice. But, the results are worth it.

Author's e-mail: lorna.earl@utoronto.ca

6 Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London, England: Routledge.

Imagining Education Tailored to Assessment As, For, and Of Learning: Theory, Standards, and Quality Improvement

William P. Fisher, Jr.
University of California, Berkeley

Abstract

The demands of a new economic and technological age require leaders in education to awaken to new challenges, and learn to manage new responsibilities. Assessment will play a key role in meeting these challenges and in taking on these responsibilities. Assessment will not, however, be implemented in just the traditional sense of assessing learning for accountability purposes but will itself become a medium embodying and setting the stage for learning. Educators have made distinguished contributions in aligning learning progressions with assessments for use in self-directed individualized instruction. Highly technical principles and mathematical operations have been made accessible to end users who do not have the training or skills needed to directly employ those principles and operations themselves. Accordingly, the time has come to draw out some of the developments likely to follow from or impact the technical advances in assessment as, for, and of learning that are being put in practice. Three areas have the greatest promise for the future. They are theory development, standardized units of measurement, and systematic quality improvement methods. Putting these in place around self-directed learning-oriented assessment practices offers real hope for significant positive changes in education.

Introduction

We are in a new economic and technological age. Leaders in education who manage school systems and individual schools must awaken to new challenges, and to being accountable for new responsibilities. They must invent new educational outcome products, and, most importantly, they must match the supply of those products with growing demand in the world for responsible citizens with productive skills.

Teamwork and collaboration are increasingly required in industries that thrive on innovation. Schools need to teach more than just basic and critical thinking skills, they need to foster emotional intelligence and social skills, too. Schools must provide environments in which all of these skills are incorporated into day-to-day practices.

Passive reading and listening to lectures are far less effective for learning than self-directed participatory involvement and dialogue. Because they can support the latter with highly effective tools, online distance education and e-learning methods are being shown just as or more effective for learning as traditional classrooms.

Already today, the school a student attends matters less than what that student knows and can do. Leaders in education should be doing everything they can to make the market for educational outcomes as efficient as possible. Schools are partnering with employers, linking educational and training opportunities with recruitment efforts.

Assessment is playing a key role in meeting these challenges and in taking on these responsibilities. The changes taking place are paradigm-shifting (Gipps, 1994; Shepard, 2000). Assessment is not, however, being implemented in just the traditional sense of assessing learning for accountability purposes like grades, graduation, admissions, certification, or licensure. Instead, assessment has itself become a medium embodying and setting the stage for learning

(Black & Wiliam, 2009; Cheng & Mok, 2007; Earl, 2013; Feng, Heffernan, Koedinger, 2009; Ma, 2012; Mok, 2011; Wilson, 2009). Implementations integrating assessment and instruction in the classroom build on the large effect sizes observed in research on what works in education (Hattie & Timperley, 2007).

The research aligning learning progressions with assessments for use in self-directed individualized instruction is complex and technical. Putting the right methods and tools in the hands of the teachers and students – methods like self-directed learning strategies and tools like the S-P Chart, Item-Person Map, Kid-Map, and Unexpected Persons Map – will require innovative approaches but will have lasting effects (Mok, 2011).

The accomplishment is one in which highly technical principles and mathematical operations have been made accessible to end users who do not have the training or skills needed to directly employ those principles and operations themselves. The philosopher, Whitehead, observed that “Civilization advances by extending the number of important operations which we can perform without thinking about them.” Indeed, everyday tools like telephones, computers, and automobiles are now so complex that even engineering experts do not have the range of knowledge needed to master all of the component parts in a single device.

The question arises, then, as to what developments might follow from or further impact the technical advances in assessment as, for, and of learning that are being put in practice. There are three areas in which research and development seem to me to have the greatest promise for the future. The three substantive areas that will have a direct bearing on self-directed learning-oriented assessment are:

- Theory, which allows for greater efficiency in item development, test assembly, adaptive administration, instrument calibration, and measure estimation.

- Standardized units of measurement, which will facilitate universal comparability and maximize the involvement of all stakeholders (teachers, colleagues, students, parents, the community at large, employers, researchers, etc.) in the outcomes produced.
- Quality improvement methods that systematically implement the common metrics in a variety of environments, building up from the students' self-directed learning to self-directed improvement efforts at weekly faculty meetings, at quarterly parent-teacher meetings, at district-wide teacher conferences, etc.

In these remarks, advances in computer hardware, software and networks are taken for granted as the mechanisms by which we will have ever faster connections and processing speeds. What are not obvious and cannot be taken for granted, however, are the paths along which theory, standardized units of measurement, and quality improvement methods will develop. Though these paths can be sketched now, their full shape and direction remain to be determined. The best way to predict the future is to invent it, and so the best place to begin is where integrated assessment and instruction is now.

Assessment As and For Learning

Where does education begin? The most important realization about education is that we learn through what we already know. That's why early education focuses on language and numbers. Alphabets, characters, words, grammar, numbers, symbols, phonemes, etc. are the media of learning, so they, along with basic social and self-management skills, have to be learned before anything else.

To figure out what to learn next, students and teachers need to know what is already known. Assessment makes its point of entry in providing this information. It might do so in a manner as informal as a short conversation, or as structured as an adaptive online assessment. Advanced measurement modeling helps ensure that the information obtained is precise and meaningful, and advanced cognitive theory relates the assessment results to broader performance expectations.

Assessment as and for learning tells the individual student what to focus on next: the easiest lesson in the learning progression that has not yet been mastered. Here assessment results converge with the curriculum. Theoretical expectations as to developmental sequences align with practical experience in the way the difficulties of items and tasks increase with their cognitive complexity. This correspondence provides the structure by which formative feedback can individualize instruction and improve outcomes (Black & Wiliam, 1998, 2009; Hattie & Timperley, 2007; Wilson, 2004, 2009).

Assessment can be the place where learning happens, but the fact that students learn while answering test questions has often been viewed as more of a problem than an opportunity. Assessments deliberately designed to tap existing knowledge can leverage that knowledge to create new learning. Decades of research in measurement and cognition are coming to fruition in practical classroom applications (Feng, Heffernan, & Koedinger, 2009; Hannafin & Foshay, 2008; Law & Leung, 2012; Mok, 2011). As these results are brought to bear, demand will likely grow for better theoretical control, more efficient and meaningful communication, and systematic improvement methods. Leaders in education must awaken to new challenges beyond assessment, and learn new responsibilities.

Theory

“There is nothing so practical as a good theory” (Lewin, 1951, p. 169). Theory is efficient. Item difficulties are increasingly accurately predicted from previous experience or from a thorough understanding of the construct's properties (Embretson, 1984, 1998; Embretson & Daniel, 2008; Gierl & Lai, 2012; Stenner, Fisher, Stone, & Burdick, 2013; Stenner & Smith, 1982; Stenner & Stone, 2010). In this context, data gathering and analysis can become a needless waste of resources. Clocks, thermometers, voltmeters, and electrical cable are calibrated from theory, not data. The practical value of predictive theory is such that there would be no electrical industry or array of convenient consumer electronics and appliances if the resistance properties of every meter of cable had to be empirically calibrated.

As the patterns observed in assessment data repeat themselves over time and space and across millions of students and thousands of items, the difficulties of assessment items will be determined less by data analysis than by means of experimentally validated predictive theory. Table 1 lists a small sample of studies, their constructs and predictive successes. Figure 1 shows a typical scatterplot from one of those studies illustrating the correspondence between observed and expected item calibrations. Plainly test items automatically generated from theory (Bejar, Lawless, Morley, Wagner, Bennett, & Revuelta, 2003; Embretson, 1999; Gierl & Lai, 2012; Gorin & Embretson, 2012; Stenner, Swarz, Hanlon, & Emerson, 2012; Ying & Yang, 2008) stand to improve assessment efficiency by an order of magnitude.

Recent estimates put the cost of developing a single high stakes item for use in a pencil and paper assessment at US$4,000 (Stenner, et al., 2012). Items built by computers on the fly in the course of administering an assessment, in contrast, cost less than US$0.01. Maintaining test security is complex and difficult, introducing another array of costs. Compromised assessment items are useless when high stakes decisions are riding on the results. Single-use theory-based items make every assessment unique without compromising comparability, making response key cheating a thing of the past. Theory-informed, targeted curriculum materials (Law & Leung, 2012; Stenner, et al., 2012) can individualize instruction and assessment over time, and provide a basis for learning growth modeling (Figure 2). If educators dream big and take perfection as a goal, even the two sigma problem (Bloom, 1984) may not be out of reach (more on this below).

Standardized Units

Universally uniform units of measurement provide common languages used by communities of research and practice to coordinate their collective learning. Metrology, the science of calibrating instruments to standards, is increasingly credited with establishing the basis for distributed cognition across social networks, and so also for much of the power of science and commerce (Ashworth, 2004; Latour, 1987, pp. 147-157). In the same way that assessment

tells students where they are and what to do next, so, too, it should allow teachers to see where they stand relative to their peers in their instructional effectiveness. Standardized units of measurement are essential to creating the needed common languages (Fisher, 2009, 2012a, 2012b).

Communication is simplified when numbers mean the same thing no matter which particular assessment items were used and no matter which particular student in which particular grade at which particular school is measured. Though this simplification is the fundamental requirement of Rasch's (1960; Andrich, 2010; Wright, 1977, 1999; Wright & Mok, 2000) concept of specific objectivity, the lack of predictive theory and the concomitant need for empirical instrument equating have made creating standard units of measurement impractical (Fisher & Stenner, 2013, p. 8). Theory-referenced standard units of measurement make it possible to plot longitudinal trajectories over time in growth charts. Growth in learning can be managed with the end in mind, so that action can be taken as soon as there is any sign of special need (Figure 2).

Consensus standards for uniform product definitions and units of measurement are of huge economic value. Measurement standards have been shown to increase productivity, reduce transaction costs, improve efficiency in research and development, enable the creation of new markets, and enhance product quality (NIST, 2009). For instance, according to the ISO (2010), the benefits of standards within the worldwide automobile industry contribute US$25 to $55 billion to the global economy annually.

Similar new efficiencies can likely be expected to accrue in education as measurement becomes more meaningful in terms of representing real change in universally uniform terms (Fisher, 2011). Educators will develop shared expectations, vocabularies, and terminologies around their own product definitions, such as the gains typically made in any given level of mathematics or reading instruction. These common languages will make it possible for educators to see more clearly what works and what does not, while also informing parents, students, and the public about the range of outcome quality

available in local education markets. The ability to compare outcomes simply will be an important step forward in bringing the cost-quality relationship in education under better control.

Quality Improvement

The emerging culture of learning (Shepard, 2000) will increasingly blend with continuous quality improvement values and methods (Deming, 1986, 1994; Heinemann, Fisher, & Gershon, 2006; Lunenberg, 2010) and the ideal of learning organizations (Senge, 2006; Senge, Kleiner, Roberts, Ross, & Smith, 1994; Kotter, 1996). Strong theory, standardized outcome definitions, and uniform units of measurement set the stage in education for powerful implementations of the ideas that transformed the manufacturing and service industries. Deming's 14 Total Quality Management (TQM) principles and four points on quality have a wide scope of application relevant in every respect to the needs for improved outcomes in education (see boxes).

In accord with TQM and CQI methods, it will be essential to separate general and systemic common causes of quality issues managed by school leadership from local and idiosyncratic special causes of quality issues managed by students and teachers. Traditional quality control methods focus on removing undesired results from production without addressing the systemic causes bringing them about in the first place. These so-called "tail-chopping" methods get their name from the process of cutting off the low-quality end of a distribution by focusing on worker training or local, special causes of problematic results. The problem with this approach is that the systemic causes of poor quality remain unaddressed, and nothing is done to prevent the tail of the quality distribution from simply reappearing in the next round of production.

TQM and Continuous Quality Improvement (CQI) methods, in contrast, focus on shifting the entire curve to a higher overall level. This is accomplished by modifying the general, systemic causes of problems to improve the process as a whole.

In education, Bloom (1984) grasped the essence of TQM and CQI curve-shifting methods in his definition of what he called the two sigma problem. To improve group instruction to the point that its outcomes are equivalent to the outcomes of one-on-one tutoring, it is necessary to move the average performance up the scale by two standard deviations. Figure 3 shows the distributions commonly associated with classroom instruction, mastery learning, and tutoring. Bloom focused so intently on the two sigma problem that Ben Wright, speaking at Bloom's retirement celebration, referred to these curves as Bloom's "personal logo" (Bloom, 2006, p. xvi). The convergence within education of networked information technology, strong theory, standardized metrics and outcome definitions, and systematically applied TQM/CQI methods may in time fulfill Bloom's dream of achieving the best possible outcomes for all students.

Moving into the Future

As the research and technical foundations of assessments for and as learning are laid, new and broader issues of academic achievement will arise. Assessment for and as learning structures academic planning and goal setting, but both of these have been shown to have less of an effect on achievement than students' academic goal orientation (Zhu & Mok, 2012). Alignment with mastery goals and studying for the love of learning require students to fully possess their reasons and purpose in pursuing knowledge.

And what goes for students goes double for teachers and educational leadership. Fulfilling every student's potential will demand that educators take ownership of the full scope of the educational process, from beginning to end. Figure 4 illustrates the operational change and improvement cycle proven in its effectiveness time and again across a wide range of industries. Note the correspondence between predictive theory and guiding ideas, between standards and infrastructural innovations, and between TQM/CQI and applied theory, methods, and tools. Organizations capable of creating a culture of continuous learning around Deming's 14 principles will endlessly move through the change and improvement cycle to create and share new value.

Figure 5 shows the conditions that must be met for sustainable change to be realized, along with the negative results that accrue when any given element is omitted (Knoster, Villa, & Thousand, 2000). Vision, skills, incentives, resources, and a plan must all be in place to avoid confusion, anxiety, resistance, frustration, and treadmills. Creating a culture of learning with an advanced information infrastructure, predictive theory, the common languages of standard metrics, and systematic quality improvement methods is a huge challenge, but as Georg Rasch (1980, p. xx) recognized in a related context, “once the problem has been formulated it does seem possible to meet it."

Table 1 : Publications Featuring Item Calibrations Predicted from Theory

Authors	Scale	Variance Explained
Irvine, Dunn, Anderson, 1990	British Army Recruitment Battery	R > .70
Embretson, 1998	Abstract Reasoning Test	R > .70
Stenner & Smith, 1982	Knox Cube Test	R > .90
Fischer, 1973	Elementary calculus test	R > .85
Stenner, et al, 1983	Peabody Vocab Test	R > .80
Stenner, et al, 1997	Reading tests	R > .90
Bejar, et al, 2003	Math tests	R > .85
Fisher, 2008	Physical function surveys	R > .90

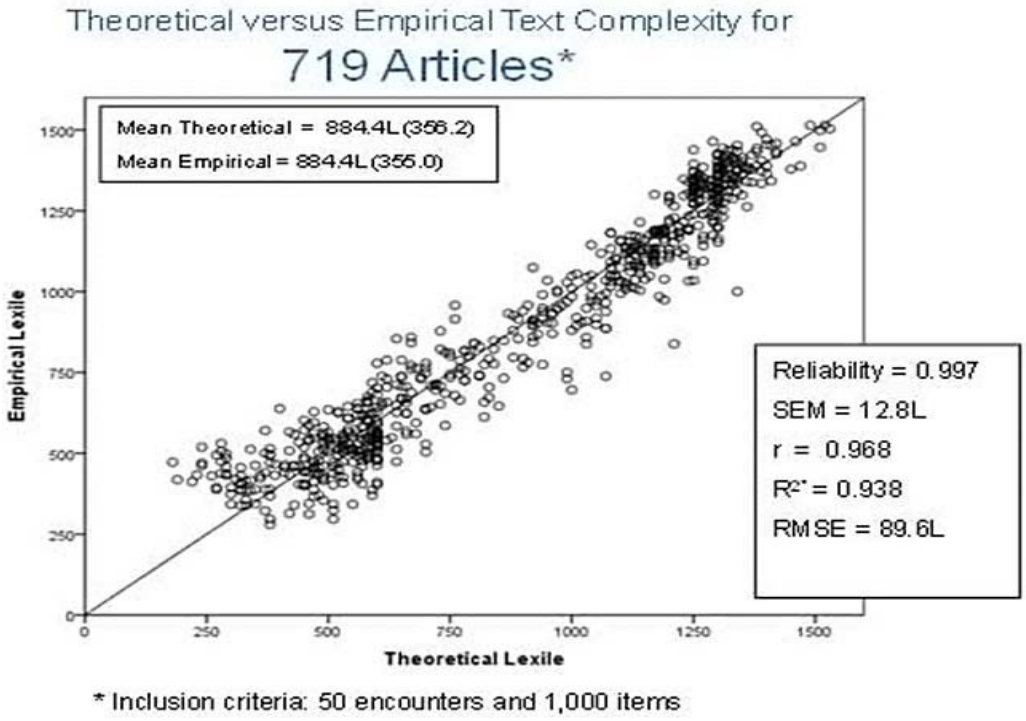


Figure 1. Empirical vs. theoretical item calibrations
(Stenner & Burdick, 1997)

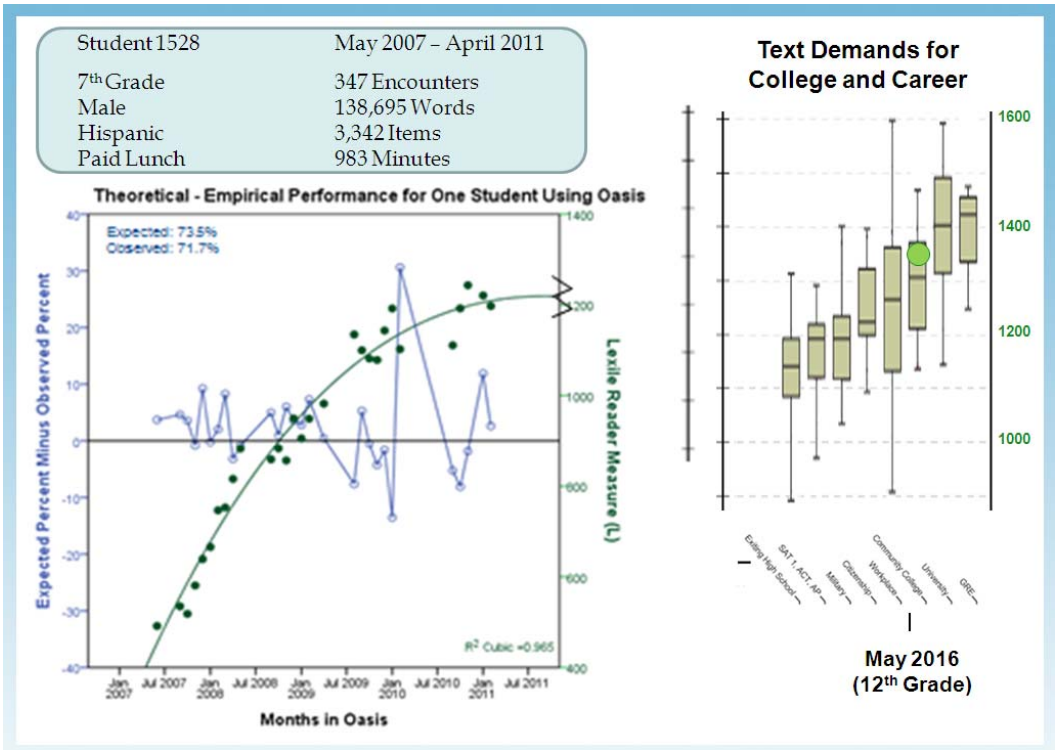


Figure 2. Individualized growth in reading relative to desired outcomes (Stenner, Swartz, Hanlon, & Emerson, 2012)

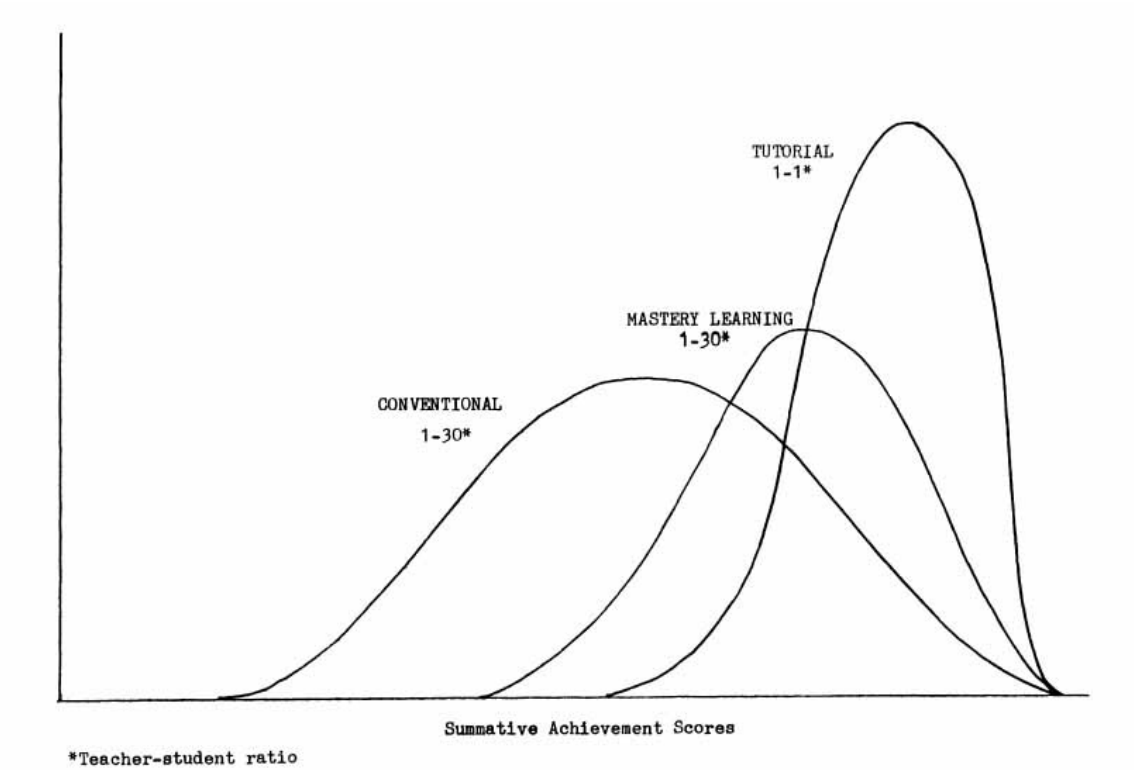


Figure 3. The two sigma problem (Bloom, 1984, p. 5)

Deming's 14 Total Quality Management Principles Adapted to Education
(Lunenberg, 2010; Deming, 1986, 1994)

1. Constancy of purpose
2. Adopt the new philosophy.
3. Augment end-point accountability with formative assessment.
4. Focus less on lowest-cost bids and more on long term relationships of trust.
5. Improve constantly and forever.
6. Institute continuous on-the-job training.
7. Lead toward learning for all.
8. Drive out fear.
9. Break down barriers between staff areas.
10. Eliminate slogans, exhortations, and productivity targets, like proficiency levels.
11. Eliminate numerical quotas and goals.
12. Remove barriers that rob students and teachers of their natural pride in their work.
13. Institute vigorous education and training for all.
14. Involve everyone in the transformation.

Deming’s 4 Points Brought to Bear in Education
(Adapted from Lunenberg, 2010; Deming, 1986, 1994)

1. Appreciate the system: Understand all of the processes and roles relevant to how educational outcomes are produced, including teaching methods, relationships with textbook and material suppliers and producers, the students and their families, the community, teachers, support staff, and the material consequences of the building housing the school, the cafeteria operations, and the physical plant.

2. Knowledge of variation: Understand the range and common vs. special causes of variation in quality, and use measures that can be interpreted and applied not just in accountability applications, but in assessment as and for learning.

3. Theory of knowledge: Understand what knowledge is and the limits of what can be known.

4. Psychology: Understand human nature so as to be able to create satisfying, joyful work for all involved.

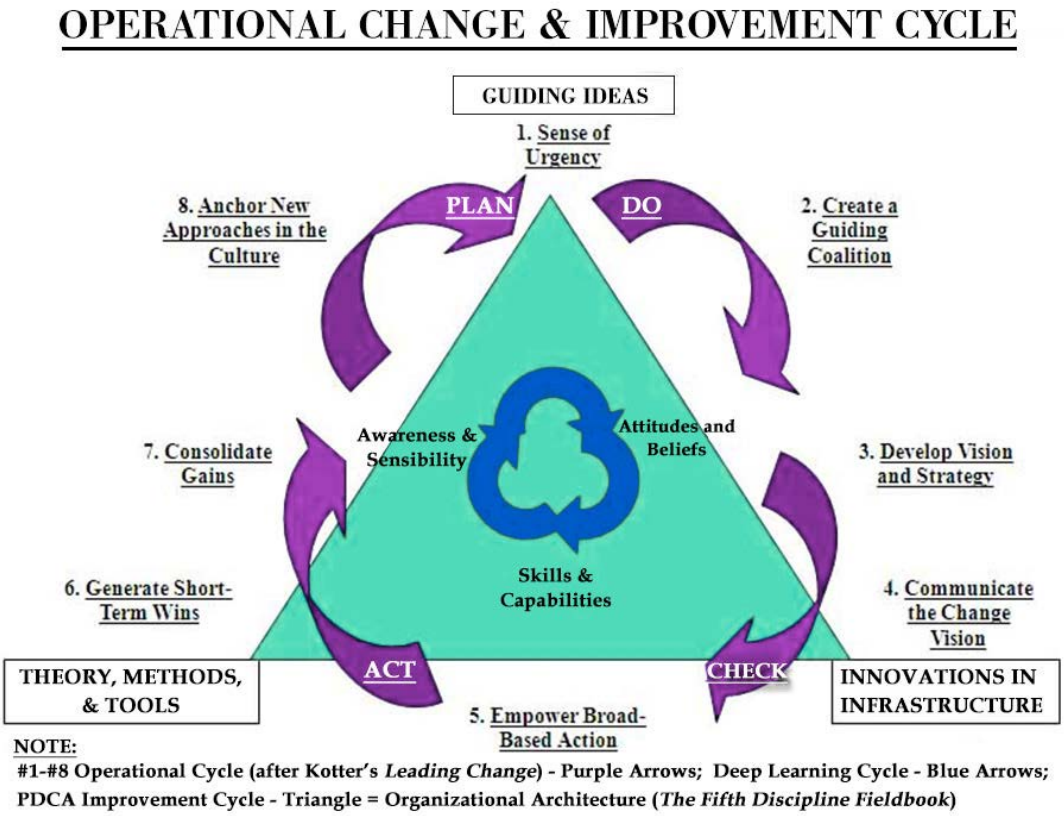


Figure 4. Operational change and improvement cycle (Kotter, 1996; Senge, 2006; Senge, Kleiner, Roberts, Ross, & Smith, 1994)

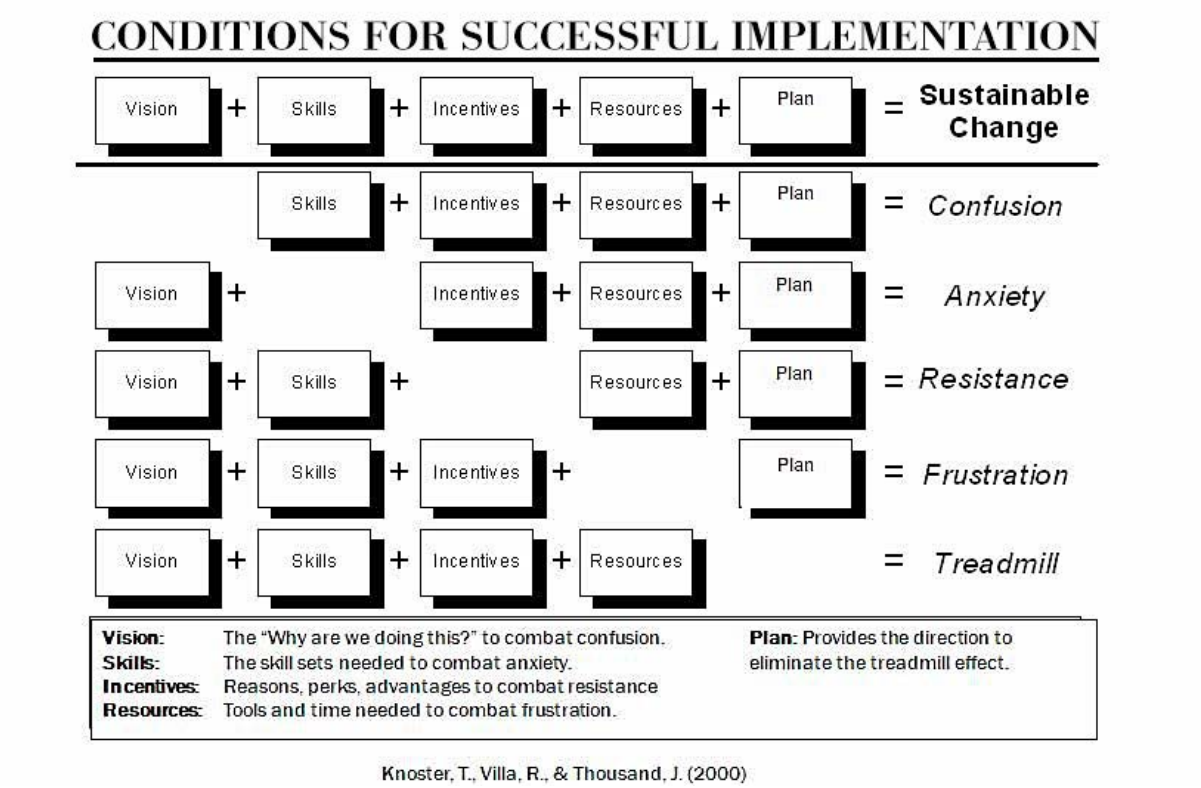


Figure 5. Conditions for successful implementation (Knoster, Villa, & Thousand, 2000)

References

Andrich, D. (2010). Sufficiency and conditional estimation of person parameters in the polytomous Rasch model. *Psychometrika, 75*(2), 292-308.

Ashworth, W. J. (2004). Metrology and the state: Science, revenue, and commerce. *Science, 306*(5700), 1314-1317.

Bejar, I., Lawless, R. R., Morley, M. E., Wagner, M. E., Bennett, R. E., & Revuelta, J. (2003). A feasibility study of on-the-fly item generation in adaptive testing. *The Journal of Technology, Learning, and Assessment, 2*(3), 1-29.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*, 5-31.

Bloom, B. S. (1984). The two sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4-16.

Bloom, S. (2006). Preface. In T. R. Guskey (Ed.), Benjamin S. Bloom: *Portraits of an educator* (pp. xi-xx). Lanham, Maryland: Rowman & Littlefield Education.

Cheng, Y. C., & Mok, M. M. C. (2007). School-based management and paradigm shift in education: An empirical study. *International Journal of Educational Management, 21*(6), 517-542.

Deming, W. E. (1986). *Out of the crisis*. Cambridge, MA: MIT Press.

Deming, W. E. (1994). *The new economics for industry, government, education* (2d Ed.). Cambridge, Massachusetts: MIT Center for Advanced Educational Services.

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize classroom learning*. Thousand Oaks, California: Corwin.

Embretson, S. E. (1984). A general latent trait model for response processes. *Psychometrika, 49*(2), 175-186.

Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods*, 3(3), 380-396.

Embretson, S. E. (1999). Generating items during testing: Psychometric issues and models. *Psychometrika, 64*(4), 407-433.

Embretson, S. E., & Daniel, R. C. (2008). Understanding and quantifying cognitive complexity level in mathematical problem solving items. P*sychology Science Quarterly, 50*, 328-344.

Feng, M., Heffernan, N., & Koedinger, K. (2009). Addressing the assessment challenge with an online system that tutors as it assesses. *User Modeling and User-Adapted Interaction, 19*(3), 243-266.

Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 37, 359-374.

Fisher, W. P., Jr. (2008). *A predictive theory for the calibration of physical functioning patient survey items*. Presented at the Second Conference on Patient Reported Outcomes Measurement Information Systems, Bethesda, Maryland: NIH and NIAMS.

Fisher, W. P., Jr. (2009). Invariance and traceability for measures of human, social, and natural capital: Theory and application. *Measurement, 42*(9), 1278-1287.

Fisher, W. P., Jr. (2011). Bringing human, social, and natural capital to life: Practical consequences and opportunities. *Journal of Applied Measurement, 12*(1), 49-66.

Fisher, W. P., Jr. (2012a). Measure and manage: Intangible assets metric standards for sustainability. In J. Marques, S. Dhiman & S. Holt (Eds.), *Business administration education: Changes in management and leadership strategies* (pp. 43-63). New York: Palgrave Macmillan.

Fisher, W. P., Jr. (2012b). What the world needs now: A bold plan for new standards. *Standards Engineering, 64*(3), 1 & 3-5.

Fisher, W. P., Jr., & Stenner, A. J. (2013). On the potential for improved measurement in the human and social sciences. In Q. Zhang (Ed.), *Pacific Rim Objectiive Measurement Symposium 2012 Conference Proceeding* (pp. 1-11). Berlin, Germany: Springer-Verlag.

Gierl, M. J., & Lai, H. (2012). The role of item models in automatic item generation. *International Journal of Testing, 12*(3), 273-298.

Gipps, C. (1994). *Beyond testing*. Abingdon: Routledge.

Gorin, J. S., & Embretson, S. E. (2012). Using cognitive psychology to generate items and predict item characteristics. In M. J. Gierl & T. M. Haladyna (Eds.), *Automatic item generation: Theory and practice* (pp. 136-156). New York: Routledge.

Hannafin, R. D., & Foshay, W. R. (2008). Computer-based instruction's (CBI) rediscovered role in K-12: An evaluation case study of one high school's use of CBI to improve pass rates on high-stakes tests. *Educational Technology Research & Development, 56*(2), 147-160.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Heinemann, A. W., Fisher, W. P., Jr., & Gershon, R. (2006). Improving health care quality with outcomes management. *Journal of Prosthetics and Orthotics, 18*(1), 46-50.

International Organization for Standardization (ISO). (2010). *Assessing economic benefits of consensus-based standards*. Retrieved 6 July 2013, from http://www.standardsinfo.net/info/benefits/Economic%20benefits%20of%20standardization-Overview.pdf.

Irvine, S. H., Dunn, P. L., & Anderson, J. D. (1990). Towards a theory of algorithm-determined cognitive test construction. *British Journal of Psychology, 81*, 173-195.

Knoster, T. P., Villa, R. A., & Thousand, J. S. (2000). A framework for thinking about systems change. In R. A. Villa & J. S. Thousand (Eds.), Restructuring for caring and effective education: *Piecing the puzzle together, 2nd Ed.* (pp. 93-128). Baltimore: Paul H. Brookes.

Kotter, J. (1996). *Leading change*. Cambridge, Massachusetts: Harvard Business School Press.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. New York: Cambridge University Press.

Law, K. K. Y., & Leung, F. F. Y. (2012). Design of web-based support materials to address learning difficulties – vocabulary building skills. *Assessment and Learning, 1*, 165-182.

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers* (D. Cartwright, Ed.). New York: Harper & Row.

Lunenberg, F. C. (2010). Total Quality Management applied to schools. *Schooling*, *1*(1), 1-6.

Ma, M. K. K. (2012). Assessment for learning: Collaborative research and professional development projects in school context. *Assessment and Learning, 1*, 157-164.

Mok, M. M. C. (2011). The assessment for, of and as learning in mathematics: The application of SLOA. In B. Kaur & K. Y. Wong (Eds.), *Assessment in the mathematics classroom* (pp. 187-215). Singapore: World Scientific.

NIST. (2009). *Outputs and outcomes of NIST laboratory research*. Retrieved 1 March 2011, from NIST: http://www.nist.gov/director/planning/studies.cfm.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests* (Reprint, with Foreword and Afterword by B. D. Wright, Chicago: University of Chicago Press, 1980). Copenhagen, Denmark: Danmarks Paedogogiske Institut.

Senge, P. (2006). *The fifth discipline: The art & practice of the learning organization* (2nd Ed.). New York: Random House.

Senge, P., Kleiner, A., Roberts, C., Ross, R. B., & Smith, B. J. (1994). *The fifth discipline fieldbook: Strategies and tools for building a learning organization*. New York: Crown Business.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Stenner, A. J., & Burdick, D. S. (1997). The objective measurement of reading comprehension. [Response to technical questions raised by the California Department of Education Technical Study Group], Durham, North Carolina: MetaMetrics, Inc.

Stenner, A. J., Fisher, W. P., Jr., Stone, M., & Burdick, D. (2013). Causal Rasch models. *Frontiers in Psychology*, in press.

Stenner, A. J., & Smith III, M. (1982). Testing construct theories. *Perceptual and Motor Skills, 55*, 415-426.

Stenner, A. J., Smith, M., III, & Burdick, D. S. (1983). Toward a theory of construct definition. *Journal of Educational Measurement*, *20*(4), 305-316.

Stenner, A. J., & Stone, M. (2010). Generally objective measurement of human temperature and reading ability: Some corollaries. *Journal of Applied Measurement, 11*(3), 244-252.

Stenner, A. J., Swartz, C., Hanlon, S., & Emerson, C. (2012). *Personalized learning platforms*. Presented at the Pearson Global Research Conference, Fremantle, Western Australia.

Wilson, M. (Ed.). (2004). *Towards coherence between classroom assessment and accountability*. Chicago, Illinois: University of Chicago Press.

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching, 46*, 716–730.

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, *14*(2), 97-116.

Wright, B. D. (1999). Fundamental measurement for psychology. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every educator and psychologist should know* (pp. 65-104). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Wright, B. D., & Mok, M. M. C. (2000). Understanding Rasch measurement: Rasch models overview. *Journal of Applied Measurement, 1*(1), 83-106.

Ying, M.-H., & Yang, H.-L. (2008). Computer-aided generation of item banks based on ontology and Bloom's taxonomy. In F. Li, J. Zhao, T. Shih, R. Lau, Q. Li & D. McLeod (Eds.), *Advances in Web Based Learning - ICWL 2008* (Lecture Notes in Computer Science) (pp. 157-166). Berlin: Springer.

Zhu, J., & Mok, M. M. C. (2012). Effect of academic goal orientation, goal setting, and planning on academic achievement of secondary students in Hong Kong. *Assessment and Learning, 1*, 11-31.

Author's e-mail:　wfisher@berkeley.edu

The Use of CADATS to Establish an Item Bank for Formative Assessment in Primary Schools

Qingping He, Peter Tymms and See Ming Tsui
Centre for Evaluation and Monitoring, Durham University

This short article proposes the use of a computer assisted assessment system, the Computer Assisted Design, Analysis and Testing System (CADATS), developed at Durham University to establish an item bank system for conducting effective formative assessment by teachers in primary schools.

Introduction

Tests are frequently used to assess students' learning progress, and feedback has become an important and integral part of almost all forms of educational assessment, particularly formative assessment. It has been recognised that feedback to students can act as an effective means for improving students' performance. Results from research have generally indicated that the use of effective feedback can improve students' learning outcomes substantially. Feedback generating process has to include the following fundamental elements for it to be effective:

- A recognised and measurable standard must be set in the assessment;
- There exists a means to identify students' performance in relation to the standard set in the assessment;
- There exists a way that applies the information obtained in order to narrow the gap between the student's performance and the standard set by the assessment.

The generation of feedback can therefore represent a demanding and complex task for teachers, since it is the teachers themselves that are responsible for developing tests; setting target standards; undertaking results

analysis and data interpretation; identifying gaps between students' performance and the target standard; and developing procedures to improve teaching and students' learning in areas where students are below the target standard. The complexity of the feedback generating process has hindered the use of effective feedback procedures in formative assessments as part of normal classroom practice. In view of the limited time available to teachers and the difficulties involved in producing feedback, a dedicated computer assisted system that can perform all or part of the work involved in the process of generating feedback would help teachers substantially in terms of saving their time and encouraging them to adopt effective feedback methods in teaching. Although the last element listed above is teacher-specific as different teachers may use different methods in dealing with the specific problems their students encounter, a computer assisted system can be developed to help teachers with the first two elements.

The computer assisted assessment (CAA) system CADATS (Computer Assisted Design, Analysis and Testing System) developed at Durham University which has been tried out in some schools in Hong Kong and in China would seem to fulfil the above requirements.

The CADATS System

CADATS can perform a range of functions to generate information that can be used by teachers to provide effective feedback to students. Main functions provided by CADATS include:

- Creating items of various types and construct item banks.
- Designing tests (Classical Test Theory – CTT based or Item Response Theory – IRT based including Computer Adaptive Tests – CATs) effectively by selecting items from an item bank.
- Conducting tests (CTT and IRT-based tests, including CATs) on standalone or networked computers.
- Analyzing test results. The system is able to undertake detailed diagnostic analysis on the performance of students and test items at individual, class and whole year group levels in order to identify curriculum areas where students are under-performing.

- Analyzing the performance of test items using CTT and the Rasch IRT model, and equating tests.

Building an Item Bank System Using CADATS in Primary Schools to Facilitate Effective Formative Assessment

Given the functions provided by CADATS, an item bank system can be built using CADATS to facilitate effective formative assessment by teachers in schools. This section explains how to build such a system.

Research has indicated that feedback in improving students' learning outcomes is most effective when it is task-oriented. Questions developed by experienced teaching professionals who understand the curricula very well have specific assessment objectives in measuring skills gained by students in specific curriculum areas and are particularly useful for goal-setting.

To make the system more effective and efficient, it would therefore be best if the system is built and used by a consortium of, say, 20 primary schools for all year groups in the schools. These schools will need to work together. The following activities will need to be taken:

Item and Test Development – Item Banking

Participating schools will provide items to cover topics taught in their schools for the selected subject (such as mathematics, English or any other subject). Items should cover the full primary curriculum and developed by experienced teaching professionals. Items will need to be classified based on year groups and curriculum topic areas. These items will then be digitised for inclusion in the item bank built using CADATS.

Tests for different year groups will be designed using the items in the item bank for each academic year. In order to equate the tests for different year groups, common items will be embedded in tests for adjacent year groups.

The item bank and the associated tests can be shared by all schools in the consortium so that the schools can use the item bank in conjunction with other

functions provided by CADATS to enhance their teachers' teaching and students' learning experience.

Test Administration and Diagnostic Analysis

Schools in the consortium will need to administer the tests included in the item bank to their students using the CADATS test delivery system. Teachers can use CADATS to conduct diagnostic analysis of the results from the tests in order to help them identify curriculum areas where individual students are having difficulties and to provide individualised help to the students.

Item and Test Analysis and Test Equating

Responses from students from all schools in the consortium will be analysed together using the analysis functions of CADATS in order to derive item and test statistics (both CTT and IRT statistics, including for example item percentage score or facility and item difficulty) for the population. To place items for all year groups on the same difficulty scale, the test equating function of CADATS can be used to equate the different tests using the common item equating procedure. The statistics of the items can be loaded into the item bank.

Development of a Computer Adaptive Test

It would be possible that after about three years, a reasonably large calibrated item bank can be created. Once a calibrated item bank has been established, a computer adaptive test for the subject covering all year groups can be created. Such a test would be particularly useful for monitoring students' study progress. Teachers can also use the items in the bank to design their own tests if they so wish.

Training for Teachers

In order to help teacher to use CADATS to conduct formative assessment effectively and make it part of their normal classroom practice, training should be provided to teachers from schools in the consortium. Focus of training should be on how to use CADATS to undertake detailed diagnostic analysis for identifying problem areas in their teaching.

Discussion

Benefits of using the system described above would include:

- The encouragement of the willingness in using formative assessment by teachers to enhance the teaching and learning process as part of their normal classroom practice.
- The establishment of a calibrated item bank (in which items have known statistics) to cover all year groups in primary schools for specific subjects. A number of CTT and IRT based tests can also be included in the item bank. Specifically, the items in the bank can have a common difficulty scale and a computer adaptive test can be created for adaptive testing across the year groups. The item bank can also be used in conjunction with CADATS to enhance teaching and learning with the aim to improve students' learning outcomes. Teachers can also create tests themselves by using the items in the bank. Because the item statistics (such as item percentage scores) for the population can be established through the use of the system, teachers can use the results from the test designed using the items in the bank to compare the performance of their students with the average (norm) performance of the population to identify areas where their students under-perform and therefore enhance teaching in those areas.

Further Reading

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A. & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research,* 61, 213-238.

Black, P. and Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education,* 5, 7-71.

Black, P. and Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139-148. Also available online at: http://www.pdkintl.org/kappan/kbla9810.htm (Accessed on 26 September, 2007).

Bennett, R. (2011). Formative assessment: a critical review. *Assessment in Education*, 18: 1, 5-25.

Gibbs, G. & Simpson, C. (2003). Does your assessment support your students' learning? *Journal of Learning and Teaching in Higher Education,* 1.

Hattie, J. & Jaeger, R. (1998). Assessment and classroom learning: A deductive approach. *Assessment in Education,* 5, 111-122.

He, Q. & Tymms, P. B. (2005). A computer-assisted test design and diagnosis system for use by classroom teachers. *Journal of Computer Assisted Learning,* 21, 419-429.

McDonnell, M. (2007). Formative feedback and learning in the English classroom. http://www.golding.ca/maggie/documents/McDonnell_LitReview.doc (accessed September 24, 2007)

Nichol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education,* 31, 199-218.

Smith, E. & Gorard, S. (2005). 'They don't give us our marks': The role of formative feedback in student progress. *Assessment in Education,* 12, 21-38.

Taras, M. (2002). Using assessment for learning and learning from assessment. *Assessment & Evaluation in Higher Education,* 27, 501-510.

Tsuei, M. (2007). A web-based curriculum-based measurement system for class-wide ongoing assessment. *Journal of Computer Assisted Learning* (in press).

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education,* 45, 477-501.

Authors' email: Qingping He Qingping.He@Ofqual.Gov.Uk
Peter Tymms Peter.Tymms@cem.dur.ac.uk.
See Ming Tsui inei@netvigator.com

Implementing Assessment *for* Learning: An Application of the Rasch Model for the Construction of a Mathematics Assessment to Inform Learning

Jingjing Yao[1, 2] and Magdalena Mo Ching Mok[1]

[1]Assessment Research Centre, The Hong Kong Institute of Education
[2]Department of Psychology, Zhejiang Normal University

Acknowledgement

This study was supported by a grant from the General Research Grant sponsored by the Research Grants Council of Hong Kong (Project Number 844011).

1. Introduction

Mathematical literacy is a key competence for individuals and organisations in society today (OECD, 2010; National Mathematics Advisory Panel, 2008). In order to prepare students' capacity for understanding and applying mathematics, there has been an increase of attention in recent years on the stronger integration among mathematics curricula, instruction, and assessment of mathematics, especially in the foundation stage of primary school (Abakhani, 2011; Arslan & Ozinar, 2010; Bulut, 2007).

The subject of mathematics has had a long history of using quizzes, tests and examinations for the assessment *of* learning. That is, the identification of standards reached by students, particularly at the end of key learning stages, used to be the sole purpose of assessment. Needless to say, assessment *of* learning is important for accountability purposes; and in this regard, schools and teachers are accountable to taxpayers, parents, and school sponsoring bodies. Assessment *of* learning is also necessary to maintain academic standards of the education system and for articulation between education systems. Nevertheless, assessment *of* learning alone is inadequate to prepare our students in facing

challenges unique to the twenty-first century workplace. Furthermore, we now know a lot more about the power of assessment in transforming teaching and learning. Notably, large scale reviews undertaken by Black and Wiliam (1998), Hattie and Timperley (2007), Kluger and Denisi (1996), Mory (2004), Narciss and Huth (2004), and Shute (2008), and independent classroom-based research conducted by Berry (2008), Carless (2007), Lee (2012), Mok (2010), Salvage (2011), and others have repeatedly shown that through providing diagnostic feedback, assessment can inform and support further learning of students. These studies provided empirical evidence that quality feedback helps students to diagnose their learning progress, how well they have learned, identifies the gaps and the nature of misconceptions. As a result, students are supported with information to improve their learning; a development that would not occur without such feedback information. This means that whereas validity, reliability, and discrimination are essential features of assessment of learning, diagnostic feedback is a critical component of learning assessment.

Despite consistent research on the importance of feedback to learning, the implementation of assessment *for* learning is often hindered by the lack of tools required for the generation of diagnostic feedback. If we step back and reflect on the question: When teachers ask their students to take a test, what feedback information do they want from this test? The answer is two-fold. First, teachers should want to know more about their students; namely, how well each individual student is doing, the level of target knowledge mastery and skills of the whole class, and how to help each and every one of them to further their learning from where they are, based on the diagnostic information derived from students' misconceptions and non-mastery. Second, teachers may also want to know about the quality of the test; namely, the overall test difficulty and the difficulty of individual items. Some teachers may also want to address the questions of reliability, fairness, and validity of the test and its individual items.

A certain amount of the feedback information above can be generated by experienced teachers themselves through inspecting the distribution of raw scores for the whole class, or by inspecting individual item- and student-responses. But other information (e.g., the issue of "where to go from

here?" for individual students) can hardly be obtained without the help of analytic tools. This is the time when such analytic tools as the Rasch model (Bond & Fox, 2007) could be helpful.

Moreover, the literature of Embretson (1996) warns against traditional methods of using raw scores in the analysis of assessment data. Research has shown that raw scores cannot be assumed to be interval-level data (Stevens, 1946), and treating them as such will lead to the misinterpretation of test quality and of student achievement. For instance, to improve from a score of 98 to 99 in a test, with a maximum score of 100, is much harder than to improve from a score of 71 to 72 in the same test. This means the distances between the two pairs of raw scores might be both one unit mathematically, but they are not of the same distance. It is therefore erroneous to compare students directly based on raw scores (Embretson, 1996).

In view of the need for diagnostic tools in support of assessment for learning, this study aims to illustrate, through an example of Rasch analysis on students' responses to a 35-item mathematics assessment designed for Primary 5 students in Hong Kong, how the Rasch model (Bond & Fox, 2007) could be used to optimize the effectiveness of assessment for learning in school-based assessments. In this example, the easily accessible Winsteps software (Linacre, 2011) was used and the main steps of Rasch analysis for extracting diagnostic information in support of teaching and learning were introduced didactically.

2. Method

2.1 Participants

This is part of a larger longitudinal study on assessment feedback, self-directed learning, and mathematics achievement of primary students. Participants for the current study comprised a sample of 1368 Primary 5 students from 16 Hong Kong schools. There were 648 males (47.4%), 716 females (52.3%), and 4 students (0.3%) did not report their gender (Table 1). This study observed all ethical compliances set by the university where the

authors worked, and informed consents from parents and schools were obtained before the commencement of the study.

Table 1. Sample Distribution by Gender

	N	%
Male	648	47.4%
Female	716	52.3%
Missing	4	0.3%
Total	1368	100%

2.2 Instrument

A 35-item mathematics test was developed after careful analysis of the Hong Kong mathematics curriculum, and in consultation with teachers on the suitability of the test for Primary 5 students by the end of Semester One. All the items in the test were multiple choice questions with four options and only one of the options was the correct answer. The test consisted of 25 items in the Number domain, three items in the Shape and Space domain, and seven items in Measures domain. Within the Number domain, nine items involved the understanding of the basic concepts of whole numbers and fractions, 12 items involved performing addition, subtraction, multiplication operations, as well as mixed operations on whole numbers and fractions, and another four items involved solving application problems. In the Shape and Space domain, the three items were on direction and location. In the Measurement domain, the seven items were on the calculation of perimeters and areas (Table 2).

Table 2. Item Domain Contents and Numbers

Domain	**Contents**	**Items**
Number	basic concepts of whole numbers and fractions	9
	addition, subtraction, multiplication and their mixed operations	12
	application problem solving	4
Shape & Space	direction and location	3
Measures	calculation of perimeters and areas	7

2.3 Procedure and Analysis

First, we analysed the local curriculum carefully, selected the representative contents and typical items from textbooks, exercises and other related materials. On the basis of this analysis, we designed items for the assessment. The purpose of the study and drafts of the assessment were presented to participating primary mathematics teachers in order to consult with them on suitability and practicability of the assessment for the targeted cohort of students. After several rounds of consultation and revisions, the final version of the assessment was administered at the end of Semester One to 1368 Primary 5 students from 16 Hong Kong schools under the supervision of mathematics teachers during normal school time.

Rasch analysis was conducted using the Winsteps software (version 3.72.3) (Linacre, 2011) to validate the mathematics assessment. As responses to the items in the assessment were scored, either right or wrong, a dichotomous Rasch model (Rasch, 1960) represented in equation (1) was used to estimate difficulties of the item, or the item measures, and mathematics ability of the student, or the person measure, on a common interval scale of mathematics ability.

$$P_{ni1} = \frac{\exp(\theta_n - \delta_i)}{1 + \exp(\theta_n - \delta_i)} \qquad (1)$$

where P_{ni1} is the probability of person n making a correct response to item i. Correspondingly, P_{ni0} (= 1 − P_{ni1}) is the probability of this person n making a wrong response to the same item i. In the Rasch model expressed in equation (1), the probability of a correct response is a logistic function of the difference between the person ability θ_n and item difficulty δ_i. Thus, we can place item difficulty and student ability on the same measurement scale for interpretation. It will be discussed in later sections of this manuscript that other diagnostic information about the students and the assessment, for example the extent to which items making up the assessment fall into a single dimension, can be generated using the Rasch model.

3. Results

The Rasch analysis conducted below was divided into two parts: (a) the validation of the mathematics assessment, and (b) the data analysis for diagnostic information. In order to validate the assessment, the unidimensionality of the mathematics assessment according to the Rasch model was first tested to ascertain the extent to which the assessment was underpinned by a single Rasch measurement. Next, other indices of validity were generated, including person and item reliabilities, item difficulty, item fit, and gender differential item functioning (DIF) (Wang, 2008). This validation process aimed to guarantee the quality of the mathematics test as an appropriate and valid instrument for assessing the students. It also informed the teachers of the characteristics of the assessment, so that the teachers could decide whether or not the items could be included in an item bank for future assessment purposes.

The diagnosis aspect of the data analysis included the generation of estimated mathematics ability of students for the whole group as well as for individual students, the individualised diagnostic map (called the Person-Kid-Map, abbreviated as PKMAP in Winsteps), which provided information on the Zone of Proximal Development (Vygotsky, 1978) of each student, and the person Keyforms for each student, from which observed and highly unexpected responses could be identified and diagnosed. Collectively, this diagnostic information gathered is made possible through the powerful functions of Winsteps software (Linacre, 2011). The results reported below illustrate how teachers could make effective use of diagnostic information for formative purposes of assessment.

In terms of validation of the mathematics assessment, the results showed that: (1) unidimensionality of the assessment was supported by the data; (2) Rasch person and item reliabilities were acceptable; (3) with one exception, all items had item-fit between 0.5 and 1.5; (4) item difficulties ranged from -2.30 to 2.83; (5) there was good alignment between item difficulty and student ability; and (6) there was no gender DIF detected among the items. In terms of diagnostic information, the analysis showed that (7) student ability ranged between -2.58 and 4.19; (8) person diagnostic PKMAPs provided information on: the Zone of Proximal Development, the items being mastered comfortably, and future learning goals for each of the students; and (9) the person Keyforms provided information on: the extent to which responses to each item for each

student were within or out of expectation. Details of the information are reported in the sections that follow.

3.1 Unidimensionality of the Assessment

Test of unidimensionality of the assessment was conducted through a Principal Components Analysis of Rasch residuals subroutine in the Winsteps software (version 3.72.3) (Linacre, 2011). The Principal Components Analysis of Rasch residuals is used to detect if there is more than one factor that can explain the response structure (i.e., unidimensionality) by comparing differences between the observed and the expected responses (Raîche, 2005; Linacre, 2011). Simulation studies by Raîche (2005) found eigenvalues of the first contrast in the Principal Components Analysis of test from 20 to 60 items to be in the range of 1.4 to 2.1. The results were subsequently replicated by Linacre and Tennant (2009). The literature of Linacre (2011) recommends researchers to use eigenvalue of the first contrast being less than 2.0 as an acceptable criterion for establishing unidimensionality. In this study, the Principal Component eigenvalues in the first contrast was 1.8 (below 2.0), and 31.2% of raw variance were explained by the Rasch measures. The result indicated that the mathematics assessment was likely to be underpinned by a single dimension, which was consistent with the assessment design intent.

3.2 Reliability of Item and Person Measures

Internal consistency of assessment items in terms of Cronbach's alpha was 0.82 for the Primary 5 mathematics assessment, which was an acceptable reliability in accordance with classical test theory. The Rasch analysis also found that the assessment had a Rasch item reliability of 1.00, an item separation index of 17.17, a Rasch person reliability of 0.80, and a person separation index of 1.98. These results mean that the assessment had excellent item reliability, and the items could be separated into nearly 17 groups according to responses by students. On the other hand, the person reliability was just acceptable, and the students could be separated into almost two groups by the items in the assessment. If we take the different number of items and students into consideration, the differences in reliabilities and separation indices between items and persons could be interpreted easily. It is easy to separate 35 items by 1368 students, but it is comparatively more difficult to separate 1368 students by only 35 items.

3.3 Item Fit, Item Difficulty, and Alignment between Item and Person

Item statistics are presented in Table 3, including estimates of item difficulties, their standard errors, item goodness of fit (both Infit and Outfit), and point-measure correlation for each item. These statistics support the validity and reliability of the assessment. More details are given in the sections below.

Table 3. Item Difficulty, Standard Error, Fit, and Point-Measure Correlation

			Infit		Outfit		PTME Corr.
Item	Difficulty	*SE*	MNSQ	ZSTD	MNSQ	ZSTD	
Q16	2.83	0.08	1.16	2.85	2.01	8.28	0.07
Q12	2.23	0.07	1.01	0.33	1.19	2.73	0.32
Q34	1.48	0.06	1.11	4.00	1.23	4.99	0.27
Q22	1.37	0.06	1.03	1.11	1.07	1.73	0.37
Q31	1.17	0.06	1.16	6.21	1.29	7.37	0.24
Q15	1.14	0.06	1.02	0.95	1.10	2.74	0.37
Q25	1.12	0.06	0.96	-1.58	0.98	-0.59	0.44
Q19	0.98	0.06	0.89	-5.12	0.87	-4.02	0.51
Q6	0.96	0.06	1.04	1.87	1.05	1.63	0.37
Q27	0.85	0.06	1.01	0.28	1.02	0.76	0.40
Q13	0.73	0.06	1.04	1.75	1.07	2.36	0.37
Q35	0.73	0.06	1.10	4.59	1.13	4.17	0.31
Q18	0.66	0.06	0.94	-2.94	0.93	-2.31	0.47
Q33	0.50	0.06	1.02	0.87	1.03	0.93	0.39
Q7	0.43	0.06	1.18	7.69	1.24	7.02	0.24
Q20	0.38	0.06	0.99	-0.46	1.00	-0.05	0.42
Q29	0.36	0.06	0.94	-2.68	0.93	-2.21	0.46
Q26	0.26	0.06	0.92	-3.73	0.90	-2.99	0.48
Q28	0.03	0.06	1.19	7.26	1.36	8.55	0.20
Q2	-0.13	0.06	1.01	0.41	1.03	0.66	0.38
Q21	-0.17	0.06	0.89	-4.32	0.83	-4.11	0.50
Q32	-0.21	0.06	0.90	-3.78	0.83	-4.01	0.49
Q24	-0.47	0.07	0.98	-0.81	0.89	-2.18	0.41
Q23	-0.83	0.07	0.93	-2.01	0.91	-1.43	0.42
Q10	-0.91	0.07	0.82	-4.96	0.66	-5.79	0.53
Q11	-0.93	0.07	0.95	-1.20	0.99	-0.11	0.38
Q5	-1.11	0.07	0.95	-1.06	0.95	-0.64	0.37
Q30	-1.21	0.08	0.92	-1.80	0.80	-2.68	0.41
Q8	-1.30	0.08	0.94	-1.20	0.92	-0.95	0.36
Q17	-1.36	0.08	0.88	-2.47	0.76	-2.94	0.43

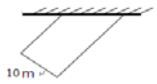
Q1	-1.40	0.08	0.99	-0.25	1.01	0.16	0.32
Q3	-1.62	0.09	1.04	0.65	1.22	2.05	0.22
Q9	-2.07	0.10	0.94	-0.82	1.08	0.68	0.29
Q4	-2.18	0.10	0.93	-0.92	0.80	-1.49	0.31
Q14	-2.30	0.11	0.86	-1.68	0.56	-3.54	0.38
Mean	0.00	0.07	0.99	-0.10	1.02	0.40	0.37
SD	1.24	0.01	0.09	3.10	0.24	3.60	0.10

3.3.1 Item difficulty and Wright Map

In Table 3, the item difficulty estimated values are listed in ascending order of difficulty and they range from -2.30 to 2.83. Their standard errors are all small and in the order of 0.1. The three most difficult items are items Q16 (item difficulty 2.83), Q12 (2.23), and Q34 (1.48), while the three easiest items are Q14 (-2.30), Q4 (-2.18), and Q9 (-2.07). These items are presented in Table 4.

Q16 and Q12 are questions dealing with fractions (number domain) and Q34 is a measures problem dealing with the perimeter and the area of a trapezoid. Although these three items are different in domain types, they share a common feature that the problem expression is a little complicated and needs to be well understood before accurate computation. For instance in Q16, multiple choice option A attracted most students who just added 3 grams of sugar to the existing 2 grams of sugar, and divided the sum by the 100 grams of water. In so doing, they did not take into account that the term sweet soup meant a mixture of water and sugar, and so the 3 grams of sugar needed to be included both in the numerator and in the denominator. Since sweet soup is a common diet in Hong Kong, it is unlikely that students committed the error for cultural reasons. Rather, the error was more likely to have arisen because students were confused by the complicated language expression of sugar, water, and sweet soup. The result suggests possible interference to students' mathematical abilities by way of their language abilities. On the other hand, we found these students were good at calculation of whole numbers (Q14), understanding concepts of numbers (Q4) and identifying positions (Q9). Each of these items had a success rate of over 90%. Overall, the Primary 5 mathematics test is an appropriate and valid instrument to detect the mathematics performance for Primary 5 students.

Table 4. The Three Most Difficult Items and the Three Easiest Items

The 3 most difficult items and the 3 easiest items		Item difficulty	Key
Q16	There are 2g of sugar and 100g of water in a bowl of sweet soup. If 3g of sugar is added, then sugar becomes（　　）of the sweet soup. A.　　B. $\frac{1}{21}$　　C.　　D. $\frac{5}{102}$	2.83	B
	Percentages choosing the Options for Q16 are: A: 48%　　**B: 14%**　　C: 17%　　D: 22%		
Q12	3 is added to the numerator of $\frac{1}{5}$. In order to make the fraction unchanged, which of the following should be done to the denominator? A. + 3　　B. - 3　　C. x 3　　D. x 4	2.23	D
	Percentages choosing the Options for Q12 are: A: 31%　　B: 11%　　C: 37%　　**D: 21%**		
Q34	Referring to the picture below, if we string a rope 58m long from the left side of the figure to the wall, what is the area cordoned off by the rope?  A. 480 m^2　　B. 240 m^2　　C. 580 m^2　　D. 290 m^2	1.48	B
	Percentages choosing the Options for Q34 are: A: 14%　　**B: 33%**　　C: 36%　　D: 17%		
Q9	Referring to the figures in Q8, Tom walks (…) to go from his home to the school and then walks (…) to go to the hospital. A. east, south　　B. east, north C. west, south　　D. west, north	-2.07	A
	Percentages choosing the Options for Q9 are: **A: 91%**　　B: 2%　　C: 5%　　D: 1%		
Q4	In which of the following numbers does “5” have the largest value? A. 15706439　　B. 16905347 C. 79654310　　D. 96574130	-2.18	A
	Percentages choosing the Options for Q4 are: **A: 92%**　　B: 1%　　C: 1%　　D: 6%		
Q14	14 children are playing together. They have spent $840 in total. How much has each child spent on average? A. 840 ÷ 14 = $60　　B. 14 × 14 = $196 C. 840 + 14 = $854　　D. 840 × 14 = $11760	-2.30	A
	Percentages choosing the Options for Q14 are: **A: 93%**　　B: 2%　　C: 2%　　D: 3%		

The item difficulty ranging from -2.30 to 2.83 indicated an appropriate difficulty level span. This is further supported by the Wright Map (Figure 1), which shows that the Primary 5 mathematics test was well matched against the sample. In the Rasch approach, a Wright Map is a visual representation of the distribution of the respondents' abilities in relation to the distribution of the item difficulties. Each # on the left panel represents seven students in this study and the numbers (e.g. Q16) on the right represent the items.

Items were plotted on the Wright Map according to their difficulties along the vertical straight line, which represented the mathematics ability scale, in the middle of the figure. Items at the top of the scale are more difficult items than those at the bottom of the scale. Students were plotted into the map according to their estimated mathematics abilities. More able students were at the top and less able students were at the bottom of the scale.

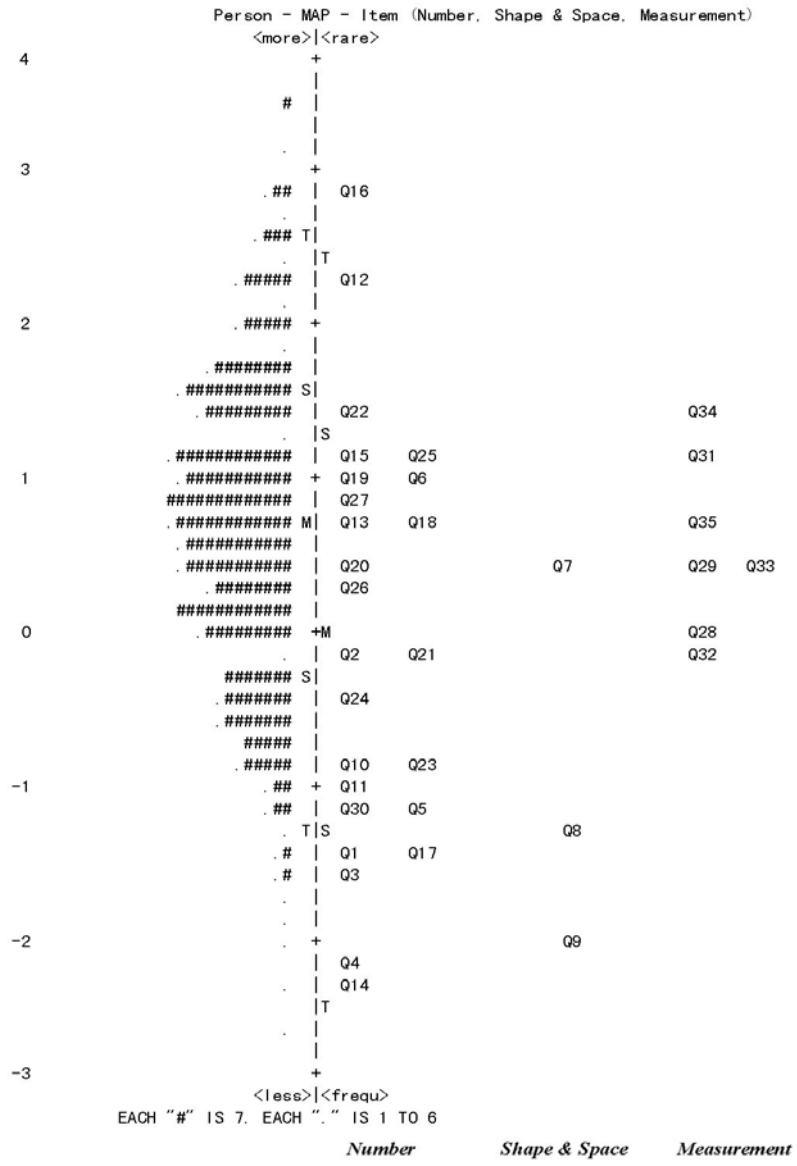


Figure 1. Wright Map of Items and Persons

In Figure 1, the mean (indicated by M on the right panel in Fig. 1) of item difficulty was close to the mean (indicated by M on the left panel in Fig. 1) of student ability within the first standard error (indicated by S). This means there was good alignment between student ability and item difficulty.

In Figure 1, items are listed as clusters according to different domains. The three items of the shape and space domain are listed in the middle of the right panel, while items of the number domain are in the left panel and items of measurement domain are in the right panel. On average, the three domains do not have substantially different mean difficulties. That is, no single domain presented more difficulty than another domain for the group of students.

Most of the items have difficulties near the mid-range of the vertical scale, which is around the mean item difficulty and within one standard deviation of the mean, while the two most difficult items Q16 and Q12 are located two standard deviations above the mean item difficulty. The three easiest items (Q9, Q4, Q14) are located outside two standard deviations toward the bottom of the scale. The majority of students on the left panel have ability levels above these three easiest items. It can be seen from this analysis that because of the good alignment between item difficulty and student ability, teachers can obtain a picture of the performances of the Primary 5 students as a whole, as well as their individual profiles.

3.3.2 Item fit

Fit statistics (Table 3) show the difference, or the residual, between the observed data and the estimated measure according to the Rasch model. Outfit mean square (MNSQ) is a mathematical function based on the mean of squared residuals. The computation of Infit MNSQ is similar to that of outfit except that each observation is weighted by its statistical information or the model variance. Statistical information is higher around the middle of the measurement scale where the observations tend to concentrate, and is lower towards the top and bottom tails of the scale where there are fewer observations. Infit ZSTD and outfit ZSTD are the standardized forms of Infit MNSQ and Outfit MNSQ

respectively (Bond & Fox, 2007; Linacre, 2011). Infit and Outfit statistics provide evidence of construct validity in Rasch measurement. Linacre (2011) recommended that Infit and outfit MNSQ values falling within the range of 0.5 to 1.5 can be taken as indication of good concordance between the data and the Rasch model. Items with goodness of fit values less than 0.5 or more than 1.5 are considered as having poor fit to the Rasch model. Table 3 shows that all items in the Primary 5 mathematics assessment, with the exception of item Q16 (which has an outfit MNSQ value of 2.01), have infit and outfit MNSQ values ranged from 0.56 to 1.36, which is well within the range of good fit.

When we referred to other statistics indexes, we found Q16 had a low point-measure correlation 0.07, while other items were all from 0.2 and 0.53. This result means that except for Q16, items in the assessment are internally coherent. As discussed in an earlier section, students might have failed Q16 because of a deficiency in common sense, which is not directly connected with knowledge and skills about fractions. However, mathematics thinking needs strictness and rationality. Item Q16 revealed deficiencies in the daily learning and training of Primary 5 students, and highlighted possible areas for enhancement in future instruction.

3.4 Differential Item Functioning

Differential Item Functioning (DIF) occurs when test-takers with same abilities in some measured latent trait have different probabilities of achieving a correct response to an item, which is considered an important issue in establishing test fairness (Wang, 2008). The magnitude of DIF signifies the extent to which the item parameter differs between different groups, such as gender, location, or social-economic status, even though the groups under comparison are of equal ability (Wang, 2008). In Rasch models, a DIF value of 0.5 logits or larger could be considered a substantial DIF (Wang, 2008). Recent research (Paek & Wilson, 2011) showed that for short tests with a small sample size, the Rasch model approach to DIF is more effective than the traditional approach of using Mantel-Haenszel probability. This analysis found that all items, with the exception of items Q9 and Q14, had very low DIF (less

than 0.5) for gender. The two exception items Q9 and Q14 had DIF values of 0.52 and 0.74 respectively, although both had low Mantel-Haenszel probability values (0.02 and 0.01 respectively). On this basis, it is concluded that items in the Primary 5 mathematics test assessment revealed no substantial DIF in gender, which means the mathematics assessment is fair to both boys and girls who take the test.

3.5 Person Diagnosis Information from Winsteps

From the teachers' perspective, diagnostic information on students' ability is perhaps the most precious. Discussions in earlier sections have already shown that the Rasch analysis using Winsteps (Linacre, 2011) can generate information on the mathematics ability of individual students. The Wright Map presents students' abilities alongside the items in the assessment and provides two frames of reference for the interpretation of each student: against the other candidates taking the same assessment, and against the ability requirement of the assessment items. Indeed, the Rasch analysis can generate at least two additional pieces of diagnostic information invaluable to teachers; namely, the Person-Kid-Map, and the Person Keyforms. These are discussed in the following sections.

3.5.1 Person-Kid-Map (PKMAP)

The Person-Kid-Map (PKMAP) is a graphical display of the zone of proximal development and response pattern of each individual student. An example taken from this study is the PKMAP of student number five presented in Figure 2. In the PKMAP, the estimated ability level of the student is represented by "xxx." Using this estimate as a focal point, the PKMAP divides the figure vertically into two panels, and horizontally into three panels, resulting in six regions of the graph. Located in the left panel are those items which the student answered correctly and items in the right panel are those which the student answered incorrectly. Items in the top panel were difficult for the student because their difficulty levels are at least 0.5 logits more than the ability level of the student. Items in the top panel are easy for the student because their

difficulty levels are at least 0.5 logits less than the ability level of the student. The middle panel contains items with difficulty levels within ±0.5 logits of the student's ability level. This is the Zone of Proximal Development of the student. Utilising these categorisations by the panels, and counting clockwise from the top-right region, the six regions in the PKMAP are:

1. **Non-mastery Future Goal Region**: Items that are difficult for the student, who answered them incorrectly. The items involved in this region (Q12 in this example; Fig. 2) give direction for future learning goals of student.
2. **Zone of Proximal Development Need-Scaffolding Region**: The student answered items in this region incorrectly (Q22, Q31, Q25, Q6 and Q35 in this example). This is the region where the student has not yet mastered the necessary knowledge and skills required to answer the items correctly, but if given support, the student will be able to achieve mastery.
3. **Carelessness/Special Learning Needs Region**: Items (Q33, Q28, Q23, Q10, and Q11) are easy for student, but the student still answered them incorrectly. The teacher should check items in this region carefully to see if there is something wrong with the items themselves. If the answer is negative, the teacher should find out if the student has made careless mistakes, lacks examination skills, has special learning difficulties (e.g., dyslexia), or misconceptions, and then seek to provide appropriate remediation.
4. **Mastery Region**: This is the region of mastery. Items (Q7, Q20, Q29, Q26, Q2, Q21, Q32, Q24, Q5, Q30, Q8, Q1, Q17, Q3, Q9, Q4, and Q14) in this region are easy for the student, who answered them correctly.
5. **Zone of Proximal Development Need-Consolidation Region**: Although the student answered items (Q15, Q19, Q27, Q13, and Q18) in this region correctly, learning is shaky and needs consolidation.
6. **Pleasant Surprise Region:** Items (Q16 and Q34) in this region are beyond the student's ability level. Nonetheless, the student answered them correctly. The teacher has to check to see if there are elements of luck, dishonesty, or that the student has learned the topics involved at other settings such as a tutorial school.

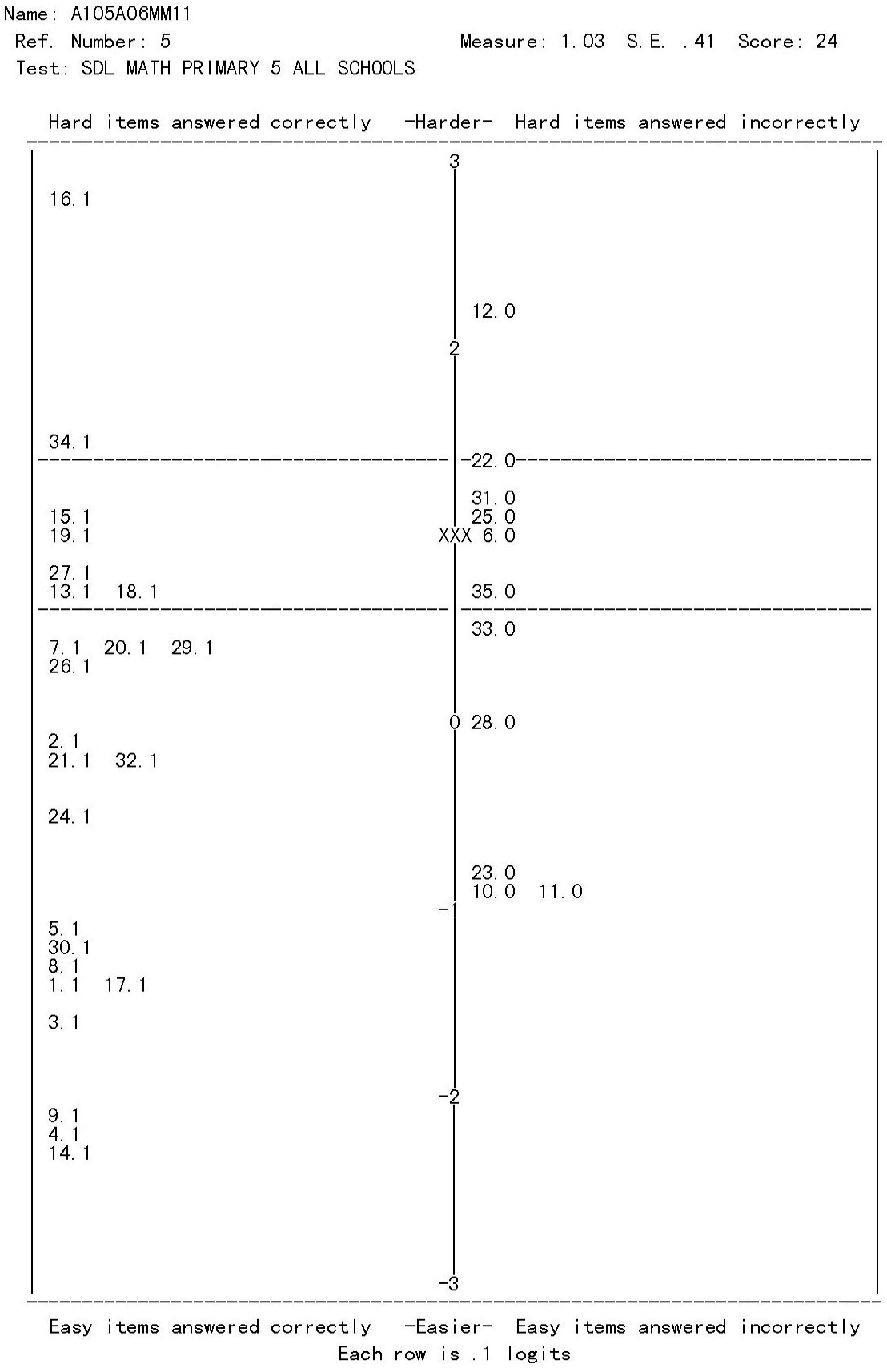


Figure 2. Person Diagnostic PKMAP

3.5.2 Person Keyforms

Person Keyforms is another graphical device produced by the Winsteps software (Linacre, 2011) that provides diagnostic feedback information for individual students. An example from this study is presented in Figure 3. In the Person Keyforms, the Rasch measurement scale is presented horizontally and extends from left (less able students) to right (more able students). Actual responses of the student to items in the assessment are printed vertically at the student's ability location.

In the example from this study, student number five has a person measure of 1.03 logits. Based on the student's person measure, the expected responses to Q16, Q12, Q34 . . . Q14 are, A, B, A . . . A, respectively and these predicted responses are printed in a column in the Person Keyform of this student at the location of their ability estimate. Observed responses to the right of the estimated ability column are those items that are more difficult than the ability of the student. The further away an observed response from the estimated ability column, the more of a discrepancy there is between the predicted and actual response.

The actual responses made by the student are printed in two forms; namely, either (a) with a period before and after the character of the response (for example, .B.) in situations where the actual response is not equal to, but not too highly unexpected, from the predicted response, or (b) with round brackets before and after the character of the response (for example, (B)) in situations where the actual response is highly unexpected compared to the predicted response. For instance, in the example presented in Figure 3, the student is predicted to choose option A for Q16, but the student has chosen option B instead, and so the symbol (B) is used in the figure. By referring to the horizontal ability scale, it can be seen that only students with an ability estimated at about 3.8 logits are predicted to choose option B in Q16, so it is highly beyond expectation that student number five, whose ability is only 1.03 logits, would choose this option. The teacher might want to investigate the reasons behind such a large discrepancy. Would cheating, luck, or other reasons be the answer?

Similarly, according to the Rasch model, student number five is estimated to choose option A for Q10, but instead the student has chosen option D. Only

students with ability estimated at around -2 logits, which is much lower than the ability estimate of student number five, would make such a choice for Q10. This choice of student number five is therefore highly unexpected and the choice is represented by (D) in Figure 3. Again, the teacher might want to find out more about such a large discrepancy. Would the reason be carelessness, under-preparation, lack of test-taking skills, previously unidentified misconceptions, or special learning difficulties?

By inspecting the actual and predicted options of each student, the teacher would get very specific information, based on which the teacher is then able to strategise the next course of action in support of the student's learning.

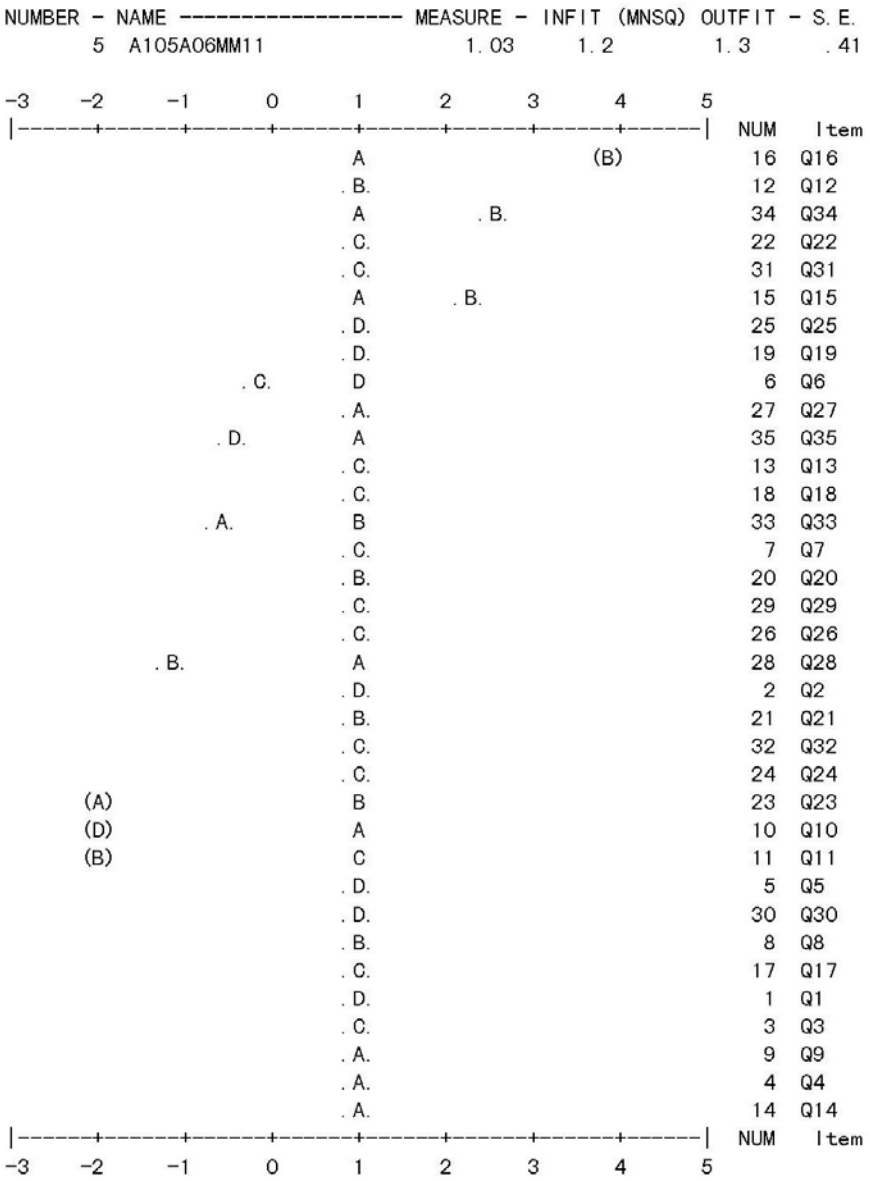


Figure 3. Person Keyform of Student Number Five

4. Conclusion

This study was part of a larger longitudinal study on the effect of feedback and self-regulated learning on mathematics achievement of primary students. The current study focused attention on the development of mathematics assessment for Primary 5 students in Hong Kong. It was undertaken to optimally use diagnostic information generated from Rasch analysis of assessment data in order to inform teaching and learning. The Rasch analysis conducted in this study made use of the Winsteps software (version 3.72.3) (Linacre, 2011) and some of the outputs were unique features of the software.

The Rasch measurement approach was used to develop and validate a 25-item mathematics assessment for Primary 5 students in Hong Kong in this study. Data analysis with Winsteps (version 3.72.3) (Linacre, 2011) showed the mathematics assessment is underpinned by a unidimensional construct, has acceptable item and person reliabilities, satisfactory item fit indices and item difficulties, good alignment between item difficulty and student ability, and has no gender DIF. The analysis undertaken demonstrates the procedures necessary for scientific inquiry into the validity of test scores, which are of key importance in all forms of testing (Messick, 1989; Ariffin, Omar, Isa, & Sharida, 2010). Establishing validity in test scores is particularly important to teachers in their implementation of assessment for learning because test scores form the basis of subsequent instruction.

As illustrated in this study, Rasch analysis can generate rich and imperative information for teachers about the assessment items and the students taking the assessment. Multiple frames of reference are available to the teachers to get both specific and holistic understanding of each student’s performance profile. The frames of reference include each individual student, performance of the entire group being assessed, difficulty of individual items, as well as all the items that constitute the test.

As an illustration, the study discussed selected items at the two extreme ends (most difficult and easiest items), as well as selected student responses in

order to show teachers how to detect issues arising from the item- or student-levels, and how to collect information for further teaching and remedial instruction. Amongst the information generated from the Rasch analysis, the diagnostic information provided by the Person-Kid-Map (PKMAP) and person Keyforms are most crucial for the identification of evidence regarding individual students' achievement. Through the PKMAP and the person Keyforms, teachers could get to know the Zone of Proximal Development of each student, areas of mastery and areas needing remediation.

5. References

Ariffin, S. R., Omar, B., Isa, A., & Sharif, S. (2010). Validity and reliability multiple intelligent item using Rasch measurement model, *Procedia Social and Behavioral Sciences, 9*, 729-733.

Arslan, S., & Ozpinar, I. (2010). Assessment in mathematics course and evaluating primary school coursebooks in terms of assessment, *Procedia Social and Behavioral Sciences, 2*, 4157-4163.

Babakhani, N. (2011). The effect of teaching the cognitive and meta-cognitive strategies (self-instruction procedure) on verbal math problem-solving performance of primary school students with verbal problem-solving difficulties. *Procedia Social and Behavioral Sciences, 15*, 563-570.

Berry, R. (2008). *Assessment for learning*. Hong Kong: Hong Kong University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Bulut, M. (2007). Curriculum reform in Turkey: a case of primary school mathematics curriculum. *Eurasia Journal of Mathematics, Science & Technology Education, 3*(3), 203-212.

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

Earl, L., & S. Katz. (2013). Getting to the core of learning: Using assessment for self-monitoring and self-regulation. In M. M. C. Mok, (Ed.), *Self-directed Learning Oriented Assessments in the Asia-Pacific,* pp. 123-137. Dordrecht Heidelberg, New York, London: Springer.

Embretson, S. E. (1996). Item response theory models and spurious interaction effects in factorial ANOVA design. *Applied Psychological Measurement, 20*, 201-212.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81–112.

Kluger, A., & DeNisi, A. (1996). The effect of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254-284.

Lee, K. O. (2012). Physical Education in higher education in Hong Kong: The effects of an intervention on pre-service sports coaches' attitudes towards assessment for learning used in sports. In M. M. C. Mok (Ed.), *Self-directed Learning Oriented Assessment in the Asia-Pacific*. New York: Springer.

Linacre J. M., & Tennant A. (2009). More about Critical Eigenvalue Sizes (Variances) in Standardized-Residual Principal Components Analysis (PCA). *Rasch Measurement Transactions, 23*(3), 1228.

Linacre, J. M. (2011). *A user's guide to Winsteps/Ministep Rasch-model computer program*. Chicago, IL: Winsteps.com.

Mok, M. M. C. (2010). *Self-directed Learning Oriented Assessment: Assessment that Informs Learning & Empowers the Learner*. Hong Kong: Pace Publications Ltd.

Mory, E. H. (2004). Feedback research review. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745-783). Mahwah, NJ: Lawrence Erlbaum.

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegermann, D. Leutner, & R. Brunken (Eds.), *Instructional design for multimedia learning* (pp. 181-195). Munster, NY: Waxmann.

National Mathematics Advisory Panel (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*. U.S. Department of Education: Washington, DC.

OECD (2010). PISA 2012 Mathematics Framework. Paris: OECD Publications. Retrieved 23 May 2013 from http://www.oecd.org/dataoecd/8/38/46961598.pdf

Paek, I., & Wilson, M. (2011). Formulating the Rasch differential item functioning model under the marginal maximum likelihood estimation context and its comparison with Mantel-Haenszel procedure in short test and small sample conditions. *Educational and Psychological Measurement, 71*(6), 1023-1046.

Raîche, G. (2005). Critical eigenvalue sizes in standardized residual principal components analysis. *Rasch Measurement Transactions, 19*(1),1012.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Institute of Educational Research. (Expanded edition, 1980. Chicago: The University of Chicago Press.)

Savage, J. (2011). *Cross-curricular teaching and learning in the secondary school*. New York: Routledge.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 677–680.

Wang, W. C. (2008). Assessment of differential item functioning. *Journal of Applied Measurement, 9*(4), 387-408.

Weiland, C., Wolfe, C. B., Hurwitz, M. D., Clements, D. H., Sarama, J. H., & Yoshikawa, H. (2012). Early mathematics assessment: validation of the short form of a prekindergarten and kindergarten mathematics measure, *Educational Psychology: An International Journal of Experimental Educational Psychology, 32*, 311-333.

Authors' email: Jingjing Yao jingjing@ied.edu.hk

Magdalena Mo Ching Mok (Corresponding Author) mmcmok@ied.edu.hk

The Assessment as Learning (AaL) Framework for Teaching and Learning – The AaL Wheel

Rita Berry
Hong Kong Institute of Education

Introduction

Currently, there is an increasing need in society for individuals with the ability to deal with dynamic challenges, to be able to constantly learn and to produce effective and innovative solutions through critical thinking, which is one important aspect of Learning How To Learn (LHTL) (Crick, 2007). The LHTL theory maintains that, essentially, there is an overriding need for the individual to develop not only knowledge-based understandings of materials in the classroom, but more importantly, metacognitive skills which help them manage their learning. Learners need skills which they can apply to other challenges in life, for example, the ability to observe, analyze and learn, what Black et al. (2006) refers to as 'lifelong learning' and what Garrett (2011) argues as elements in widening students' capability in learning. LHTL depicts a learning process during which students learn how to tackle learning course materials in addition, all the while developing metacognitive skills and knowledge applicable to other challenges beyond the classroom, from the workplace to everyday life (James et al., 2007). The last half a century saw a number of education reforms that took LHTL as a main overarching educational goal on the reform agenda. The agenda highlights the significance of helping students take charge of their learning. The assessment policies stipulated in these reforms suggest the use of self- and peer-assessment to increase learners' metacognitive abilities (Berry, 2011b). The assessment practices in many classrooms, however, are still very much less effective than others in promoting the kinds of learning outcomes that are needed by students today and in the future (James, 2006).

The pedagogies used by the teacher could impact significantly on the quality of student learning (Bronkhorst et al., 2011). Education constantly looks for innovations that bring about improvement in teaching with the express purpose of improving student learning. What is quite noticeable in all this transformation is that those in the midst of it need catalysts such as frameworks, models, advice, and other guidance. Frameworks are very useful to effect improvement as they provide a useful way to examine possibilities (Bonk & Dennen 2007; Eun, 2011). They help focus attention on the characteristics of teaching and learning that are salient to each individual theory. They provide systematic, well delineated ways of describing and explaining the teaching/learning process, often with the support of a distinct vocabulary representative of underlying epistemological and ontological perspectives (Young, 2008). Marsh (2009) stresses that a well-developed framework should provide strong links between theory and practice and that it should be inspiring to teachers. Developing a framework demands strategic alertness (Entwistle and Walker, 2000). Strategic alertness requires a shift of attention to previously unattended-to factors through the use of a selected frame of reference. The framework designed will thus help refocus attention and provide a fresh way of conceptualizing teaching and learning. The Assessment as Learning (AaL) Framework proposed intends to look at teaching and learning from a new perspective – the perspective of assessment in respect to students taking an active role of their learning.

Teachers' and Students' Roles in Assessment as Learning

There are three widely recognized assessment approaches in current literature, namely, Assessment of Learning (AoL), Assessment for Learning (AfL) and Assessment as Learning (AaL) which reflect three different focuses of learning conceptions. AoL represents the assessment conception of measurement. Judgments of performances are taken at the end of learning. Assessment practices in many educational contexts are often inclined to link teaching and learning with this kind of assessment. Both AfL and AaL take the learning process as being significant and emphasise the roles of assessment in

supporting learning. What makes AaL different from AfL is that AaL places special attention on the role of the learner and promotes active engagement of learners while AfL places stronger emphasis on the role the teacher plays in promoting learning. AaL could be said to be an “assessment as learning to learn paradigm” while AfL an “assessment in support of learning paradigm” (Berry, 2008a).

Earl (2003) says that AfL can go a long way in enhancing student learning. By introducing the notion of AaL, the intention is to extend the role of AfL by emphasising the role of the student, not only as a contributor to the assessment and learning process, but also as the critical connector between them. The student is thus the link between teaching and learning. Being an active, engaged and critical assessor, the student makes sense of information, relates it to prior knowledge, and deliberates the strategies and skills involved in taking their learning forward. S/he self-analyses, self-references, self-evaluates and self-corrects in the learning process. These and other metacognitive strategies help him/her raise their awareness of what s/he is doing so that s/he can plan what s/he needs to do to move learning forward (Berry & Adamson, 2011). Students’ roles may also include working out what their teachers expect of them and doing it well. Brookhart (2001) calls this knowing the art of ‘studenting’.

For the teachers, this entails a major change of their roles, from a presenter of content to a practitioner of more productive pedagogy, involving shared responsibility for learning by student and teacher (Klenowski, 2007). Vygotsky’s conception of the Zone of Proximal Development (1978) suggests that the aim of teaching is to encourage the learner to be ever more independent from the teacher. AaL may involve the teacher aligning to a set of procedures that allow the learner to move forward independently in the required learning. During the learning process, students are helped to use assessment information to set goals, make learning decisions related to their own improvement, develop an understanding of what quality work looks like. They self-assess, seek feedback from their peers and teachers, and reflect on how these take them to the next step of learning (Chappuis and Stiggins, 2002). Although AaL concepts have been in discussion for quite some time, there is little information on how

the concepts can be transferred into actions. The AaL Framework for teaching and learning, or the AaL Wheel, is subsequently proposed with an aim of bridging the gap between theory and practice.

The Basic Structure of the Assessment as Learning Framework

Weaver and Farrell (1997, p.45) identify four essential elements in developing paradigms, models or framework – assumptions/beliefs, values, vocabulary, and behaviors/activities.

- Assumptions/Beliefs: one's perceptions about what is real or true; the foundation for the behaviors and activities that are chosen by him or her.
- Values: one's views of what are important to him or her; these become the basis for setting priorities and making choices of what goals to pursue and how problems are to be solved.
- Vocabulary: the words that are used to communicate, for example, about how problems are posed and solutions described.
- Behaviors/Activities: are those worked out approaches and solutions that display the world view as a coherent whole.

The framework proposed in this article is built upon the above mentioned concepts, using the terminology of Contextual, Societal, Communication, and Action Domains. The term *Contextual Domain* is chosen based on the understanding that no matter how widely recognized some education conceptions may be, political, economic and cultural contexts should play a part in policy development and implementation in individual educational settings. In the AaL Framework, the contextual domain represents the assumptions and beliefs of the government, often expressed explicitly in official documents and hence become the policies to direct and govern the activities to be designed by the personnel working in different education sectors in the society. The term *Societal Domain* is selected to acknowledge the influence that society may exert on educational policies. This domain describes the values, including shared perceptions and expectations, across one or more groups in the society. The domain covers the beliefs and philosophies of these groups. Whether in-line

with the vision stipulated in the government policies or not, these perceptions would consequently govern the stakeholders' decisions and the actions to take in delivering the government directives. The term *Communication Domain* is used for its self-explanatory function – expressions and negotiations of social meaning. This domain contains the vocabulary or the words that are used to communicate, for example, about how problems are posed and solutions described. It is the language, including common terminology and understandings through which the beliefs are conferred and understood. *Action Domain* is picked as the term implies change and progression (Angyal's System Dimensional Model (Angyal, 1941)). Linking theory to practice is often regarded as a challenge (Berry, 2008b; Munns, 2005; Rose, 2002). Actions are the catalysts to link conceptions into classroom practice. The Action Domain of the AaL Framework describes the behaviors or activities happening at the implementation frontline. The activities can consist of simple, singular tasks that are carried out on a daily or regular basis.

The Contextual Domain is core to the framework. The stored information in the context is retrieved by the other three domains for deliberation, interaction, and delivery of actions. The context domain, in turn, draws observations from the societal, communication and action domains in order to update the information stored within itself, acting as a dynamic archive. For example, if the framework was the clothing industry, before a skirt is designed and manufactured (*Action Domain*), the factory draws upon existing knowledge of information from the *Contextual Domain,* which can be the policies and regulations set by the government and respective organizations. To determine what designs and manufacturing processes would be more successful and what costs would result, the factory will need to refer to the *Societal Domain* to check what views the society has on skirts. The views may include public opinion on the appropriateness of, fashion trends about, and gender implications towards skirts. At the same time, the factory will need to be sure that the terminology used, such as the term 'skirt' and other messages, including problems and solutions, in relation to the design and production are established and can be communicated with relevant parties (the *Communication Domain*). *The Action*

Domain represents the actions, for example, sewing a hem on a skirt in a garment factory. Methodologies in this domain are a combination of many ‘do’s, for example, how to run a factory or, even the garment industry as a whole, such that there is a tangible product. After production of the skirt, the information stored in the *Contextual Domain* would be updated with the observations that were made across the board – how successful the production process was (*Action Domain*), whether the manufacturing process caused any public disturbances or damage (*Societal Domain*), if the language of the industry had changed over time (*Communication Domain*), etc., allowing for the development of all four domains. Figure 1 below presents the basic structure of the four domains in the AaL Framework.

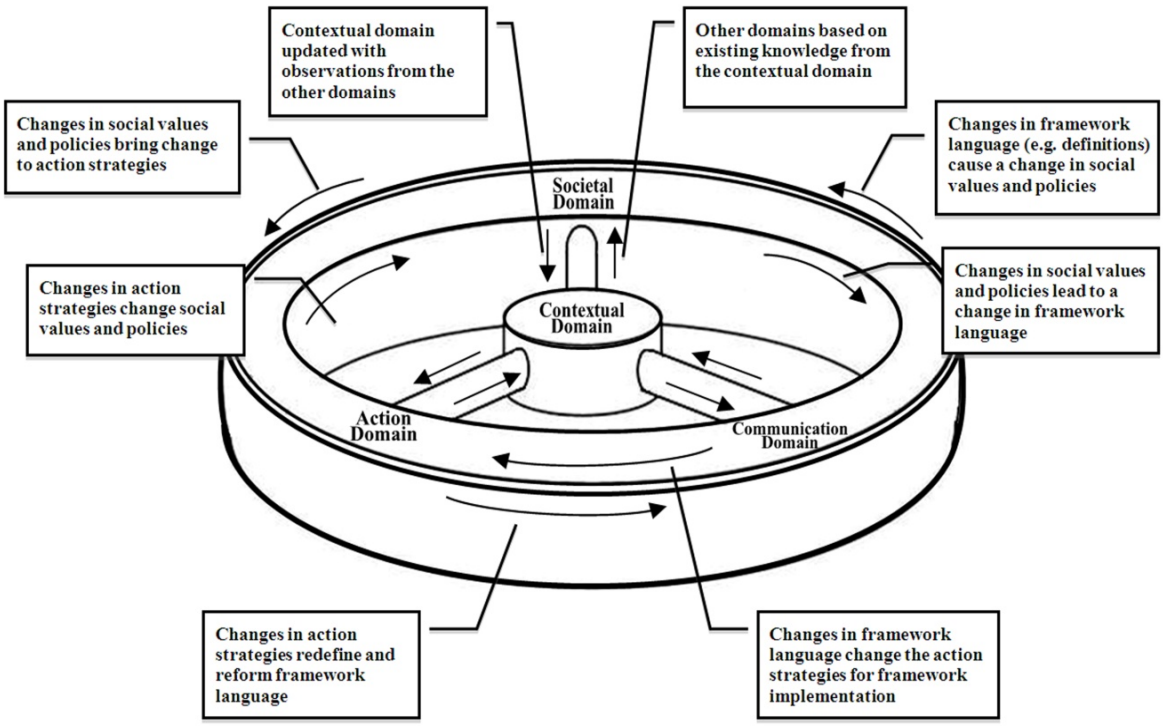


Figure 1. The basic structure of the Assessment as Learning (AaL) Framework

The Assessment as Learning Framework for Teaching and Learning

Though the structure of a framework can be varied, in constructing a new framework, it is deemed important to know the reality of the situation and the parameters of the expectations. According to Stansfield (2001), the approach towards constructing a new framework should begin with the definition of the desire (expectations) – what one believes to be the truth. This is followed by observing what the reality is in this particular domain, and a comparison between the two.

The Contextual Domain

In the *contextual domain* of the AaL Framework, a desirable situation is that government policies embed the development of "Learning how to learn" in students, with a vision on making assessment an agent for activating student learning. There will be emphasis on taking assessment as a process of metacognition for students. Assessment policies will revolve around learning process, taking student-centeredness as core. Official documents may contain a section with detailed guidelines on how teachers' roles can be de-centered to students in the purpose of making students take more responsibility for their own learning. Drawing on the government assessment policies, the assessment policies for different education sectors can focus on developing students' ability to learn, for example, enhancing their critical thinking, analyzing and general skills. More directions can be given to create opportunities for students, either in groups, pairs or individually, to reflect and analyze their own performance and subsequently work on the next steps for learning. Assessment is seen as a partnership between students and teachers, where the former is not only active but also responsible for their own learning and assessment while the latter acts as a facilitator, providing opportunities for learning and self development and guidance when necessary. Teachers are to be supported to use assessment to promote LHTL. The government may provide funding for teacher training and offer additional resources to reduce teacher-student ratio to enable more classroom interactions between the teacher and the students. Through an agora

between the government and the personnel from different educational sectors and contexts, the development of LHTL in connection with assessment may be established.

In many education *contexts*, assessment policies at government level are more focused on giving detailed descriptions of rules and regulations of examinations, mechanisms for marking and moderation of scores, avoiding plagiarism and cheating, etc. There is typically little or no discussion on how assessment can contribute to learning, in particularly how it may help students to become active learners (Boud, 2007). Many countries, particularly in recent years, embarked on educational reforms with LHTL highlighted as the way forward for the overarching educational aim. In their official documents, these countries provide the overarching policy to link assessment with learning but generally miss concrete ideas on how to make assessment a useful tool to promote active learning in parts or all of their policies and guidelines. At schools, similarly, assessment policies are usually presented as a form of official document which may not contain a section with detailed guidance on how, and on what bases, judgments about the quality of student performance should be raised through increasing learners' self awareness of their learning. Assessment policies revolve around processes like examinations, grading, as well as quality assurance, with less emphasis on linking assessment with learning (Saddler, 2005; Berry, 2011b).

The Societal Domain

Regarding the *Societal Domain*, one would like to see that society perceives assessment as a tool to help develop potentials and abilities in dealing with challenges in life. Employers would be more aware of the fact that grades and numbers shown on the qualification documents can only depict some of the qualities of their staff. Parents would come to understand that assessment is not simply a tool for measuring their children's performance and abilities or checking the return on their financial and emotional investments, but also a tool to help their children develop metacognitive skills for their future. They will recognize that assessment is both a responsibility of the teacher and the student,

with an emphasis on the latter, and hence help with the psychological and mental preparation of their children towards dealing with self-assessment while the child is under their care. Teachers, similarly, are to see assessment as not only their responsibility, but also that of the students. As such, assessment requires teachers not only to allow, but to encourage and facilitate student participation in monitoring and critiquing their own work and progress, and by association, their own learning. They would see their role in assessment as facilitators, helping students, for example, understand the criteria to assess themselves, self reflect their performance and make educated decisions on what to do next in enhancing their learning. Students will realize that assessment procedures are opportunities for them to develop LHTL. They would learn that assessment is a tool to help them monitor learning and understand what learning stage they are at. With the updated information, they will direct efforts towards improving their work. They will acknowledge the importance of the internal processes of assessment in their own learning process and uses external assessment as a necessary but relatively auxiliary form of support.

In recent years, there has been an increased advocacy across education sectors on having students become more active players in assessment as part of the process of learning (Berry, 2006; Craddock & Mathias, 2009). Accordingly, assessment can be used as a tool to help develop individuals into people who are more able to deal with different challenges. Although this view of assessment is gradually gaining more recognition, the larger public still places the values of assessment strongly in certifications, qualifications and accountability (Berry, 2011a; Knight, 2003; Murphy, 2006). Employers often make decisions based on the grade and marks achievements of the candidates. Many parents see assessment as a measurement of the return on their financial investment in their offspring (Race, 1999). Teachers are under pressure to feed their students with a certain amount of academic and community needs information and the simplest way to do it is to adopt the old and traditional approaches to teaching. Assessment methods are not tailored to student needs and students are rarely, if ever, given an active role in their own assessment (Carless, 2006; Gibbs & Simpson, 2004). Teachers come to see teaching and students' learning as

something done to them rather than something teachers and students can be in control of (Watkins et al., 2007). Students are very used to taking assessment as the teacher's responsibility. This is particularly likely to happen if their teachers also believe that this is all they are capable of doing (Kember, 2004).

The Communication Domain

The *Communication Domain* of the AaL Framework would like to see in the society an increased consensus of the function of activating learning in assessment. Assessment is communicated across different parties as a tool which students and teachers can use to enhance learning and develop students' metacognitive abilities. Both teachers and students can be involved in the assessment and learning process. Teachers are the supporter in students' learning process, guiding and helping them to develop the right mentality to learn. Through various kinds of dialogues between teachers and students, students are helped to understand quality of work. Students are given opportunity to establish the ability to check their progress against standards, and make plans to improve when the standards have not been met (Assessment Reform Group, 2006; Expert Group on Assessment, 2009).

The society, however, may not see assessment in the same way. Nicol & Macfarlane-Dick (2006) point out that students, more often than not, take on a passive role in the assessment process. Klenowski (2009) says that there are often variations in interpretation and terminology of assessment. For example, assessment might be treated as an equivalent of tests and examinations. Assessment may be interpreted as merely a tool to generate grades/marks at the end of the learning process. Teachers could be understood as judges to the final product of learning. Formative Assessment (FA), usually used interchangeably with Assessment for Learning (AfL), may be misinterpreted as testing students continuously with a keen focus on checking learning outcomes at the end of numerous teaching intervals. AfL advocators would prefer FA or AfL interpreted as assessing students continuously to understand how students learn so that timely support can be given to them. The interpretation of grades and marks is another example. Grades and marks are often treated as a direct conversion of

feedback. Grades or marks are better understood as one form of feedback, which when used alone, are judgmental of performances and do not give directions for improvement.

The Action Domain

In the *Action Domain,* when applied to day-to-day teaching, teachers can use multi-faceted and various types of assessment to provide students with different kinds of learning experiences. Teachers select, develop, or adapt assessment methods for use based on students' learning needs and different learning styles. They provide self-assessment opportunities for students and help them develop good quality self-assessment. Teachers can try using smaller tasks to make timely feedback possible. Feedback, informal or formal, should be constructive which aims at, in addition to acknowledging students achievements, helping students understand what has been achieved and how to advance from there. Through dialogues or written forms of communication, teachers help students identify the types of strategies which are useful for their learning. For students, the actions will entail a greater involvement in their own learning. There will be opportunities for students to practice assessing skills. They will be supported to understand different standards, for example, the standards required by the teacher and what 'higher standards' entail. Students set their learning goals for assignments and choose strategies to complete the assignments. They record the progress and make notes of the issues that worth attention. They may then work out how they could improve their work.

Shipman, Aloi and Jones (2003) point out that, in many classrooms, students are given a minimal or non-existent role in assessment. Formative assessment and feedback are still largely controlled by and seen as the responsibility of teachers and feedback is still generally conceptualized as a transmission process (Nicol & Macfarlane-Dick, 2006; Hargreaves, 2011). Syllabuses are often provided to students with minimal or insufficient explanation of assessment criteria. There is very little communication of assignment requirements between the teacher and students. Feedback is frequently given in the form of grades with very little communication to students what the grades imply and how students can move on to the next level of learning. The common phenomenon in this kind of classroom is that students

are not able and/or not willing to take control of their learning. This could have resulted from being treated incessantly as passive participants throughout their major time of education, when assessment used is largely traditional, number-based, with specific purposes such as grading, selection, certifications, and qualifications. This kind of assessment usually associates with standardized summative assessments frequently in the form of MC, short answers, etc (Berry, 2010). The learning mode adopted by students usually reflects their mentality of learning. Provided with a passive learning environment, students tend to rote learn. They may perceive that this is what their teachers and their course expect them to do, or that it is what the assessment requires. Students are unaware of what active learning is and how assessment can be of help in making learning active. To meet academic requirements, they normally streamline their study methods and study for the tests, often causing surface learning (Gibbs, 1999). The following table (Table 1) summarizes the ways on how the current situations could be made better.

Table 1. Assessment as Learning (AaL) Framework for Teaching and Learning – The Targets

AaL Domains	**The Targets**
Contextual:	**<u>Policies by the government:</u>** - Embed the policy in the development of "Learning how to learn" in students, with a vision on making assessment an agent for activating student learning; - Share the vision with the personnel in different education sectors; - Provide resources and concrete ideas for teacher training and allocate greater funding to reduce teacher-student ratio so as to facilitate AaL development. **<u>Policies at the educational frontline:</u>** - Take developing students' abilities to learn how to learn as one main focus of assessment; - Specify in the policies that assessment is also a process of metacognition development in students; - Encourage partnership between students and teachers, where the former is not only active but also responsible for their own learning and assessment while the latter acts as a facilitator, providing opportunities for learning and self development and guidance when necessary.

Societal:	**Views of the society:** - Sees assessment as a tool to help develop potentials and abilities to deal with challenges in life. **Views of the teacher:** - Sees assessment as an agent for enhancing student learning, in addition to its other functions such as certification; - Sees assessment a shared responsibility between the teacher and the student; - Sees assessment a dialogue between the student and the teacher regarding student learning; - Sees the teacher's role in assessment as facilitator, helping the student learn, for example, what standards to meet, how to self-assess, and what kinds of strategies to take to move learning forwards. **Views of the student:** - Sees assessment as an opportunity to take responsibility and action in learning; - Sees the teacher a supporter of their learning but understand the support will decrease over time; - See assessment as a tool to help monitor learning and understand what learning stage s/he is at. With the updated information, the student will direct efforts towards improving his/her work. **Views of the parent:** - Sees assessment, in addition to its many other functions, as a tool to help their children develop metacognitive skills for their future; - Sees assessment as both a responsibility of the teacher and the student, with an emphasis on the latter. **Views of the employer:** - Understands that the assessment results shown on qualification documents can only tell part of the abilities of his/her staff.
Communi-cation:	**Learning as:** a process that allows students to take control of their learning. Students can set their own learning goals, check their progress against standards, and make plans to improve when the standards have not been met. **Assessment as:** a tool through which students and teachers can use to enhance learning and develop students' metacognitive abilities. Both teachers and students can be involved in the assessment and learning process. **Teachers as:** facilitators in the learning and assessment process, guiding and helping students to develop the right mentality and skills to learn and assess. **Students as:** active participants in the process of assessment and stewards of their own learning. They are able to set their own learning goals and select the strategies which are helpful for their learning. They know how to self and peer assess and understand the purpose of self- and peer-assessment.

Action:	**For teachers:** ***Learning opportunities, for example:*** - Use multi-faceted and various types of assessment to provide students with different kinds of learning experiences; - Select, develop, or adapt assessment methods based on students' learning needs, for example, different learning styles; - Provide self-assessment opportunities for students and help them develop good quality self-assessment. ***Feedback and support, for example:*** - Try using smaller tasks to make timely feedback possible. Feedback can be informal or formal; - Give constructive feedback that helps students understand what and how to advance; - Through dialogue, help facilitate students planning of strategies to improve learning; - Acknowledge students' achievements.
	For students: ***Self involvement, for example:*** - Understand different standards including the standards required by the teacher; find out what 'higher standards' entail; - Establish own goals for assignments; - Choose strategies deemed appropriate to tackle the assignment; - Record entire progress, making note of any issues arising and have been resolved; - Write self reflections; - Write peer reviews; - Prepare questions based on self, peer and tutor evaluations prior to communication; - Communicate with the teacher directly or through self reflections after the completion of the exercise; - Modify learning strategies appropriately. ***Student development, for example:*** - Seize opportunities to practise assessment skills; - Learn how to set appropriate learning goals and reflect or report on own progress against the goals; - Involve in group work and practise peer assessment.

How Students Can Be Helped to Become Active in Learning

The four domains of the AaL Framework for teaching and learning are engaged in a dynamic relationship, with constantly evolving definitions, understanding of social perceptions and plans of actions in various contexts, aiming at helping learners take an active role in their learning. The *Context Domain*, represents the policies that convey the beliefs and assumptions of AaL

in documents, provides the three other domains with AaL information through policies, directives and guidelines but is constantly updated with observations from three other domains to make AaL understandable in the society and implementable at the education frontline. The *Communication Domain* establishes the definition of AaL terminologies within the framework, such as that of assessment and active learning, setting the language through which the framework will be communicated. Overarching social attitudes, including the perceptions and attitudes of educators, administrators, teachers and students, parents and employers, are contained within the *Societal Domain*. The change in framework language (e.g. definitions) may cause a change in social values in assessment. Learning concepts which highlight student-centredness relates assessment as the activities used by students for gathering information, analyzing and interpreting it, drawing inferences, making wise decisions, and taking appropriate actions in the service of one's learning. Through different channels, this interpretation of assessment is communicated, which may gradually make an impact on how the society sees assessment. A change in the perception of assessment can change assessment practices, reflected in the *Action Domain*. The *Action Domain* represents the responsibilities, roles and characteristics of the student and the teacher. These include the strategies and implementation of AaL practised by the teacher. Students are helped to become active participants in the process of assessment and stewards of their own learning, setting their own goals and developing the skills necessary to achieve them through self- and peer- assessment as well as teacher assessment. Students are allowed to take control of their learning and are helped to set realistic and useful learning goals. When using the new action strategies in the classroom, teachers have a direct understanding of the kind of impact of their actions on students, which may in turn make a change in their perceptions of assessment or even redefine the assessment language in itself for communication with their counterparts. The context domain, as mentioned previously will draw the information from the three domains and update the policies, directives and guidelines which suit the needs of the education community. The figure (Figure 2) below, which builds on the basic structure of the AaL Framework (Figure 1), presents the key features of how students can be helped to become active in learning.

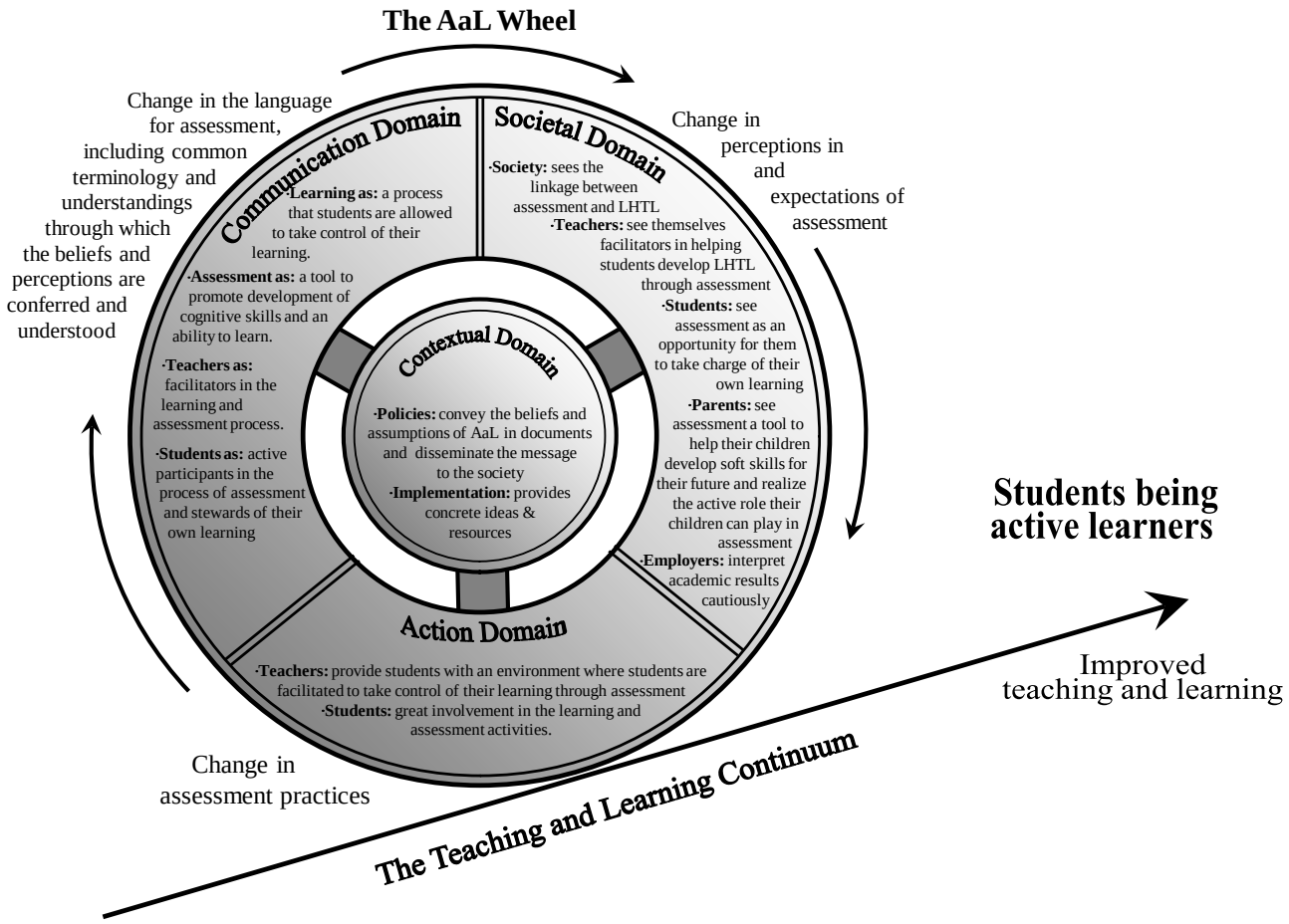


Figure 2. The Assessment as Learning Framework - The AaL Wheel for teaching and learning

Conclusion

Despite the frequent changes and development of school-level curricula and teaching methodologies with an increasing focus on the need for greater student participation in the assessment process and a formative approach to learning, developments with regard to assessment and instruction leaves something to be desired. Many current assessment and instruction practices encourage students to demonstrate current knowledge and to play a passive role in the assessment process, rather than developing critical thinking abilities and being active in their own learning. The Assessment as Learning Framework places strong emphasis on the role of the learner and highlights the use of assessment to increase learners' ability to take control of their own learning. This framework is built upon the combination and integration of the four

domains: Contextual, Societal, Communication and Action Domains. In the Assessment as Learning Framework, the four domains are engaged in a dynamic relationship, with constantly evolving definitions, plans of actions and understanding of social perceptions in various contexts, aiming at helping learners take an active role in their learning so that students can be more able to tackle their challenges in and beyond the classroom.

References

Angyal, A. (1941). *Foundations for a Science of Personality.* Cambridge, MA: Harvaard University Press.

Assessment Reform Group (2006). *The Role of Teachers in the Assessment of Learning.* London: University Institute of Education.

Berry, R. (2011a). Assessment Trends in Hong Kong: seeking to establish formative assessment in an examination culture. *Assessment in Education: Principles, Policy & Practice, 18*(2), 199-211.

Berry, R. (2011b). Assessment reforms around the world. In R. Berry, & B. Adamson (Eds.), *Assessment Reform in Education: Policy and Practice* (pp.89-102). Dordrecht, Netherlands: Springer.

Berry, R. (2010). Teachers' orientations towards selecting assessment strategies. *New Horizons in Education, 58*(1), 96-107.

Berry, R. (2008a). *Assessment for Learning.* Hong Kong: Hong Kong University Press.

Berry, R. (2008b). From theory to practice: Curriculum for autonomous learning. In M. F. Hui, & D. Grossman (Eds.), *Improving Teacher Education through Action Research* (pp.117-136). New York, USA: Routledge

Berry, R. (2006). Activating learners using the learner autonomy approach: An action research on the relevance of teaching to classroom practice. *Curriculum Perspectives, 26* (3), 34-43.

Berry, R. & Adamson, B. (2011). Assessment reform past, present and future. In R. Berry, & B. Adamson (Eds.), *Assessment Reform in Education: Policy and Practice* (pp.3-14). Dordrecht, Netherlands: Springer

Black, P., McCormick, James, M. & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education, 21*(2), 119-132.

Bonk, C. J., & Dennen, V. (2007). Frameworks for design and instruction. In M. G. Moore (Ed.), *Handbook of Distance Education* (2nd Ed.) (pp. 233-246). Mahwah, NJ: Erlbaum.

Boud, D. (2007). Reframing Assessment. In D. Boud & N. Falchikov, *Rethinking Assessment in Higher Education* (pp. 14-25). New York, NY: Routledge.

Bronkhorst, L. H., Meijer, P. C., Koster, B. & Vermunt, J. D. (2011). Fostering meaning-oriented learning and deliberate practice in teacher education. *Teaching and Teacher Education, 27*(7), 1120-1130

Brookhart, S. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education, 8*(2), 153-169.

Carless, D. (2006). Differing perceptions in the feedback process, *Studies in Higher Education, 31*(2), 219-233.

Chappuis, S. and Stiggins, R. (2002). Classroom assessment for learning. *Educational Leadership, 60*(1), 40-43.

Craddock, D. & Mathias, H. (2009). Assessment options in higher education. *Assessment & Evaluation in Higher Education, 34*(2), 127-140.

Crick, R. D. (2007). Learning how to learn: the dynamic assessment of learning power. *The Curriculum Journal, 18*(2), 135-153.

Earl, L. M. (2003). *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning.* Thousand Oaks, Calif.: Corwin Press.

Entwistle, N., & Walker, P. (2000). Strategic awareness within sophisticated conceptions of teaching. *Instructional Science, 28*(5-6), 335-361.

Expert Group on Assessment (2009). *The Report of the Expert Group on Assessment*. Retrieved on 29 August 2011 from: http://publications.dcsf.gov.uk/eOrderingDownload/Expert-Group-Report.pdf

Eun, B. (2011). From learning to development: a sociocultural approach to instruction. *Cambridge Journal of Education, 40*(4), 401-418.

Garratt, D. (2011). Reflections on learning: widening capability and the student experience. *Cambridge Journal of Education, 41*(2), 211-225.

Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown and A. Glasner (Eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 41-54). Buckingham: Open University Press.

Gibbs, G. & Simpson, C. (2004) Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 1*(1), 3-26.

Hargreaves, E. (2011). Teachers' feedback to pupils: 'Like so many bottles thrown out to sea'? In R. Berry, & B. Adamson (Eds.), *Assessment Reform in Education: Policy and Practice* (pp.121-133). Dordrecht, Netherlands: Springer.

James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner, *Assessment and Learning* (pp. 47.-60). London: Sage.

James, M., Black, P., McCormick, R. & Pedder, D. (2007). Promoting learning how to learn through assessment for learning. In M. James, and R. McCormick et al. (Eds.), *Improving Learning How to Learn: Classroom, Schools, and Networks* (pp.1-29). NY: Routledge.

Klenowski, V. (2009). Assessment for learning revisited: an Asia-Pacific perspective. *Assessment in Education: Principles, Policy and Practice*, *16*(3), 263-268.

Klenowski, V. (2007). Book Review: Assessment and learning. *Assessment in Education: Principles, Policy & Practice*, *14*(2), 269-274.

Kember, D. (2004). Misconceptions about the learning approaches, motivation and study practices of Asian students, In M. Tight, *The Routlegefalmer Reader in Higher Education* (pp.37-52). London: RoutlegeFalmer.

Knight, P. T. (2003). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education, 27*(3), 275-287.

Marsh, C. (2009). *Key Concepts for Understanding Curriculum (3rd Ed.).* London & New York: RoutledgeFalmer.

Munns, G. (2005). School as a cubbyhouse: tension between intent and practice in classroom curriculum. *Curriculum Perspectives*, *25*(1), 1-12.

Murphy, R. (2006). Evaluating new priorities for assessment in higher education. In *Innovative Assessment in Higher Education* (pp. 37-47). New York, NY: Routledge.

Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Race, P. (1999). Why assess innovatively? In S. Brown & A. Glasner, *Assessment Matters in Higher Education* (pp. 57-70). Philadelphia, PA: Open University Press.

Rose, R. (2002). Teaching as a “research-based profession”: Encouraging practitioner research in special education. *British Journal of Special Education 29*(1), 44-48.

Saddler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education, 30*(2),175-194

Shipman, D., Aloi, S. L. & Jones, E. A. (2003). Addressing key challenges in higher education assessment. *Journal of General Education*, *52*(4), 335-346.

Stansfield, M. (2001). *Introduction to Paradigms: Overview, Definitions, Categories, Basics, Optimizing Paradigms & Paradigm Engines*. Victoria, B.C: Trafford.

Vygotsky, L. S. (1978). *Mind and Society: the Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

Watkins, C., Carnell, E. and Lodge, C. (2007). *Effective Learning in Classrooms*. London: Paul Chapman.

Weaver, R. G., & Farrell, J. D. (1997). *Managers as Facilitators: A Practical Guide to Getting Work Done in a Challenging Workplace*. San Francisco, CA: Berrett-Koehler Publishers.

Young, S. (2008). Theoretical frameworks and models of learning: tools for developing conceptions of teaching and learning. *International Journal for Academic Development, 13*(1), 41-49.

Author’s e-mail:　rsyberry@gmail.com

教育的成功與成功教育——課堂教學的反思

莊紹文
荔景天主教中學

教育的目標就是啓發學生的個人潛能、培養人文素養、建構全人發展，追求成功。但是，良好的願望並不能保證教育目標的實現，究其原因，可以從多方面及不同角度作出思考。作爲前線的教育工作者並擔任中學校長一職多年，在日常工作所體會到的、所觀察到的及印象最難忘的，可說是課堂內的學與教活動。有效能的課堂教學，乃是促進學生學習、啓發學生個人潛能，及追求教育成功的要訣。本文嘗試以學生的課堂學習活動作爲基礎，探討如何提高課堂效益，有效提升學與教素質及學生的學習成效。

教育的改革，由傳統以教師爲中心的「教與學」到以學生爲中心，強調營造多元情境，體驗不同學習經歷的「學與教」活動，目的都是強調學生的創新精神與實踐能力的培養。「跑出課室」以激發學生學習活力，可以說是成爲了當前教育改革的新路向。然而，在「跑出課室」後，學生們仍要「返回課室」，將學習經歷整合及梳理，把所獲知識轉化爲自己的真知，從而鞏固其學習。 事實上，學生們主要學習的場所總離不開「教室」及老師們的課堂教學。 因此，課堂教學可以說是教學活動的主要組織形式，是培養學生的主要陣地。

課堂教學作爲學與教在動態中的組合，其中包括了多種複雜的因素。在教學的實踐中，課堂教學主要目的在於充分發揮教師的主導作用和學生的主體作用，從而達到提高課堂效益、提升學生們的學習成效，更重要的是培養學會學習的終身學習態度。正如我國著名教育學家葉聖陶先生[1]認爲：「教師的教就是爲了不教，起主導作用，學生學會自學，這是教學成功的最高境界。」因爲，學生進入這樣一種境界，能夠自己去探索，自己

1 見維基百科 葉聖陶介紹：
http://zh.wikipedia.org/zh-tw/%E5%8F%B6%E5%9C%A3%E9%99%B6#.E2.80.9C.E6.95.99.E6.98.AF.E4.B8.BA.E4.BA.86.E4.B8.8D.E6.95.99.E2.80.9D

去辨析，自己去歷練，從而獲得正確的知識和熟練的能力，豈不是就不需要教了嗎？ 而學生所以要學要練，就爲要進入這樣的境界。葉聖陶先生更認爲：「教育學說雖然深奧萬端，也可以用一句包括，就是要學生『生』。」因爲「讀書忌死讀，死讀鑽牛角，復孜孜，書我不相屬。活讀運心智，不爲書奴僕，泥沙悉淘汰，所取惟珠玉。」然而，究竟甚麼樣的課堂教學，才是最佳的課堂教學結構呢？ 相信，沒有可能找到一種到處都可以套用的統一模式。由於教學對象、教學環境的不同，在一種場合是最佳的，而在另一種場合可能並不能發揮其相應的果效。因爲在教學的過程及其產生的教學效應，是受多種複雜變量因素制約和影響的。然而，不管是甚麼情況，課堂教學仍然有一些共同的規律可循。

正如孔子在《論語》中所言：「學而時習之，不亦悅乎？」意思是，在學了之後及時、經常地進行溫習和實習，不是一件很愉快的事情嗎？ 很明顯，按照孔子和其他中國古代教育家的看法，學習這一複合名詞中的「學」，就是聞、見與模仿，是獲得訊息、技能的過程。主要是指接受感官訊息（圖像訊息、聲音訊息及觸覺味覺等等訊息）與書本知識，有時還包括思想的含義。「學」是自學或透過導師的教授， 而「習」所指的就是鞏固知識、技能的行爲。一般來說有三種含義：溫習、實習、練習。由此看來，「學」偏重於思想意識的理論領域，「習」則偏重於行動實習的實踐方面。 因此，學習就是獲得知識，形成技能，獲得適應環境，改變環境的能力及過程[2]，實質上就是學、思、習、行的總稱。

「學」是指知識和經驗的累積，「習」是指知識和經驗的實踐。因此，學屬知，習屬行。中國思想家王陽明所提倡的「知行合一」， 對課堂教學結構的規律，能否給我們提供一些啓示呢 ？ 而這些規律，是否有助提升學與教的素質？

中國現代教育家段力佩[3]在上海育才中學的銳意改革事蹟，是全國教育界所矚目的。他對提升學與教素質，主張教師要充分發揮和提高「四十五分鐘」(即每一教節)的效能。因爲教師引導學生把已知的科學真理轉化爲自

2 Holt, John (1983). *How children learn*. UK: Penguin Books.

3 中国教育家的故事。《段力佩的大胆尝试》。
http://www.mj.org.cn/mjfc/hyfc/200410/t20041029_77930.htm

己的真知，把知識轉化爲能力的教學過程絕大部分是在課堂上完成，也就是在「四十五分鐘」內完成。個人認爲，要提高學與教素質，加強學生的能力培養，教師們必須充分提高課堂的效能，至於是否應以「四十五分鐘」作爲每一課堂學習活動的框架，相信仍有待深究。 然而，教師對每一課堂教學結構的規律、理念，甚至課堂內的教學方法等等，實在宜多花心血、多想辦法、多作研究，採取多種形式的教學方法，切忌使學生感到枯燥乏味，失去學習動力。

源於研究「弱勢學生」的上海市閘北八中的成功教育，在「三相信」(即堅信每個學生都具有成功的願望、都有成功的潛能、都可以獲得多方面的成功)的基礎上，通過「三要素」(即積極的期望、提供成功的機會和鼓勵性評價)，經過「三階段」(即幫助成功、嘗試成功、自主成功)，抓住「一個作用點」(即自我概念)，形成學生們學習的內在動力機制，對促進學生的自主學習，幫助學生正確認識和改造失敗，最終形成自主學習或提升學生學會學習的能力[4]。

成功教育，作爲一種教育思想，亦可以作爲一種課堂教學方法。在課堂教學過程中的基本要求包括：

(1) 制訂恰當的教學目標和要求;

(2) 按照教學目標和要求設計並組織實施分層教學，爲每一學生創設成功機會;

(3) 創設情景，讓學生更多、更主動地參與學習活動;

(4) 注重學與教的多向交流，及時回饋、及時矯正;

(5) 實施鼓勵性評價，以激勵學生，使學生獲得成功感;

(6) 指導學生掌握正確的學習方法，培養學生具有良好的學習習慣[5]。

在構建成功教學模式方面，以「幫助成功」，即「低、小、多、快」作爲教學策略。 「低」即「低起點」。教師可於教學前預先透過問卷調查、

4 劉京海 (2005)。成功教育。福建：福建教育出版社。

5 王慶茂、余拱焰、金雍城 (2001)。成功教育研究。山東：山東教育出版社。

談話、診斷性測試、預習檢查、口頭或書面提問等方式，了解學生的水平，摸清學生相關知識、基礎、能力和心理準備或狀況，以確定適當的教學起步點。只要把教學起點放在學生努力一下就可以達到水平之上，才可使學生對所掌握的新舊知識產生聯繫。 事實上，「工欲善其事，必先利其器」。要提高學生的課堂學習效能，關鍵在於精心備課。一般來說，教師對課前的備課皆能盡心盡力，準備充足。普遍來說都會深入鑽研課程大綱和教材，分析教材中的重點和難點，考慮講授的重點和方法，訂定教學目標。然而，卻欠缺「摸底」的部份，研究學生的水準和差異，考慮講授的重點和方法，從而選擇不同的教學方式，根據學生的實際情況擬定計劃，爲學生訂立學習目標，琢磨如何引起學生的注意、興趣和想像，貫徹教材的科學性和思想性，調動學生的思維活動，循序漸進，因材施教，設計不同層次的提問及內容，作出適時的小結及總結，以鞏固學生記憶，從而發展「優生」的學習能力，適應「弱勢學生」的實際水準，力求達到傳授知識、發展能力及促進自學能力的教學目的。顏明仁 (2009) 在《促進學習的評核及考試改革》[6] 一文中，提出了促進學習評核的特徵，當中包括了：(1) 和學生分享學習目標及 (2) 協助學生知悉和認定達到目標的標準。這正好就是「摸底」的方法，以確定適當的教學起點，亦可「幫助學生成功」，知道只要努力一下，就可以達到水平之上，更可以使學生對所掌握的新舊知識產生聯繫。

構建成功教學模式的「小」，即「小步子」。乃根據學生的實際能力，確定能達到的實際進度。教學的步子要小，把教學內容首先進行整理，把學習內容作適合的「拆解」，按由易到難，由簡到繁的原則「分解」成合理的層次，然後分層漸進、按部就班，將學生們有機會產生挫折的頻率降到最低的程度，使學生層層有進展，處處有成功，處於積極學習的狀態，感到自己有能力進行學習，從而不斷增強學習的自信心和動機。然而，在課堂內，教師的課堂語言也要精練，一般來說，老師爲能確保所教的知識能真正「入腦」，因此習慣了將簡單的要點不斷重複，卻忘記了採用不同的方式，如恰當的提問技巧，提出問題，從而就既定的學習內容作出小結，以確定學生們的學習成效，亦可作爲鞏固學習的一種方式。除此之外，教師的講話亦宜生動，多帶啓發性，忌平鋪直敘塡鴨式的結論性敘述，宜透過

6 顏明仁 (2009)。促進學習的評核及考試改革。輯於吳迅榮、黃炳文主編：廿一世紀的學校領導：持續與創新，頁 97-112，教育領導與變革叢書。香港：學術專業圖書中心。

促進學習評估的方式，爲學生的學習進行階段式的評核，從而總結所學，確定學生們的學習成效。另一方面，在進行課題的分析時要精闢，忌嚕囌；解題要簡潔，忌繁瑣；一題多解，「舉一反三」，鼓勵學生多角度思考。在時間的安排上，教師宜讓學生可按部就班，有充分獨立思考的餘地和時間，忌「一桿子」講到底。「滿堂灌」是教學工作的大敵，原因是個別老師擔心課程緊迫，只要「教完」課程，甚至學校或科本的要求，已「交足貨」了，完全未能掌握學生們真正來說學了多少？學習的成效如何？

改變「滿堂灌」，教師大段講解的傾向，使師生活動交替進行，正好是構建成功教學模式的「多」，即「多活動」或「多練習」。「多」可以說是爲針對學習能力稍遜學生在專注力薄弱、記憶「容量小」、概括能力較爲弱的特點。因爲「多活動」或「多練習」不僅可以調節學生的專注力，更重要的是學生可以大量參與學習活動，自我表現的機會多了，能力的發展也通過逐步積累而得以實現。在良性的循環下，又大大促進了學生各方面的發展。然而，從日常觀課所見，有些教師唯恐學生不明白，在短短時間的課堂中，經常「滿堂灌」，學生完全沒有或很少有獨立思維的餘地。教師聲嘶力竭，講得滿頭大汗，但效果卻往往相反。據本人的經驗，整體上時間安排可以「四六開」至「六四開」之間，即在一般的講授課，教師講授時間控制在百分之六十至四十左右，學生親自動腦、動口、動手的時間在百分之四十至六十左右。教師須同時鼓勵並歡迎學生隨時提出問題來，即使是較爲幼稚的問題也要加以歡迎和鼓勵，因爲正是這種「快速反饋」的策略才有利於促進教學互動、提高學生的學習主動性和學習興趣，以及活躍的課堂氣氛。另一方面，師生可以從學生的表現及進度下檢視及反省 (顏明仁, 2009)，有利於提升教學效果及有利於培養學生的能力。在日常的觀課所見，教師努力備課，對教學內容及教學流程，可以說是充份掌握。然而卻欠缺「靈活性」，每遇到學生打斷自己講課而嚴加訓斥，有時甚至有「被挑戰」的感覺，將課堂氣氛成爲「一言堂」變得沉悶。事實上，學生的學習氛圍對促進學習的成效至爲重要。因爲活躍的課堂氣氛，營造「生生互動」的課堂氛圍，往往都是教師引導兩個「轉化」的最佳時刻。

至於在構建成功教學模式的「快」，即「快回饋」。在每一層次的教學過程中，既有教師的「講」，也有學生的「練」，還有教師的「查」。這種快速的反應，既可提供回饋使學生知道如何改善、引領學生安排進一步做事 (顏明仁, 2009)，亦可以把學生取得的進步變成有形的事實，使之受到鼓勵，

讓學生有信心可以改善，甚至樂於接受下一個任務，亦可以及時發現學生存在的問題，及時矯正乃至調節教學的進度，從而有效地提高課堂教學的效益 [7]。Marzano (2007)[8]在進行給學生回饋的研究分析中指出，其效應值 (Effect size) 乃介於 0.26 – 1.47， 肯定了「快回饋」在教學過程中的作用，而讓學生透過自評和同儕互評，利用評估數據改進學習和教學，均是現時的教師需要加以利用的。 因此，教師宜適當地增加課堂上的練習。除此之外，教師所提出的問題、例題或習題，要能促使學生動腦筋、多思考，而忌一味追求數量，機械重複。另外，教師所有要講解的例題或習題，不管是新課還是複習課，都應堅持在講解前讓學生有親自動腦、動口、動手的機會和時間。在此基礎上才能達到「講一題、懂一類、識一批」的效果，讓學生取得的進步變成有形的事實，使學生有信心可以改善。一般來說，教師的「通病」是「心中已有既定的答案」，在提出問題時未讓學生有足夠思維的機會和時間，只在不斷地「追求」心中既定的答案，只要遇到「知音人」能「答中」心中既定的答案，便「點到即止」，沒有追問其答案的原因。而有些教師甚至以「自問自答」的形式，先講解題目(有些學生往往喜歡教師如此，懶得動腦筋)，便提供即時的答案。事實上，這點對學生的能力培養也是至關重要的。

成功教育除了構建成功教學模式，以「幫助成功」，實踐「低、小、多、快」作爲教學策略，在課堂教學過程中的基本要求也包括了指導學生掌握正確的學習方法，培養學生具有良好的學習習慣。因此，教師須培養學生在課內外獨立自學的能力和習慣。 閱讀，透過對課題預習的方式，是自學能力的基本條件。然而，有相當多的學生缺乏閱讀的習慣或能力，這是一個要花大功夫的工作。學生出於多方面的原因，如不習慣、不認真、不重視、有依賴性、看電視和玩電子遊戲的吸引力大等等，對看書興趣不大，閱讀能力較低。因此需要教師諄諄引導，說明看書閱讀的目的和益處，指導、講解一些學習閱讀的方法，同時採取相應措施，如在預習課本後的提問，分享閱讀心得等等。一般來說，在課堂時間的規劃方面，教師盡可能

7 王慶茂、余拱焰、金雍城 (2001)。成功教育研究。山東：山東教育出版社。

8 Marzano, R.J. (2007). *The art and services of teaching : A comprehensive framework for effective instruction.* Alexandria, Va.: Association for Supervision and Curriculum Development.

在每節課中安排一定的時間讓學生認真閱讀 (包括數學堂)，將書本上的課題進行閱讀，再配以適當的提問，確定學生對課題的掌握，從而鞏固所學。如教師對這問題予以足夠的重視，同時對其難度也作充分的準備，堅持兩個月左右的時間，學生便能逐漸養成習慣，而且閱讀的效果也會慢慢顯示出來，對教師的講課有意料之外的幫助。

以上各點心得，都只是本人多年教學及觀課的一些經驗體會而已，古今中外「教無定法，各師各法」。個人認爲不能厚「一法」而薄其他，更不宜強行推行某一「定法」(這本身就違反了教學規律)。只要能充分發揮教師的專業精神，體驗「知行合一」，通過實踐，証明能最大限度地提升學與教效能的，就是一種好的教學法。

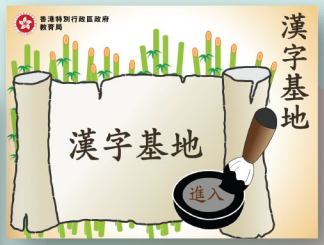
最後，必須強調的是：教師若要在課堂的框架內達到預期效果，則需要「吃透兩頭」，一頭是教材，一頭是學生。所謂「知己知彼，百戰不殆」。當然，「要給學生一杯水，自己需有一桶水」。因此教師工作主要的「功夫」是在課外，否則一切都只是「紙上談兵」，中看不中用了。

作者電郵： ambrosechong@gmail.com

WLTS 促進學習評估資源庫
網上學與教支援

wlts.edb.hkedcity.net

- ✓ 針對學習難點發展的互動課業
- ✓ 實踐「促進學習的評估」的理念



教學簡報 · 互動課業 · 網上遊戲

研究報告 · 課堂工作紙 · 教學計劃

閱讀理解評估素養——
閱讀篇章特徵在擬題中的位置

林偉業
香港大學教育學院

一、閱讀理解評估與閱讀認知層次

閱讀，首先是「理解」(comprehension)，所謂理解，就是「明白」(understanding)，不嫌簡化的話，即讀者因爲所閱讀過的篇章而多知道一些東西；學習以求取知識，便是最常見的閱讀例子。當然，讀者也可以只爲享受閱讀的樂趣，例如閱讀笑話或小說，對讀者來說，所知的東西其實不重要，不過，通過所知，產生了愉悅感受 (Barthes, 1975)。又或者，讀者根據新知，調整對事物已有的立場和觀點，引發影響真實世界的行動；閱讀評論文章，即爲此類。然而，不論愉悅享受、觀點重整，或者其他閱讀的結果，都離不開「明白」這一閱讀的核心。

無論從常識看，還是以認知科學或學習理論的角度觀之，「明白」至少有表層、深層之別。Gray (1960) 區分「字裡」(reading the lines)、「行間」(reading between the lines)、「言外」(reading beyond the lines) 閱讀三者，是較早區分閱讀層次的見解。

對如何區分閱讀層次最有影響的學說是布魯姆 (Bloom et al., 1956) 的教育目標分類學，這學說全方位地影響教育的各個領域，閱讀亦然。Anderson 等 (2001) 修訂布魯姆的教育目標分類學，遂把閱讀的層次由低而高按序分爲「記憶」、「理解」、「運用」、「分析綜合」、「評價」、「創造」等六個認知過程維度，教師與學界可以據之設置閱讀學習和評估的目標。祝新華 (2005) 提出六層次閱讀認知能力，包括：(一) 複述——認讀原文，抄錄詞句，指出顯性的事實；(二) 解釋——用自己的話語概說詞語、句子的表面意思；(三) 重整——概括篇章內容，辨識內容關係、表達技巧；(四) 伸展——在理解篇章表層意義後，推出隱含意義；(五) 評鑒——評說思想內容，鑒賞語言表達；(六) 創意——找新方法，提新想法，運用所讀資訊解決問題，更是香港教育界十分熟悉的架構。

以上觀點把讀者讀明白篇章的能力區分爲由低至高的層級架構，上一層次依賴下一層次，達到下一層次才能晉身上一層次。另一類閱讀層次學說認爲閱讀能力固然可以分成不同層級，但不必然是線性的順序關係。如 Marton 與 Säljö (Marton & Säljö, 1976, 2005; Säljö, 2005) 區分表層 (surface) 與深層 (deep) 閱讀之別，指出讀者聚焦於篇章內容本身的閱讀是表層閱讀，深層閱讀關注讀者自身的經驗。Mayer (1987) 把「讀明白篇章」看成是一個「努力追求意義」的過程，過程中，讀者須使用三類知識，包括：運用內容知識——運用先前已有知識、運用策略知識——運用文章結構、做推論，以及運用後設認知知識——監控理解的過程、自我提問與檢核，並依不同目標調整閱讀方法。羅燕琴 (2011) 考察香港的評估實踐，提出閱讀能力先分爲表層文意基本理解、深層意義篇章理解，以及個人新知識及看法建構三層，表層文意理解包括字詞解碼能力與直接推論能力二者，深層意義理解包括分析綜合文意關係能力與推論言外之意能力，個人建構則包括應用所讀解決問題的能力、結合所知評鑒所讀內容與手法的能力，以及拓展所讀創新感悟的能力。

無論是前者的序列取向，還是後者的板塊取向，均以描述讀者讀明白篇章的能力爲目的。不過，除此之外，從現象學的角度看，凡思維活動，能思 (noesis) 必對應於所思 (noema) (Sokolowski, 2004)，因此，作爲思維活動，「明白」也必伴隨以「明白的對象」，在閱讀的過程中，這個對象離不開閱讀的材料，即篇章本身。所以，閱讀能力既是讀者讀明白篇章的能力，則一端涉及讀者閱讀認知能力的層次，另一端則涉篇章的特徵，二者結合，才能描述閱讀過程與結果的全貌，以作爲閱讀評估的基礎。

本文以下部份以 2011 年第三屆「全球學生閱讀能力進展研究」(Progress in International Reading Literacy Study, 以下簡稱 PIRLS) 其中一篇已公布的測試篇章——《飛吧！鷹，飛吧！—— 一個非洲的故事》(Fly, Eagle, Fly!) ——及所設問題爲例，探討如何在閱讀理解測試擬題時，除了思考讀者閱讀認知層次外，同時考慮閱讀理解測試篇章的篇章特徵 (textual characteristic)。

二、「全球學生閱讀能力進展研究」2001 閱讀測試篇章——《飛吧！鷹，飛吧！—— 一個非洲的故事》

《飛吧！鷹，飛吧！—— 一個非洲的故事》一文及其閱讀理解設題載於《PIRLS 2011 國際閱讀成績結果》(The PIRLS 2011 International Results in Reading) <http://timssandpirls.bc.edu/pirls2011/international-results-pirls.html> 的《附件 H：樣本篇章、問題與評分建議》(Appendix H: Sample Passages, Questions, and Scoring Guides) <http://timssandpirls.bc.edu/pirls2011/downloads/P11_IR_AppendixH.pdf> 中。《飛吧！鷹，飛吧！—— 一個非洲的故事》本是一個非洲的民間故事，敘述一位農夫在尋找因風暴走失的小牛時，無意中在峭壁的岩石後發現一隻剛出生的小鷹。農夫救了小鷹回家，讓牠與雞一起生活，要把小鷹訓練成一隻雞王，漸漸，這頭鷹的行爲像一隻雞。農夫的一位朋友看不過眼，要喚醒鷹的本性，朋友把鷹高舉頭上，對牠說：「你不是雞，是鷹，你不屬於大地，屬於天空，飛吧，鷹，飛吧。」可是鷹卻跳回地上與其他雞爭食。朋友不甘心，第二天，天未亮便把鷹帶到高山峭壁的岩石上，對牠說：「你不是雞，是鷹，你不屬於大地，屬於天空，飛吧，鷹，飛吧。」當鷹看到第一道陽光從山後射出，照遍山上山下，鷹張開翅膀，感受羽毛上的溫暖，還有一股把牠往上升的氣流，牠向前一傾，飛了起來，消失在陽光之中。

「全球學生閱讀能力進展研究」爲這篇測試擬設了 12 道問題，見第 3.3 節表一。一如其他閱讀評估，「全球學生閱讀能力進展研究」同樣以讀者閱讀認知能力層次爲閱讀評估的架構，包括：尋找文章表層明顯訊息、簡單直接推論、綜合並解釋文意，以及評價文章內容與表達方式等四者 (謝錫金等, 2005; Mullis, Martin, Kennedy, Trong, & Sainsbury, 2009)。12 道問題中，尋找文章表層明顯訊息的題目有 3 道、簡單直接推論的題目有 4 道、綜合並解釋文意的題目有 4 道，評價文章內容與表達方式的題目有 1 道。此外，12 道題目中，7 道爲多項選擇題，俱爲 1 分題，其餘 5 道問題爲短答題，除其中一題爲 1 分題外，其他短答題都是 2 分題。層次愈高，分值一般愈多，並且愈多以短答題方式擬題。這種考慮閱讀層次的分布，並按層次擬定分值與題型，已是閱讀理解評估素養 (assessment literacy) 的重要環節。

不過，正如第一節所述，這些分布於不同閱讀認知層次的問題，必與篇章有關，問題是：與篇章哪些特徵有關？

三、閱讀理解評估與與閱讀篇章特徵

設計閱讀認知的問題要與篇章特徵有關，設題者必須先爲所測考的篇章作篇章分析 (textual analysis)。《飛吧！鷹，飛吧！—— 一個非洲的故事》是一篇故事，故事由兩維度構成，一是故事的情節，另一是故事的話語 (Chatman, 1980)，因此，這故事的分析如下：

3.1 故事情節

記敘故事的情節一般由四個環節構成：背景——問題——解決方法——結果。簡言之，敘事以「問題——解決」(problem-solving) 爲情節的重心，讀者理解敘事時，須辨認出故事「人物」(character) 在所置身的「背景」(setting) 中遇到哪些重大「問題」(problem)，以及因而所要達到的「目的」(goal)，故事遂由人物如何自力或在他力的幫助下，爲「解決」(solution) 問題，做出一系列行動，作爲「方法」(plan)，在逐步開展的過程中，並以問題最終能否得以解決、或目的是否達到爲故事的「結果」(resolution) (謝錫金等, 2005; Ng & Lam, 2009)。

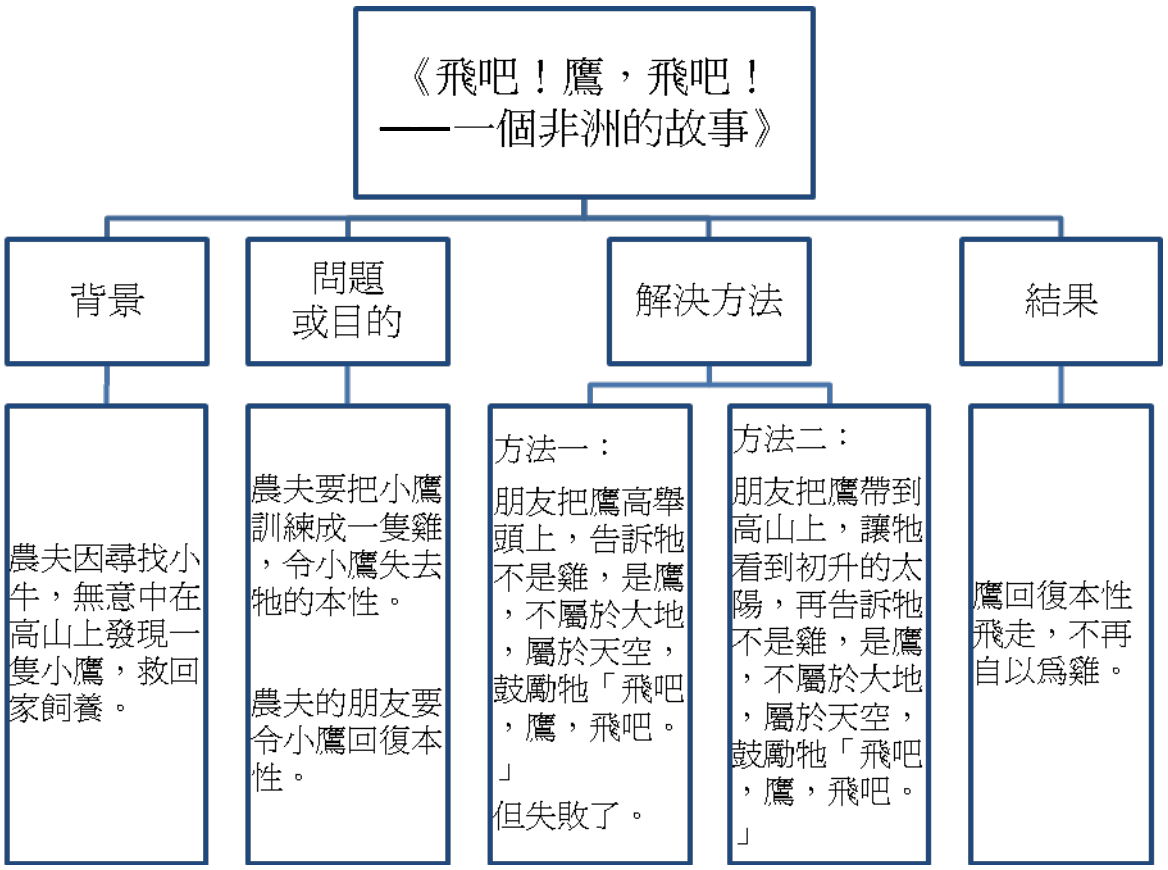
一、背景環節先介紹故事發生的時空和人物（通常稱爲時間、地點、人物），讓讀者能藉此建立與真實生活對應的情境 (Zwaan et al., 1998)。然而，背景的作用其實遠不止於交代時、地、人，更重要者，是爲故事的問題〔或目的〕設定條件，讓問題〔或目的〕成爲可能。

二、問題〔或目的〕環節是故事的核心，故事主要人物（主角）或者希望達成某個目的，或者不滿當下的現況，希望改變當下狀況中存在的問題。問題〔或目的〕是故事所以成爲故事的核心，沒有這一部分，故事常常令人難以理解，或令人感到不成故事，或令人覺得故事沒有意義 (Labov & Waletzky, 1967; Polanyi, 1985)。理解故事，首先就是能瞭解主角在故事中所要達成的目的或所要解決的問題，作爲故事發展的動機 (Stein & Glenn, 1979; Stein & Albro, 1997)。

三、解決方法環節是故事情節的發展，主角想方設法要達成目的或解決問題，過程中主角或獲別人幫助，或遭他人阻撓；或須以試誤的方式找出有效解決問題的方法；或認識到問題〔或目的〕的真正癥結、或提煉出新的問題〔或目的〕、或由一個問題〔或目的〕產生出更多問題〔或目的〕，不一而足，由是衍生出可以十分複雜、奇譎的情節結構。

四、結果環節交代主角能否成功達成或解決問題〔或目的〕，以收束故事。故事的發展不外兩個方向，要麼主角終於達成他希望獲致的目的、成功解決遇到的問題，要麼失敗，甚至造成更大的問題；成與敗，是讀者辨認故事完結與否的標記。

按以上敘事結構分析《飛吧！鷹，飛吧！——一個非洲的故事》的情節，如下：



故事的主要角色是那隻鷹，故事重心——即問題——在於救回小鷹的農夫要把小鷹變成一隻雞，改變牠的本性。故事主角所遇到的這個問題，不是由鷹自己發現的——牠自己也受蒙蔽，而是由農夫朋友這位輔助角色的觀點帶出來的，故事的發展因此在於農夫朋友如何幫助鷹回復本性。故事情節發展略有曲折處，農夫朋友所使用的第一個方法，沒有讓鷹脫離蒙蔽牠的環境，因此沒有效果；第二個方法遂讓鷹回到牠出生的地方，在牠所屬的環境中，以太陽初升既壯麗又雄偉的景象觸發鷹的內在本性，終於成功了，這小曲折一方面令故事不會太單調，另一方面，兩個方法一成一

敗的對照也能更突出主題。所以，這是一個關於主角覺醒，恢復自然本性的故事，故事提出物種有某種內在自然本性的觀點，並把改變本性視爲對物種的傷害，而且把回復自然本性與崇高 (sublime) 的體驗聯繫起來。

3.2 故事話語

故事話語與故事的語言與意義表達有關，爲方便討論，這裡只涉故事篇章的詞彙、人物角色行動，與行動意義三者。故事篇章以語言爲媒介，其中詞彙是表意的重要手段；然後是人物角色的行動，乃故事事件的構成元素；不過，不論詞彙還是人物角色行動，都爲了表達故事的意義。

3.3 故事「情節－話語」的設題藍圖

按以上「情節－話語」的故事敘事維度分析，第三屆「全球學生閱讀能力進展研究」2001 爲《飛吧！鷹，飛吧！—— 一個非洲的故事》一文所設的閱讀理解問題可分析如下：

表 1：第三屆「全球學生閱讀能力進展研究」2001
《飛吧！鷹，飛吧！—— 一個非洲的故事》所設閱讀理解問題

	詞彙	行動	行動的意義
背景		(1) 故事一開始，農夫出去找什麼？ (2) 農夫在哪裏找到小鷹？	(3) 從故事中哪裏可以看出，農夫對小鷹很細心？
問題		(4) 農夫把小鷹帶回家之後，農夫怎樣對待牠？ (5) 朋友第一次來探訪的時候，小鷹的行爲像一隻雞。舉兩個例子說明這個情形。	
方法一		(6) 農夫的朋友第一次遇到那隻鷹時，朋友做了什麼，好讓鷹飛起？	(7) 農夫的朋友對鷹說：「你不屬於大地，你屬於天空。」請解釋他這句話的意思。 (8) 農夫的朋友第一次來探訪時，爲什麼農夫哈哈大笑？

方法二	(10) 請從故事中找出幾個詞語，形容剛剛天亮的時候，天空十分美麗，把這些詞語寫在下面。		(9) 爲什麼農夫的朋友要帶鷹到高山上，讓牠飛起來？請舉出兩個理由。 (11) 爲什麼太陽升起，在故事裏很重要？
結果			
全篇			(12) 你可以從農夫的朋友所做的事情，了解他是怎麼樣的人。試描述一下，朋友是怎樣的人，並舉出一件他做過的事情說明。

從上表可見，我們可以得到以下設題的特點：

1. 故事話語的三個層次中，詞彙的問題與佔分最少，人物角色與行動意義題目與佔分相若。

2. 設題涵蓋故事的全幅情節。這裡有兩點說明：一、上表「結果」一列好像沒有提問，然而，第 11 題「爲什麼太陽升起，在故事裏很重要？」雖然屬方法二的題目，但其實與結果有關係；二、最後第 12 道是題目「你可以從農夫的朋友所做的事情，了解他是怎麼樣的人。試描述一下，朋友是怎樣的人，並舉出一件他做過的事情說明」涉及全篇，不限於情節其中一環節。

3. 設題側重於故事重要情節，所謂重要情節，一般與故事中的問題和方法兩環節有關；然而，有些故事在結果環節發生出人意料之外的事情，這類故事的重點當然在結果環節了，但《飛吧！鷹，飛吧！—— 一個非洲的故事》不屬這類故事。《飛吧！鷹，飛吧！—— 一個非洲的故事》的設題，一方面以涉及問題和方法兩環節的題目最多，12 題中佔 9 題；此外，涉及問題和方法兩環節的題目，閱讀認知層次也較高，其中特別以方法環節（包括方法一與方法二）的題目與佔分都較問題環節更多。由以上討論可見，閱讀理解評估題目的設計，首先與閱讀評估所用的篇章有關，據此發展出不同閱讀認知層次的評估題目。如果不先以篇章特徵出發，閱讀認知層次往往變得孤掌難鳴，設題時候，容易有不知從何擬起之嘆。

4. 結語

本文以第三屆「全球學生閱讀能力進展研究」2001 測試篇章《飛吧！鷹，飛吧！—— 一個非洲的故事》及其設題爲例，討論了閱讀理解評估設題原則：閱讀理解評估設題涉及兩方面，一是讀者閱讀認知能力的層次，另一是閱讀篇章關鍵特徵，對香港語文教育界來說，前者比較認識，但後者則較少提及。本文希望提出，讀者閱讀認知能力層次與閱讀篇章關鍵特徵二者結合，方能較全面了解閱讀理解評估設題的整體面貌，二者都是閱讀評估素養的必要元素。

參考用書

Anderson, W., & Krathwohl, D. R. (Eds.) (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's educational objectives.* NY: Longman.

Barthes, R. (1975). *The pleasure of the text* (R. Miller, Trans.). New York: Hill & Wang.

Bloom, B. S., Engelahar, M. D., Frust, E. J., Hill, W. H. & Krathwohl, D. R. (1956). *Taxonomy of Educational Objective, Handbook 1: Cognitive Domain.* N.Y.: David McKay.

Chatman, S. (1980). *Story and discourse: Narrative structure in fiction and film*. Ithaca, NY: Cornell University Press.

Gray, W. S. (1960). The major aspects of reading. In J. Robinson (Ed.), *Sequential development of reading abilities* (Vol. 90, pp. 8–24). Chicago: Chicago University Press.

Labov, W., & Waletzky, J. (1967). Narrative analysis. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12-44). Seattle: University of Washington Press.

Marton, F. & Säljö, R. (1976). On qualitative differences in learning: I – Outcome and process. *British Journal of Educational Psychology*, 46: 4–11.

Marton, F. & Säljö, R. (2005). Approaches to learning. In F. Marton, D. Hounsell and N. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education*. 3rd (Internet) edition. Edinburgh: University of Edinburgh, Centre for Teaching, Learning and Assessment, pp. 39-58.
<http://www.ed.ac.uk/schools-departments/institute-academic-development/learning-teaching/staff/advice/researching/publications/experience-of-learning >

Mayer, R. E. (1987). *Educational psychology: a cognitive approach.* New York: HarperCollins.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K. L., and Sainsbury, M. (2009). *PIRLS 2011 assessment framework.* Boston: TIMSS & PIRLS International Study Center.

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). PIRLS International Results in Reading. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. <http://timssandpirls.bc.edu/pirls2011/international-results-pirls.html>

Ng, D. F. P., and Lam, J. W. I. (2009). Teaching Cantonese opera script: Story schema approach. In N. H. Salas and D. D. Peyton (Eds.), *Reading: Assessment, comprehension and teaching* (pp. 199-234). New York: Nova Science Publishers.

Polanyi, L. (1985). *Telling the American story: A structural and cultural analysis of conversational storytelling.* Norwood, N.J.: Ablex.

Säljö, R. (2005). Skill in learning and organising knowledge. In F. Marton, D. Hounsell and N. Entwistle, (Eds.), *The experience of learning: Implications for teaching and studying in higher education.* 3rd (Internet) edition. Edinburgh: University of Edinburgh, Centre for Teaching, Learning and Assessment. pp. 59-71. <http://www.ed.ac.uk/schools-departments/institute-academic-development/learning-teaching/staff/advice/researching/publications/experience-of-learning >

Sokolowski, R. (2004). *Introduction to phenomenology*. Cambridge: Cambridge University Press.

Stein, N. L., & Albro, E. R. (1997). Building complexity and coherence: Children's use of goal-structured knowledge in telling good stories. In M. Bamberg (Ed.), *Learning how to narrate: New directions in child development* (pp. 5-44). Mahwah, NJ: LEA.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (Vol. 2, pp. 53-120). Hillsdale. N.J.: Lawrence Erlbaum Associates.

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading*, 2(3), 199-220.

祝新華 (2005)。閱讀認知能力層次——測試題型系統概要。
< http://www.hkedcity.net/article/iworld_sc_feature_courses/090220-002/TE_read2_02.phtml>

謝錫金、林偉業、林裕康、羅嘉怡 (2005)。兒童閱讀能力進展：香港與國際比較。香港：香港大學出版社。

羅燕琴 (2011)。閱讀評估與學習。載於岑紹基、羅燕琴、林偉業、鍾嶺崇編，中國語文評估理論與實踐 (頁 21-66)。香港：香港大學出版社。

作者電郵： 林偉業 jwilam@hku.hk

評估量表對促進學習的作用

梁佩雲
香港教育學院

有效運用多元化的評估，以促進學生學習（教育統籌委員會，2000），相信香港的教師都耳熟能詳。「以評估促進學習」的主張，源於學校、社會對考試成績過份重視，甚至罔顧學生的學習需要。評估能否發揮促進學習的效能，本文認爲取決於學生在評估過程中的學習質量。要言之，通過評估，學生能否逐步釐清學習重點，調整學習策略，最終學會學習自主。

評估是教學不可或缺的元素，沒有評估，無法得知學習的情況和效能。評估的目的，無非是確定學習困難，適時反饋，從而改善學習成效。爲了有效發揮評估的積極作用而儘量減少評估帶來的學習壓力和負面影響，學者曾主張評估形式多元化（李坤崇，1999），評估任務由多方分擔（吳毓瑩，2003），而評估更可以與課堂教學結合（Black & Wiliam, 1998）。

按 Wiliam（2011）的回顧，西方社會曾大力推動的「促進學習的評估」（Assessment for Learning, AfL）意念，最早見於 Harry Black 在 1986 年出版的同名著作，其後經歷不同教育制度的探索，衍生出紛紜的評估理論和實踐模式，討論至今仍然持續。由英國的「評估改革小組」（Assessment Reform Group, 2002）提出的主張，「以評估作爲蒐集學習顯證的手段，讓教師和學生都能認清學生在學習歷程的位置，從而檢討教與學的方法，提升學習效能」，可謂 AfL 的權威界說。近年，多國評估研究學者推陳出新，又聯合制定出一項「第二代」的 AfL 定義：

> AfL 是學生、教師及同儕日常實踐的一部分，旨在從對話、展示和觀察中，尋找、反思和回應有關資訊，以促進持續學習。（Klenowski, 2009, p.264）

新定義不但列明評估工作並不限於由教師執行，還舉出靈活的評估形式和所需的配合行動，同時強調評估的宗旨——促進學習。

緊貼世界教育潮流，香港教育當局曾明確建議變革學校的評估安排，加強採用多元化的評估促進學習，又籲請教師讓學生評估同學的課業（同儕互評）或自己的課業（自我評估）（課程發展議會，2001）。雖然面對社會風氣和公開考試的壓力，紙筆測考仍然是學校最常用的評估方式，但根據筆者的觀察，不少本港教師都正積極在課堂上嘗試促進學習的評估辦法。下文將以一個小學中國語文教學設計爲例，探討評估工具，可以如何進一步提升學習的品質[1]。

教學設計摘要

年級： 小三

教學範疇：以說帶寫

教學時間：兩教節（70 分鐘）

活動指引：

（一）看圖說話：細心觀察以下的圖畫，然後順序說出一個完整的故事。

圖一：哥哥和妹妹經過菜市場，市場人來人往

圖二：兩人看見一宗罪案發生

圖三：兩人通知附近一名警察

圖四：頒獎台上，兩人領獎

（二） 評估：（只選自評或互評其中一項，完成下表）

☐ 自評 ☐ 互評

1. 內容	2. 發音	3. 語速	4. 音量
☐ 豐富	☐ 吐字清晰	☐ 適中	☐ 適中
☐ 一般	☐ 吐字大致清晰	☐ 太快	☐ 太高
☐ 不豐富	☐ 吐字欠清晰	☐ 太慢	☐ 太低

☐ 自評 ☐ 互評

5. 個人說話	6. 評估員的態度
☐ 以上各項表現不俗	☐ 認真
☐ 以上各項表現一般	☐ 一般
☐ 以上各項表現較差	☐ 不認真

（三）寫出說話練習中，每幅圖畫包含的時間、地點、人物、事情。

（四）根據圖意，順序寫出一個約 120 字的完整故事。

1 為免侵犯版權，本文不便原文照錄，而有關圖片內容的描述及習作指引，都曾經筆者修改，以突出討論重點，探討評估設計的效能。

由以上的摘要可見，教學設計主要「以說帶寫」，先讓學生練習看圖說話，然後再把說話內容寫成文章，由淺入深，層層遞進，應該能逐步引導學生成功完成課業，教學意念合理可取，也應該是語文教師樂意採用的做法。

雖然如此，本文討論的焦點，不在於評鑑「以說帶寫」的效能，而在於檢討「促進學習的評估」的設計和實施方式。首先是配合說話活動的評估安排，附錄的評估量表項目：內容、發音、語速和音量，無疑是說話訓練的基本要求，而無論自評或互評，學生都必須認真投入，評估才有意義，所以量表的設計，應能提醒學生說話訓練的基本要求，和從事評估時的認真態度。然而，除了對不同方面的表現作高、中、低之類的印象式評價外，評估量表還可以怎樣進一步促進學習？如果類似的評估形式只是機械地重複使用，久而久之，不但形同虛設，更有可能令學生厭，弄巧成拙（Tang, Leung, Chow & Wong, 2010）。其次，案例中的練習由說話與寫作兩部分構成，既然設計意念預期評估可以促進學習，爲甚麼寫作部分又不爲學生提供自評或互評的機會，以加深學生對看圖寫作要求的認識？在設計教學活動的同時，促進學習的評估安排，似乎尚可從詳計議。

筆者認爲，當學生認識自評和互評的基本操作後，教師就要構思具體的評估要求，而要求必須與學習重點對應，評估才會產生較大的學習意義。以上述個案爲例，看圖說話的活動要求是「細心觀察圖畫，然後順序說出完整的故事」，評估量表的項目，哪一項對應活動的要求？顯然，發音、語速和音量都不是。剩下的一項「內容」，評估標準卻是「豐富、一般、不豐富」。且別說學生能否準確拿捏「豐富」的意涵，究竟內容的豐富程度，能否反映說話人如何「細心觀察」圖片，而說出來的故事又是否「順序」和「完整」？由此可見，評估的標準，並未能緊扣學習重點，即使附設了自評和互評的機會，學習意義難免成疑。

要評估發揮更大功用，切合看圖說話的要求重點，促進學習，量表的設計可以參考例一的形式構思：

在適用的 ☐ 內加 ✓。 ☐ 自評 ☐ 互評					
1. 清楚說出圖片內內容		2. 故事完整		3. 表達清晰	
圖一	☐	開頭	☐	聲線響亮	☐
圖二	☐	發展	☐	發音準確	☐
圖三	☐	結局	☐	快慢適中	☐
圖四	☐	順序	☐		

例 一 看圖說話評估量表

與原來設計的量表比較，例一的各項評估準則都貼近了看圖說話的學習重點，由於同時兼顧內容、組織和表達技巧，各項要求只考慮學生能否做到，認爲做得到的項目，才加上✓號。例如，受評者說了半天，把圖一和圖三講述得鉅細無遺，觀察入微，但卻忽略了圖二和圖四的重要環節，評估者就可以在第一欄中只勾選圖一和圖三的方格；具體反映受評者表現的不足之處。又例如故事是否完整，開頭、發展和結局都是基本元素，如果學生能全部掌握這些概念，說話表現大概會比較完善。至於設置「順序」說話的評估標準，旨在重申活動要求，不過是提醒學生注意切題而已。同理，口頭表達是否清晰，聲量、發音和語速都在考慮之列。這樣好比「是非題」的設計構思，沒有就評估項目要求評估者分辨高、中、低的表現，正因爲這並不是本活動（課業）最重要的學習意義。要通過評估促進學生看圖說話的表現，以上量表設計的作用，在於輔助學生加深對課業要求（學習重點）的認識，出發點是先要了解學生是否懂得看圖說話在內容、組織和表達技巧等方面的要求重點，然後才考慮引導學生分辨各個細項表現的水平差異。換言之，同樣是看圖說話的評估量表，設計還有很大的發揮空間，可以（應該）視乎學習需要繼續調整，因時制宜，不斷完善。

有了說話練習的基礎，原來的教學設計是先讓學生寫出每幅圖片的時、地、人、事，然後才「根據圖意，順序寫出一個約 120 字的完整故事」；準備充足，最後的看圖寫作應該比較容易。這時，如果加插互評或自評，量表的作用就不僅能提醒學生寫作的要求，只要略加修訂，還能誘導學生怎樣可以做得更好。如例二所示，除了表達的媒介不同，看圖寫作和看圖說話的要求一脈相通（王家倫，2004），所以量表在內容和組織方面的要求

與例一完全一致。例二微調的部分，主要是針對寫作要求，爲學生提供把故事寫得更好的建議，包括字體端正、善用連接詞和能發表個人感想等出色表現，讓學生檢視。通過反復的互評和自評，學生得以不斷重溫學習重點，逐步認識學習策略，然後自我調節，發展個人的學習方式——這才是 AfL 在課堂實踐的真正意義。因此，要學生在 AfL 的潛移默化過程中學習和成長，訂定明確的評估項目和準則，至關重要。

在適用的 ☐ 內加 ✓。 ☐ 自評　　☐ 互評					
1. 清楚寫出圖片內內容		2. 故事完整		3. 出色表現	
圖一	☐	開頭	☐	字體端正	☐
圖二	☐	發展	☐	善用連接詞	☐
圖三	☐	結局	☐	有個人感想	☐
圖四	☐	順序	☐		

例　二　看圖作文評估量表

以上配合不同學習階段採用的評估量表設計，旨在剖析評估課業，可以如何誘導學生達成更深層的學習意義——通過自評和互評，訂定自己的學習目標、方向和策略，發展自我監控的能力（Stiggins & Chappuis, 2005）。細心設計的評估量表，不但能配合課堂教學活動的需要，還可以讓教與學雙方都有一個較客觀的指標，了解學習的進程和表現，從而探討改進途徑。長遠而言，對培養學生良好的學習態度和習慣，影響正面。

當然，要在課堂上成功實踐 AfL，達至促進學習的最終目的，教師必須正確認識 AfL 的理念，然後針對教學重點，精心構思，才能形神兼備，切合學習需要（Marshall & Drummond, 2006）。作爲 AfL 常用的工具之一，量表因爲形式簡便，往往被濫用和誤用，甚至淪爲粉飾現代課堂的虛假手段，徒勞無功。然而，由本文的案例可見，教師只要釐清學生在不同學習階段的需要，稍加心思，就可以使量表成爲靈活多變、輔助學習的有效資源，饒有意義，實在值得同行繼續鑽研，而教學設計更加不容忽視。

參考文獻

吳毓瑩（2003）。多元評量之解毒與解讀。《課程與教學季刊》，6（1），133-154。

李坤崇（1999）。《多元化教學評量》。臺北市：心理出版社。

教育統籌委員會（2000）。《終身學習 全人發展－香港教育制度改革建議》。香港：政府印務局。

課程發展議會（2001）。《學會學習——課程發展路向》。香港：政府印務局。

王家倫（2004）。《作文就是說話》。西寧：青海人民出版社。

Assessment Reform Group (2002). *Assessment for Learning: 10 Principles*. Cambridge: University of Cambridge Faculty of Education.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Educational Assessment: Principles. Policy and Practice*, 5(1), 7-74.

Klenowski, V. (2009). Editorial: Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy, and Practice*, 16(3), 263–268.

Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133–149.

Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice*, 44(1), 11-18.

Tang, S.Y. F., Leung, P. P. W., Chow, A. W. K., & Wong, P. M. (2010). A case study of teacher learning in an assessment for learning project in Hong Kong. *Professional Development in Education*, 36(4), 621-636.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37, 3–14.

作者電郵： pleung@ied.edu.hk

促進中文科讀寫教學的評估實踐

房湘雲
香港教育局校本專業支援組

促進學習評估的課堂設計

在教學的過程中，評估一直是學與教的核心議題。評估不但回饋老師的教學效能，也同時回饋學生的學習效能，所以評估並不獨立於課程外，而是與課程緊扣重要的一環，因此近年來教育學家不但倡議「學習成果的評估」(Assessment of Learning) ，更進一步發展至「促進學習的評估」(Assessment for Learning) 以及「作爲學習的評估」(Assessment as Learning)[1]。以評估結合教學，促進學習的效能，但如何能在日常的課堂中適切滲透評估的元素，提升教學的效能卻是值得我們深思的問題。

是次的教學設計是聯同新加坡一所大學以及教育局基本能力評估小組，在三方協作下進行。教學計劃是希望透過網上討論平台引起學生的學習動機，讓能力稍遜的中三學生學習如何撰寫議論文。教學活動借用了新加坡一所大學研發的電子討論平台 —「知識研習群」，學生需依據老師事先設計好的「鷹架」，互相評論同學的文章，從而掌握撰寫議論文的方法。

老師在課堂中首先教授了議論文的結構以及要點，然後擬定題目，讓學生根據過程寫作理論（Process writing），規劃及構思整篇文章，並將個人的構思以思路圖形式記下。學生完成了思路圖後，便可將它上載到「知識研習群」的平台上，讓其他同組同學評論，其他同學可以因應思路圖的構思提出不同的意見。對於能力稍遜的學生而言，要掌握同儕互評的技巧並不容易，因爲評論的內容易流於空洞，所以老師在平台上預設了討論的框架。框架的設計是根據鷹架理論（Scaffolding theory），協助學生聚焦同學的構思作出相關的討論，讓能力稍遜的學生可以通過足夠的支援來提升討論的質量。

---

1 見香港大學教育學院中文教育研究中心「面向跨文化學習者的中文學與教：挑戰與突破」研討會 2012 會議資料。

當同學收到其他同學對他的意見後，可以進一步修訂自己的構思，再將構思完善成爲文章的大綱，然後開始寫作。學生完成文章的第一稿後，便可以將作品放上電子平台，讓其他同學細閱。同組的同學可以根據預設的討論框架，如文章的組織、修辭、論據、論證等多方面發表意見。學生參考了這些具建設性的評論後，可以直接回應同儕的意見，也可以進一步收集別人的意見後，再修訂自己的文章。待學生將完成文章的第二稿後，便可將文章上載到平台上供同學再次的評論，作品經過同學反覆的討論後，學生完成了最後的修訂，便可以繳交文章。整個計劃爲期一個月，在這一段時間中，我深被學生的表現感動，以往一直欠缺學習動機的學生竟然可以孜孜不倦的學習，甚至到了廢寢忘餐的地步；成績表現欠理想的同學也可以寫出工整的議論文，他們在學期終評估突出的表現，讓老師重新反思自己在課堂上的教學。

面對學習能力稍遜的同學，我們一直在慨嘆學生的學習能力不濟，抱怨學生的不成材，但我們有否反思自己的教學策略是否能配合學生的學習需要，尤其是面對一群學習能力稍遜的學生？是次教學計劃帶來的反思是教學策略的轉變。沒有單一的教學策略可以滿足所有學生的需要，因此老師的教學策略必須要配合學生的學習需要，哪管是哪一類別的學生。

在整個教學設計中，最重要的元素是將評估結合教學，讓學生的態度變得積極，並提升學生的自我管理能力及學習效能。過去，我們會認爲評估與教學不會同時進行，評估應該是在教學完成以後才會發生。但其實評估並不獨立存在的，在學習的過程中，課程、教學法和評估是三個最主要的元素（Ramsden, 2003），更是學習過程中不可或缺的部分，透過有效的評估能激發學生的動力，讓學生獲取知識和技能的過程變得更有意義和更自然（Berry, 2008）。若評估離開了教學而獨立存在，教與學的效能不但得不到即時的評估，教學也難以作即時修正，合適細分學習的工序，更能有效提升學生的學習動機和學習效能。對於能力稍遜的學生而言，即時和實質的回饋有助他們掌握新學的知識和技能。雖然在同學互評的過程中，老師並沒有直接參與討論，但卻一直追縱學生的討論過程和內容，每天在課堂上與學生總結他們前一天的討論，以及對他們的討論給予即時的回饋，讓他們的討論繼續朝著正確的方向前進。

實質的回饋也是學習中重要的一環。老師在學生學習的過程中，應向學生提供足夠的回饋，內容應集中於學生的學習、課業的目標或準則，並且提供足夠的細節，讓學生易於明白和掌握。回饋也應該是適時和即時的，讓學生能注意到，並依據回饋改善他們的學習 (Gibbs and Simpson, 2004) 。學生在討論的過程中，老師一直在評估他們的學習表現，提供即時的回饋，讓學生能反思他們的學習情況以及調整他們學習的方向，令學習充滿了趣味。透過這教學計劃，學生達至另一深層的學習結果，不獨掌握學科上的知識，而且發展了更多的共通能力，例如解難、溝通技巧、能與其他人合作以及持續學習的能力（Carless, Joughin, Liu and Associates, 2006）。

學生在完成第一稿的修正後，老師發現大部分的學生未能掌握議論文的結構，以至文章結構鬆散。老師因應學生的表現，參考了基本能力評估小組「網上學與教支援」其中一個課業，以「蟲蟲定律」來幫助學生理解議論文的文章結構，讓學生對文章結構的要求有更具體、清晰的掌握。透過老師適時適合的回饋以及具體的比喻，學生對議論文結構的掌握反映在第二稿的討論中，由此可見評估應是結合在教學當中，緊貼在課程和教學而存在的。

評估的模式並非單一，也不應只限於紙筆，可以是與學生討論他們的學習進展，讓他們掌握學習的成果和目標。他們不僅是評估自己的學習成效，而且還可以向同儕學習，互相分享他們朝著學習目標的學習進程，也向同儕和老師尋求回饋（Berry, 2008）。同儕互評讓學習的過程變得更有趣味，使學生更投入學習的過程，透過評估鞏固已學的知識，並能對自己的表現作出評價，警覺自己的功課，令學習過程變得豐富和有意義。因此在教學計劃的過程中，學生比以往更加積極學習，甚至工作至深夜，仍然努力不倦，正是因爲評估強調學習者在學習過程的角色，以及突顯利用評估去提升學習者對自我學習的管理能力（Berry, 2008）。

隨著學生的學習動機提升，學習態度也變得更積極，反映了教學策略能配合學生的學習需要，適切的評估更可以爲老師的教學提供反思的空間，讓老師能即時調整教學的步伐以及策略，解決學生的學習難點，爲學習能力稍遜的學生製造成功的學習經驗，激發他們持續學習的動力。

促進學習評估的遷移

教學計劃結束後，老師將計劃中的重要元素——鷹架及評估理論轉化至閱讀及讀寫結合的教學中。在閱讀教學方面，以往的教學重點集中在字詞解釋、課文解釋等方面，以至學生難以掌握全文，未能理解文章的深層意義。在教學計劃完結後，閱讀教學重點轉爲學生對通篇文章的理解和掌握，在課堂上以提問形式協助學生理解文章的內容和結構。老師透過觀察學生在課堂上的表現，評估學生的學習進程，其中包括老師與學生、以及學生與學生之間的互動，遇上學生未能回答老師的提問時，會以鷹架理論結合提問技巧，將問題細分，以轉問、追問的方法引領學生逐步找到正確的答案，爲學生的學習製造成功感，減輕他們學習的壓力，亦期望能培養學生建立適合自己的閱讀策略，讓他們將閱讀能力遷移至其他的學習上，培養他們終身學習的技巧和能力。

讀文教學強調學生能理解課文內容、文章的深層意義以及結構，待學生掌握了文體的分析能力後，便可以將所學進一步轉化至寫作的層面上。在寫作練習前，老師先與學生溫習在閱讀教學上學習到的知識，然後引導學生將已學的技巧和知識運用至寫作上。在整個轉化學習的過程中，仍以鷹架作爲學生學習的扶手，協助他們歸納在篇章上學到文章結構、佈局和表現手法，在寫作練習中運用出來。經過不斷的練習，學生的努力最終能在公開試中反映出來。

由於教學策略的改變，老師對學生的評估更全面，更準確掌握學生學習能力的起點，有助調適相應的學習步伐和學習進程，結合課堂教學的評估促進學生的學習，而且能發展他們評估自己學習的能力（Carless, Joughin, Liu and Associates, 2006），培養了他們學習的信心和毅力。

反思

借調至教育局校本專業支援組後，參與了更多的校本支援工作。在過去一年的工作中，有不少成功的支援經驗，但也曾遇上不少的困難。反思校本中文讀寫教學的實踐，結合今年相類似的支援經驗，讓我體會到若要提升學與教的效能，必定涉及評估與課程的處理，當中包括了增加的工作量或是翻江倒海的教學變動，老師面對校外老師或專業團隊的支援抱有不同的態度，而他們的態度直接影響支援計劃的成效。

在支援的過程中，支援者或課程改革的舵手的支援重點是鼓動老師的積極參與，而非在學校強行推展支援計劃。支援團隊就老師的教學策略、課程和評估設計提供客觀而專業的意見，有助提醒老師在長期繁忙工作下容易忽略的要點，讓老師有機會停下來重新反思自身的教學工作，與支援團隊協作，實踐不同的教學策略，進而檢討實踐的成效，重整教學策略和步伐。

在這個過程中，老師不應是被動地接受支援，應是擔當主動研究員的角色；因爲不管是教學策略，還是課程與評估的設計都需要老師的參與、潤飾和推行，他們是課程改革的研究者，而不是附從者。老師從實踐中歸納出適合學生的教學策略，因應學生的學習需要進一步調適和完善課程、評估和教學的技巧，這都依靠老師的專業判斷以及貢獻，好比是從事教研人員一樣對新的事物進行無止境的探究，不斷提升學與教的成效。這一種勇於創作及探究的精神並非外力可以強加在老師的身上，是需要老師從自身誘發出來，勇敢主動踏出嘗試的第一步，接受不同的意見，懷著探究的精神，經過內化的過程，最終能演繹出屬於他們個人的教學模式。由於沒有一套萬用的教學程式，所以老師只能因應本身的特質和學生學習的需要加以轉化和昇華，成爲課程的研究者發展適合他們的教學策略，這正也是新課程改革的精神。

繁重的教學工作像是永無止境，但老師也並不是孤立無助的，因爲在整個教學的世界中總會找到同行者，勇於面對著眼前的困難，披荊斬棘爲我們的學生開創前路。未來的課程改革也得依靠老師不斷的努力，對教學策略不斷的鑽研，才可以照顧不同學習需要的學生，發展他們的潛能，讓他們茁壯成長。

參考文獻

Berry, R. (2008). *Assessment for learning*. Hong Kong: Hong Kong University Press.

Carless, D., Joughin, G., Liu and Associates (2006). *How assessment supports learning: Learning-oriented assessment in action.* Hong Kong: Hong Kong University Press.

Gibbs, G. and Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3-31.

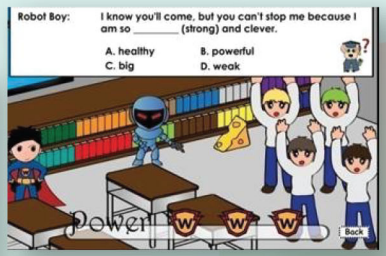
Ramsden, P. (2003). *Learning to teach in higher education (second edition).* London: Routledge.

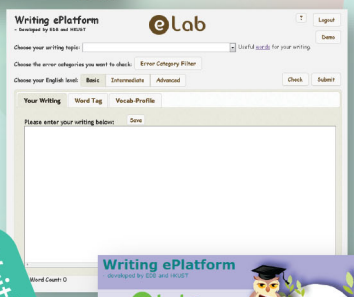
作者電郵： sheungwan.fong@gmail.com

WLTS | Assessment for Learning Resource Bank
Web-based Learning and Teaching Support

wlts.edb.hkedcity.net

- ✓ Interactive resources to address students' learning problems
- ✓ Units of work to support assessment for learning (AfL)

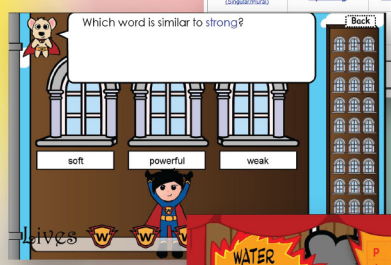




Sight Word Park

Super Word Kid

Writing eLab & eTutor



Ready-to-use Presentations • Interactive Activities • Online Games

Research Reports • Drag & Drop Exercises • Lesson Plans

Formative Assessment and Effective Teaching and Learning*: Consolidating Gains from Singapore, Brunei and Hong Kong

Seok Moi Ng
Ministry of Education, Singapore
and The University of Auckland

Abstract

Several extensive reviews of the research literature have concluded that formative assessment is the most powerful factor in raising student achievement. This paper explores the questions:

- What is formative assessment?
- How does formative assessment operate in the classroom?
- Under what conditions will it work?

This paper discusses findings from studies of formative assessment and the experience of several English language projects: how classroom discussions, questions, activities and tasks can be used to elicit evidence of student learning, how this feedback can be used to propel students' learning and how to set up conditions for activating students to own their learning.

INTRODUCTION

The paper draws on the author's 20+ years of experience in large scale professional development projects, working in collaboration with classroom teachers, asking the same questions as those asked by teachers and seeking best fit 'solutions' for the day to day challenges faced by teachers. All these projects aimed at improving children's English language development and their engagement with learning (LLELP, 2006; Ng, 2001; Ng & Sullivan, 2001) and

**Updated version, originally published in Cane, Graeme. (Ed.) 2008. Strategies in language learning and teaching. Singapore: SEAMEO Regional Language Centre and reproduced here with the kind permission of SEAMEO Regional Language Centre.*

engagement they all involved children who were learning English as speakers of other languages (ESOL). The earliest programme in Singapore (1985-1990) worked at the lower primary levels (Primary 1-3) with the later programmes in Brunei (1989-1994) and Hong Kong (2000-2006) covering Primary 1 to 6. The Hong Kong project had also a simpler version of the programme at the preschool level (1997-1999) and Singapore (where English has become the main language for classroom instruction), has developed an updated version for Primary 1 through to 6 (2006-2011). Exploration is currently ongoing in Singapore to adapt the basic conceptual framework to the Lower Secondary (from 2010) and Kindergarten levels (from 2012).

The lower primary (Primary 1-3) projects were book-based programmes involving elements of Shared Book and Language Experience Approaches, suitably adapted to the educational systems they were operating in. The teaching techniques in the upper primary programmes (LLELP, 2006) were expanded to include Sustained Silent Reading (SSR), Guided Reading (GR), Know-Want to Know-Learn (KWL – a reading comprehension technique, Ogle, 1989) and Retelling (Brown & Cambourne, 1987).

There was also a greater emphasis in the upper primary on explicit teaching of text-types and grammar. All the projects involved mainly students from average income home backgrounds in the three educational systems. In the recent Singapore adaptations, the goal is to make freely available to all schools the resources of the developed programmes.

These programmes were based on sound pedagogical principles consistent with the findings of research into effective reading and language classrooms (see for example Duffy, 2003; Duke & Pearson, 2002; Hoffman, McCarthey, Elliot, Bayles, Price, Ferree & Abbott, 1998; National Reading Panel, 2000; Pressley, 2006; Pressley, Rankin & Yokoi, 1996). Features promoted by these programmes include:

- promoting high academic engagement in the classroom through an emphasis on motivating students.
- a combination of whole class and small group teaching that provides scaffolding for students as they read and write. This allows students to work within their ‘zones of proximal development’, that is, being challenged without being frustrated.
- a greater proportion of time spent on actual reading and writing than on workbook exercises.

- a range of levels of authentic literature and other texts that allow students to read and experience according to their strengths and needs.
- language skills instruction that provides a balance of holistic reading and writing activities, often in response to the specific needs of the students.
- instruction in comprehension strategies.
- frequent opportunities to write within a plan-draft-revise process that focuses on the coherence of ideas as well as the mechanics of English.
- explicit teaching of word-identification skills and the strategies of using letter-sound associations and word parts to sound out words, knowledge of high frequency words as well as picture, text structure and syntactic cues to make sense of the text.

Learning is facilitated through integration of language features (at the micro and macro levels) to provide reinforcement and practice of target teaching goals. In the Hong Kong programmes for example, integration at a macro level, was achieved by using similar themes and topics as those used in the Chinese language curriculum. Learning activities were also integrated through the different lesson components, employing the language skills of listening, speaking, reading and writing. The different parts of any one lesson (e.g. song, book reading, language activities) were also integrated with the language and content of the book for a particular unit of work. The strategy of integration is a response to the challenge of how to provide adequate coverage of the features and characteristics of the English language. It seems that in order not to overcrowd the ESOL curriculum, the only way to cope with the multifaceted nature of the English language is through integration in as many ways as possible. Besides, such integration facilitates understanding for the learner. For ease of reference, the three programmes will be referred to as the Integrated ESOL (IESOL) programmes, underscoring the importance of integration in all three programmes, even though the three programmes have different features. Most of the examples and transcripts of classroom interactions are drawn from the project in Hong Kong.

While these IESOL programmes could be considered successful as shown by the results of internal and external monitoring and testing (e.g. Ng & Sullivan, 2001; Ng, 2001; Martin & Abdullah, 2003), it has to be pointed out that finding appropriate teaching techniques and materials is only a part of effective teaching. It was unfortunately observed that these IESOL tools and

procedures were used by a few teachers in the manner that has been institutionalised over many years – using traditional drill and rote memory. The search for improving IESOL programmes is a continuing effort and upcoming research has been and is being scoured for other strategies/factors that affect learning. The following sections of the paper will discuss important guiding principles of the IESOL programmes and their relation to an important aspect of teaching that applies across specific approaches, techniques and content areas, and assessment.

Assessment, Examinations and Teaching Practice

In the three Asian educational systems I have worked in, examination pressures exert a great influence on classroom practices. Accountability to parents and administrators is a big challenge and the most common assessment procedures are those measuring outcomes. Often, there are mid-year and end-of-year examinations and pre-examination preparation and tests. In many schools at levels where high stakes examinations are placed, there are monthly and sometimes weekly assessments. The traditional teaching style is to teach for two or three weeks, and at the end of that period, students are assessed. Whatever the results of the assessment, the teacher then goes on to the next topic, because there is a syllabus to be covered for an examination.

Outcome measures such as these yield composite marks or grades. They are useful for the classification and comparison of children, classes and schools, but the teacher needs to know more than just the end points of instruction. Teachers need to know what happens between assessments in order to obtain feedback information for instructional purposes. Additionally, teacher professional development programmes (whether pre- or in- service) have been shifting attention from teaching to learning – to a focus on what the students are getting out of the process rather than on what teachers are putting into it. In traditional lessons, teachers are often doing all the work. Besides managing classroom dynamics, reading set texts, they are often observed to be explaining all the words and concepts that are thought to be difficult for the class, often doing the asking and even answering their own questions! It’s no wonder the teacher gets better at the learning outcomes being promoted – the teacher is doing all the work, not the student. This paper examines what can be done to involve students more so that they are not mere recipients of teacher initiatives in the classroom but become more active in deciding what goes on in their learning, including the assessment of their learning.

FORMATIVE ASSESSMENT AND FORMATIVE FEEDBACK

In 1998 Paul Black and Dylan Wiliam of King's College London published their wide-ranging analysis of research into classroom-based assessment across curriculum areas (Black & Wiliam, 1998a; 1998b). Over a period of nine years, they surveyed many books, more than 160 journals and earlier reviews of research. Out of this, they studied 580 articles and chapters and published a review (Black & Wiliam, 1998a) with this main conclusion:

> 'There is a body of firm evidence that formative assessment is essential to classroom work and that it can raise standards. We know of no other way of raising standards for which such a strong prima facie case can be made on the evidence of such large learning gains.' (Black & Wiliam, 1998b. p.19)

About the same time, John Hattie from the University of Auckland in New Zealand was studying factors that made a difference in the classroom (Hattie, 1999). He made a synthesis of 337 meta-analyses, including 200,000 effect-sizes from 180,000 studies, covering nearly all methods of innovation. He concluded that the most powerful single factor for raising achievement is feedback based on evaluating students' understandings, a finding similar to that of Black and Wiliam.

Conclusions from earlier reviews and research studies into the effects of feedback in schools, colleges, and workplaces, (see for example, Crooks, 1988; Kluger & DeNisi, 1996) are rather sobering – much of the feedback that students receive has, at best, no impact on learning, and can be counter-productive. There were several other findings that run counter to common classroom wisdom about feedback. Some of these are:

- The incidence of feedback for each student in the typical classroom is very low, usually in seconds per day at best (Hattie, 1999).
- Global grades or simply confirming correct answers have little effect on subsequent performance (Crooks, 1988). When students were given only marks they made no gain from the first to the second lesson (Butler, 1988). Students given only comments scored on average 30% higher. Giving marks alongside comments cancelled the beneficial effects of the comments. This suggests that evaluative feedback including a mark/grade does not help pupils improve as much as descriptive feedback without a mark/grade.

- Repeated explanations that have previously led to failure are less effective than employing alternative strategies (Fuchs, 1993; Fuchs & Fuchs, 1986).
- An over-emphasis on feedback related to conduct or neatness is likely to have a detrimental effect on the intellectual quality of children's work (Dweck, 1986).
- Assessment of pupil performance is often in affective terms. Feedback which draws attention to self results in negative effects on performance (Bennett & Kell, 1989; Hattie, 1999).

These studies showed that while on average feedback did increase achievement, in 40% of the studies feedback often made people's performance worse than it would have been without feedback. This occurred when feedback involved the students' self-esteem, where the feedback focused attention on the person rather than the quality of the work. For example assessing students' work through marking and feeding back grades, marks or other forms of report that encouraged comparison with others is likely to label students and reduce their self-esteem.

These studies show that students' achievement is not enhanced if the teacher separates assessment from teaching (Harlen, 1998). The biggest positive impact of assessment on performance occurred when feedback told not just what to do to improve, but also how to go about it (Black & Wiliam, 1998a; Harlen, 1998; Hattie, 1999; Kluger & DeNisi, 1996, Torrance & Pryor, 1998; Wiliam, 2007).

Some Formal Formative Assessment Tools

Books for language teachers often contain recommendations for measuring tools (e.g. Calfee & Perfumo, 1996; Johnston, 1992; Glazer & Brown, 1993) to record observations of learners at the beginning of the year with repeated measures to check their progress. In the IESOL programmes (LLELP, 2006; Ng, 1988; Ng, 2001; Ng & Sullivan, 2001; Ng & Sullivan, 2008), teachers were introduced to observational checklists for recording early literacy behaviours (see for example, Jabatan Perkembangan Kurikulum, 1990). These consisted of categories of behaviour such as knowledge of the letters in the alphabet, some concepts about print, colour, numbers and oral responses observed during lessons. The intention of these tools is formative – to find out children's language needs so that teachers can craft instruction to meet those needs.

Other recommended tools include anecdotal records/notes, a "portfolio" or file for each student containing dated samples of his/her work and performance tasks that show what the student can do. Anecdotal records or notes trace students' progress throughout the year to provide evidence of growth or difficulty. They are usually kept in a positive tone about a student's progress, a sort of "kidwatching" but recorded on paper. These records are made as soon as possible after the event and should be kept strictly confidential, especially if they contain sensitive information. Performance tasks show what students can do rather than what they can memorize or recall. For example, listening to a student read aloud periodically can indicate progress in word identification strategies.

While most of the above tools can be used for measuring outcome, they can also be used formatively in that the teacher can use the information obtained to monitor the learners' ability to use oral and/or written language and design teaching moves to match students' need. As defined by Black & Wiliam (1998a, p.2), formative assessment or assessment for learning:

> "... refers to all those activities undertaken by teachers, and by the students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged. Such assessments become formative when the evidence is actually used to adapt the teaching to meet the needs."

Hattie (1999) argues that if we accept his conclusion that the most powerful single moderator that enhances achievement is feedback, the prescription for improving education must be "dollops of feedback" – providing information about how and why the student understands and misunderstands, and what directions the student must take to improve. This is not rocket science, as for decades colleges of education have stressed that the teacher needs to find out where the learners are at and what they know in order for effective teaching moves to be made. What is new is the powerful potential of this seemingly simple principle. More recent studies have found that students taught by teachers who used 'assessment for learning' concepts approximately doubled their speed of learning – these students achieved in six or seven months what would otherwise have taken a year to learn in other classrooms (Wiliam, Lee, Harrison & Black, 2004). More significantly, these improvements appeared to be consistent across countries (including Canada, England, Israel, Portugal, and the United States), age levels and content areas. The researchers also found after working with teachers in England, that these achievement gains

could be sustained over extended periods of time and when measured with external standardized tests (Wiliam, Lee, Harrison & Black, 2004).

The features of formative assessment and assessment for learning (Stiggins, 2005) are consistent with the principles used in development of the IESOL programmes. The following sections describe these features, which many IESOL teachers have been implementing in their classroom instruction for effective feedback and improved learning. These features are interconnected and in reality hard to separate one from the other, but for the purposes of discussion it is convenient to consider them separately.

Having Clear Learning Goals and Intentions

For a start, teachers should be very clear about the learning goals, the expected outcomes of classroom tasks and how the learners can achieve those goals. These goals and expected outcomes must also be communicated effectively to students. Prior to implementation of the IESOL programmes, many classrooms were 'teaching the textbook' and it would not be unusual to have students answer 'Page twelve' to a question as to what they were doing. For many students then, the purpose of English lessons was to go through the textbook, and what mattered was that their written work was accurate, neat and of the specified length required by an exam.

The following transcript shows a typical classroom interaction found in those traditional classrooms.

Transcript 1

Teacher:	Where's the pencil? Mary. (Teacher pointed to a pencil placed in a jar)
Mary:	The pencil on the jar.
Teacher:	Please repeat your answer. The pencil…
Mary:	The pencil in the jar.
Teacher:	is
Mary:	is
Teacher:	in or on
Mary:	in or on (Mary looked puzzled)
Teacher:	in or on? (Teacher pointed to helping words on the chart)
Mary:	in or on the jar! (Mary with a bright smile)

This scenario is played out in countless classrooms all over the world – the "guess what is in the teacher's head" game. The learner is left in the dark as to the real purpose of the whole lesson.

Because language learning is complex with its multiple levels, the IESOL programmes have various lesson components (see Table 1 for a typical IESOL unit plan) that interweave between whole texts, paragraphs, sentences, phrases, words and word parts (e.g. letters). Each IESOL unit is built on a cycle of activities, with whole class teaching as well as instruction in small groups and individual work. Instruction proceeds from an intensive engagement with whole text followed by deconstruction of whole text for a detailed exploration of smaller bits of language (e.g. letters and sounds, words, sentences) and then moves on to the more difficult task of constructing text in writing. With these many levels in language learning, it is crucial to make transparent to the student the learning intentions for each of those components, and what would count as success. For example, in an IESOL unit of work that may cover 1-3 weeks, the first component is often the class reading of a piece of text. In the emergent literacy stages at Primary 1, the targets would initially be to enjoy and understand the text (often a story in the early stages) in a Shared Reading lesson, complemented by follow-up activities aimed at developing basic decoding skills for word identification. [See Table 1]

A learning goal in English lessons will often be covered over a number of sessions, but teachers should explicitly share the goal with students at each lesson. Each lesson may have a different focus, with the level and needs of the students determining whether the learning intentions involve, for example, strategies for word identification, working out the meaning of words or other comprehension strategies. Especially when there is explicit and direct instruction on smaller bits of the language as in Part 2 (see Table 1), the students should be clear about how that study of small bits of language links to the larger pieces, as students can easily get lost in 'seeing the trees and not the forest'. In the IESOL programmes, when novices are exploring parts of words, they will also be shown how they can apply them to reading and to remembering unfamiliar words. The note below is extracted from typical lesson guidelines for the IESOL project teachers.

> Note: This activity helps students understand that some English vocabulary items are organized by and can be divided up into smaller meaningful units. This is a very useful vocabulary-learning technique which the students can employ (1) when they come across difficult and unknown vocabulary and (2) for retrieval and retention. When you are introducing new vocabulary to the students, get them to break the new words into syllables and look to see if there are small words in the big words.

It is advisable to have learning goals revisited as the lesson progresses and at the conclusion of the lesson. For example, at the end of a reading lesson, the session is briefly summarised and the learning goal and success criteria established for this lesson are reviewed. Students are encouraged to articulate and reflect on their own learning and to set goals for the next reading lesson.

IESOL classrooms have been rather successful in conveying the main goals of language learning, as revealed in student interviews conducted at the end of the Hong Kong project. When the students were asked about what they had learnt in the project lessons, very few mentioned specific items of knowledge (e.g. particular words or poems learnt). Most of them described the language processes they had learnt (see examples in Transcript 2) and many were able to link what they had learnt to what they would be able to use when they went to secondary school.

Table 1. IESOL Activities in Order of Presentation

Time-line	(1-3 weeks in total) →					
Part/ Component	1	2	3	4	5	6
Activity	Shared Reading, Guided Reading, etc.	Language Activities	Language Experience and Class Writing	Group Writing	Individual Writing	Learning Centres for listening, speaking, writing, reading, viewing and presenting
Examples of Targeted Skills/ Knowledge	Reading strategies, with opportunities for speaking, listening, writing and visual literacy.	Explicit instruction in word study, grammar and language use at the various levels of language (word part, word, phrase, sentence, text, inter-textual).	Writing, both composing and editing processes with opportunities for speaking, reading, listening & visual literacy.	Writing, both composing and editing processes with opportunities for speaking, reading, listening & visual literacy.	Writing, both composing and editing processes with opportunities for speaking, reading, listening & visual literacy.	Differentiated practice of all language skills, supported and/or independent, at the various levels of language (word part, word phrase, sentence, text, inter-textual, etc).
Amount of Teacher Support provided	More → Less					Differentiated support, depending on progress level.

Transcript 2

Andy: I have learned phonics in the lessons. Phonics is very useful, when sometimes we meet some unknown words, we are able to break them into syllables and pronounce it correctly.

Janice: I like the retelling part most. First of all, our teacher gives us a text. Then when we are reading the text, we would use the highlighter to highlight the important points. After reading, we need to use our own words to rewrite the text. After rewriting, we need to share our work to our group members. Our group members will give us some comments. Then I can know which part I can do better. Sometimes, I will share my work to the whole class. My classmates will give me some comments. I can learn from them and also I can build up my confidence.

The learning goals and intentions of formative assessment are not the same as those of the traditional behavioural input/output models. What is important in formative assessment is an insight into the constructions that students have made from their learning activities. Rather than evaluating in order to reinforce set learning objectives, this insight is used for feeding into instruction (Hattie, 1999). Another point of difference is that English language learning is not viewed as chiefly a matter of acquiring items of knowledge. Some teaching conversations may revolve about consonant blends, new words or concepts, sentence structures or the theme of the text, but an equally strong emphasis in formative assessment and IESOL programmes is on the acquisition of processes and strategies. For example, in writing there are processes of brainstorming, drafting and editing and in reading there are strategies for monitoring comprehension processes and repair strategies when comprehension fails. Here, the learning outcome is not only the comprehension of the specific piece of text used in the lesson but rather the reading strategies of prediction and the skill of seeking evidence to confirm predictions made.

Using Information from Assessment for Teaching and Learning

Harlen (1999) pointed out that assessing and gathering the information from assessment is only part of the formative assessment strategy. It is the use of the information gathered that distinguishes it from other purposes, and teachers should adapt their teaching to what has been identified about students' learning needs and strengths. This is an essential aspect of teaching-learning in formative assessment.

Such use of information is exemplified at one level in IESOL project teachers' regular planning of the curriculum for the coming year using the students' assessment results from the previous year. The yearly plan could be broken down into terms or semesters and could be modified, trimmed or expanded as the year progresses, depending on students' progress. Other forms of planning could take place on a weekly or bi-weekly basis, when teachers of the same grade level get together to reflect on the recent lessons taught in the classroom and the information obtained from students' progress, and use this to plan for the coming lessons. This is similar to practices promoted in other countries such as New Zealand (Bell & Cowie, 2001).

In comparison to these pre-planned (and more formal) uses of formative assessment, an example of daily formative assessment use is found in the powerful teaching procedures of the Reading Recovery lessons developed by Clay, one of the greatest teachers of our time (Clay, 2005). The Reading Recovery programme is an early intervention programme directed towards providing individual help for children having difficulty in reading and writing, to 'recover' them and restore them back into normal classroom teaching. The programme is able to do this through structuring lessons that are different for every child. The various teaching activities allow the teachers to evaluate students' understanding and then match next teaching acts to the present understandings of students. The best Reading Recovery lessons occur when the teacher responds always to what the child is trying to do, using moment-by-moment assessment and teaching decisions. Clay drew a useful analogy between conversation and Reading Recovery teaching in these words:

> "Sensitive and systematic observation of young children's reading and writing behaviours provides teachers with feedback which can shape their next teaching moves. Teaching then can be likened to a conversation in which you listen to the speaker carefully before you reply."
>
> (Clay, 1985, p.6)

An injunction from a Piagetian psychologist, Duckworth (1981) stresses the importance of teachers understanding what and how children understand. While one-to-one tutoring is not the norm in a classroom, the Reading Recovery example demonstrates the power of the basic formative assessment principle: teachers should find out where the learners are at and what they know, before teaching moves can be made to take them from where they are at to somewhere else.

Often teachers are full of what is in their own heads, like the teacher in the next transcript of part of a shared reading discussion (Transcript 3). The teacher was looking for 'big' as the answer and missed the opportunity to work on the student's contribution 'giant' and perhaps extend it to 'gigantic' – a word that could be within the child's grasp.

Transcript 3

John: It's giant.
Teacher: No, not giant, big. Say it's big!

In Transcript 4, the error made indicates that the child is probably already looking very carefully at the letters because only one letter is incorrect ('h' for 'l' in sleeping). It is likely that if the child is given feedback to refer to the context and picture clues as well as attention to letters, she will be able to read the word and learn something about the skill of using the various cues to help her read (see Clay, 2001; 2005 for a discussion use of cues in reading strategies).

Transcript 4

Sue (reading): Mary is sheeping
Teacher: Aiyah, I keep telling you, look carefully, s-l is sleeping.

IESOL teacher development workshops discuss opportunities for applying formative assessment principles on a daily basis in the classrooms. Novice teachers have found it rather hard to monitor individual students' understandings on the run, as individual students learn different things as a lesson is being presented. Gathering responses to assess the understanding of the whole class in real time is obviously a challenge and task sheets have been provided to help teachers gather information about students' learning for feedback into teaching decisions. This is of some help, but currently an IESOL team is trying to develop strategies to help them learn from other teachers who have found techniques that work for whole class and/or small group teaching.

An effective technique used by those teachers (Leahy, Lyon, Thompson & Wiliam, 2005) is to have all students write their answers on individual dry-erase boards, which they hold up at the teacher's request. The teacher can then scan responses for novel solutions as well as misconceptions. Another technique is to give each student a set of cards labeled for multiple-choice question-answer format. If the question is well designed, the teacher can quickly judge the different levels of understanding in the class and make various teaching decisions. These could be to move on, to reteach the concept or to review

particular aspects of that topic. Teacher preparation for these techniques should take about the same time as that required for the alternative – developing task sheets or assessment papers to provide feedback about students' thinking and understanding for shaping students' next learning steps (Leahy, Lyon, Thompson & Wiliam, 2005).

Their teachers have discovered that through careful planning and thinking about the questions they ask in class, teachers can check on students' understanding while the students are still in the class rather than after they have left, as is the case with grading (Leahy, Lyon, Thompson & Wiliam, 2005). The time saved from marking can be used in the pre-lesson planning of effective classroom discussions, questions, and learning tasks.

Some teachers apply formative assessment processes more successfully in the writing components of the IESOL unit (Parts 3-5 in Table 1) than in the oral interactions and discussion because the writing activities are slower in real time. Feedback about writing could occur when the class works together with the teacher on a piece of class writing or when done in small groups. However, teacher feedback operates best in individual writing conferences, when the teacher works with students on their drafts and can provide constructive suggestions for improvement based on diagnostic information from a student's individual writing. It should be two-way between the teacher and student, where the teacher can discuss the student's writing with the student and get direct feedback about what the student understands or misunderstands. More discussion of what goes on in individual conferences can be found in the section below.

Another IESOL component that sits well in the formative assessment framework is Learning Centres, in which students are grouped according to progress level. Through preceding lesson components, the IESOL teacher has gathered information about how the students are achieving the goals of a particular unit. For example one of the learning outcomes could be that students construct instructions for a simple recipe. Throughout the unit, students may demonstrate their understanding (or not) of the procedural text-type, its text features and language structures, together with knowledge of the vocabulary of a certain kind of recipe. In Learning Centres that occur at the end of the unit, the teacher may direct students depending on their learning needs to activities that reinforce what has been learnt, visit any aspect of learning not attained thus far or extend students in areas that they can control independently. Smooth management of Learning Centres also allows the

teacher to work with a lower progress group to provide better support for their learning, or to stretch a higher progress group to greater challenges.

The critical element in all these varied examples of formative assessment is the gathering of feedback to provide information about the learners' understandings (or misunderstandings), and the use of that information to show what directions the students must take to improve. The next section discusses in greater detail the kind of feedback that is most useful to learners.

Generating Informative Feedback in Quality Classroom Conversations

Aural-oral skills are important in life. In the classroom, speaking and listening are the most often used skills (Brown, 1994) and out of the classroom, listening is used twice as often as speaking, which in turn is used twice as much as reading and writing (Rivers, 1981). Speaking requires that learners not only know how to produce specific aspects of language such as grammar, pronunciation, or vocabulary but also that they understand when, why, and in what ways to produce language. Speech has its own skills, structures, and conventions different from written language (Burns & Joyce, 1997; Carter & McCarthy, 1995; Cohen, 1996). A good speaker integrates this array of skills and knowledge to succeed in a given speech act.

While learning to talk, the young child is also learning through talk (Roskos, Tabors & Lenhart, 2004) and the quality of talk that each child hears and participates in is an important part of that child's education (Cazden, 2001). Marshall and Wiliam (2006) maintain that even for secondary students, talk is an important aid in developing understandings, which are best fostered through interactions with those who already have those capabilities. For all these reasons, there should be abundant opportunities for rich and deep conversations between teachers and students in every classroom.

However, classrooms may not be conducive for such conversations as an individual child engages in a limited amount of conversation during a classroom day, even in preschool centres (Dickinson & Tabors, 2001; Smith, 1999). With one teacher to many students, formal routines are necessary for managing behaviour. Cazden (2001, p. 82) argues that the teacher's speaking rights are inherent in their institutional role.

> "Teachers have the role-given right to speak at any time and to any person; they can fill any silence or interrupt any speaker; they can speak to a student anywhere in the room and in any volume or tone of voice".

Hiebert and Raphael (1998) used the term ‘scripts’ to identify particular patterns of interaction and socially acceptable modes of communication. Scripts in and of themselves are not inherently good or bad. The problem arises when one script, like the one below, is the only script for classroom talk.

> No. Listen. Are you listening? Listen.
> We’ll put one in each cup... you see like this.
> And then we count them.
> And then we put them in the oven.

The quality of classroom talk has been the subject of many investigations, especially the quality of questions used in classroom conversations. In a longitudinal study of capable children, top on the list of items that showed positive relationships with children’s competencies at age 10 was that staff asked the young children open-ended questions (Wylie & Thompson, 2003). In contrast, there are the ‘school questions’, often asked where the questioner already knows the answer – questions that are commonly known as closed questions (e.g. What is in this story? How many animals can you see on the page?). These questions are not genuine questions but a means of moving children into playing a game of guessing the approved answer the teacher already has in mind (Commeyras, 2001; Ng, 2006).

Questioning techniques is an early topic for IESOL teacher workshops, discussed in relation to shared reading activities, when the teacher introduces a book to stimulate children to share their feelings and experiences in a group situation (see Transcript 5).

Transcript 5

Teacher:	Let’s read this book about a boy who was scared.
	[Teacher reads.] ***There’s A Nightmare In My Closet.***
	What do you think is hiding in the closet?
Lottie:	It’s nightmare. Don’t like nightmares.
James:	I’m not scared.
Teacher:	Aren’t you scared of nightmares?
James:	I don’t... I’m... I don’t... I’m not scared.
Teacher:	You’re not scared of anything?
James:	I’m not scared of anything!
Teacher:	I’m scared when it thunders, when it’s very loud.
Sue:	I’m scared of the rain.
James:	I love thunder. I love thunder, but I don’t like lightning!
Lottie:	I’m scared of snakes
Teacher:	We are scared of different things. Let’s keep that in our heads as we read what happened to this boy when he feels scared.

This is an example of a teacher's efforts to connect the learners' experience and knowledge with the knowledge about to be presented in the text. The story is about a child's fears and the conversation links it to the children's own fears – children talk about what they are (and are not) afraid of. Note that the teacher does not limit the children's responses to yes/no answers and children can talk openly about their fears.

The next transcript shows more clearly the teacher assessing the students' responses and using that information to shape students' understanding of a particular text type. The teacher had asked her students to research the topic of polar bears, following lessons that examined information reports about such topics as bats and spiders. The students were offering what they had found from the internet to contribute to a piece of class writing (refer to Table 1).

Transcript 6

Mahn: They have four large paws.
Teacher (repeating and typing in what was contributed)
Chai: They have two layers of white fur covering their black skin.
Teacher (repeating and typing in what was contributed): OK, Bess?
Bess: Their fur is so soft.
Teacher (repeating and typing in what was contributed): Any more we have to add to the paragraph? Yes, Ka Wing?
Wing: We can use 'and' because there are many 'they have, they have.'
Teacher: Excellent, can you repeat?
Wing: They have purple tongues and four large paws.
Teacher: We delete this and put the connective, 'and'? Do you all agree? Does that make it sound better?
Students: Yes.
Teacher (typing in what was contributed): Does this sound better? OK, any more ideas? Rex.
Rex: Add, and we can give them a hug.
Teacher: OK, where Rex?
Rex: the last...
Teacher: The last...?
Ken: But in information report, we can't use 'we'.
Teacher: OK, do you know what they are talking about, class? OK, Rex suggested that we put 'We can give them a hug' because they are so soft and fluffy. But Ken, you have a point.
Ken: But in information report, we cannot use 'we' and 'I'.
Teacher: Yes, because that shows just our personal feelings. But Rex, good try. Maybe next time when we write a story about polar bears, then you can write "We like to hug the polar bears', okay? Good point. So, any more ideas?

Note that except for the topic and the text type that were specified, the above interaction was relatively spontaneous and unplanned. The teacher did not know ahead of time what the students were going to contribute and she had no prescribed model answers to her questions. She provided a lot of wait time to give students the opportunity to explore their ideas so she could listen to where they were going before providing appropriate support. These behaviours are conducive to generating genuine discussion. Davis (1997) would describe this teacher as one who is listening interpretively rather than evaluatively. Instead of mere rehearsal of existing knowledge, an interpretive teacher listens to children's conversation about something that is important to them and then asks open-ended questions that will encourage more language use and creation of new knowledge.

In IESOL programmes, teachers found that the richer the language task, the better the interactions. This is in contrast to some programmes that purport that texts and tasks for ESOL elementary grades should be kept simple. The following transcript is of two students comparing their written retellings of a piece of text on pollution – a difficult piece that employed language not normally encountered at the primary levels in Hong Kong schools. While the transcript showed a drop in the accuracy of sentence structures in the retellings (in comparison to story retellings), it also showed that the students rose to the challenge, eagerly grappling with the concepts couched in academic language and, among other cognitive skills, employing intense listening and evaluation of each other's language use and interpretation of meaning (cf. Brown & Cambourne, 1987). Both students showed some competence in their development of assessment skills when critiquing their peer's retellings; pointing out specific differences between the retellings and what was good about them.

Transcript 7

Connie: I remember that we need to make a clean environment to let the animals have a good green house. Acid rain is made by producing ah... goods in factories and fires from BBQ. Oil spills are made by the ship or ship accidents. It will make life less. Global warming is made by cars and factories. It will let our temperature be higher and higher. Endangered species that we see them fewer and fewer are they, such as coral effect, wild... It was made by wild hunting and oil spill. We need to protect our earth since today... Pollution in the sea are made by people because we throw or dump the rubbish into the sea every day. It will lead to animals' lives in the sea will die, so we need to protect our earth.

Ngan: The report talking about some pollution, just like acid rain, oil spills. It tells us acid rain is some dirty thing go into the cloud and mix with the water and the cloud will fall down and the earth will die. It tells us oil spills, it is come from the boat, when the boat that sailing, it will some oil going into the sea, some thing animals live in sea that will die.

Connie: I think some...first of us are the same, such as we have talked about the pollution, acid rain and the oil spills, many. And the difference between us I think is I have talked how to stop the pollution, and how can we do and you have missed this point. And I think you have, have done the good job because although you are short, but you are clearer than me so I think you have done a good job.

Ngan: I think you have done a good job because you say that, that some people throw the rubbish in the sea and you say we should protect the earth to clean it.

Connie: Thank you.

In line with Kluger & DeNisi's finding (1996) of when the biggest impacts on performance occurred, IESOL teachers are encouraged to provide feedback for not only what to improve, but also how to go about it. The research findings also indicate that it is better to concentrate on giving descriptive feedback rather than mixing it up with the grading aspect. This can be illustrated by working through a few examples of what teachers should do in their individual conferences with children about their writing.

Example 1:

That's a good piece of writing because it describes polar bears in sections …. Now is there anything else you can find out about the appearance of the polar bears?

The first comment in Example 1 is slightly evaluative and describes what constitutes good writing. The question that follows can promote real conversation about the next step to take to improve on the student's work. Example 2 is not evaluative and is intended to get at the students' knowledge of text structure (i.e. narrative elements) and to use that knowledge in his/her writing plan.

Example 2

What is going to happen next? How are you going to get from where the story is now to the next part where a problem develops?

Comments such as "Make your story more interesting." are not very useful, especially for low progress pupils. They should be followed by examples that

show students what to do, such as:

- What does Pokemon look like?
- If you were Pokemon, what would you have done to the bad wizard?

The idea is to provide specific suggestions for improvement and depending on the progress level, more support and scaffolding in the form of suggesting sentence beginnings may be required, such as:

- Pokemon looked very _____. He had _____ on his face and his hair was _____.

There is a limit to the amount of time a teacher has to provide such feedback to each student in the class, but a potential lies in the activation of students as learning resources for one another and to move learners to independence as soon as possible in peer and self assessment (see Transcript 8).

Transcript 8: Extracts from a Conferencing Session between Students

Jen: Ming, your information text is excellent. You have written enough... your information is very good and you have written enough information on one sub-heading. Er, for subheading Life Cycle you might – I think you should put 'Reproduction' instead.

Ming: *Oh ya, because they have two cubs and they are…*

Ming: *... then do I need to change this to 'it'?*

Jen: Maybe, 'they' because polar bears. So, they…What do you mean by greenhouse effect?

Ming: *As in... Because they say that green house effect warms the Earth, melts the freshest snow and the ice that the polar bears live – it's something like the sun that warms the Earth and that's why... and melts the snow.*

Jen: OK, so maybe you can write it down. The greenhouse effect which… Your information report is like... is so wordy...why didn't you add some pictures? You can add some pictures... Where do you get all this information from?

Ming: *From the internet.*

Jen: Ya, you can write down in this place.

Ming: *But I can't really remember the website because they are all from different websites...*

Jen: Maybe you can go to the history of the computer so you can find what websites did you go to...

Using Assessment to Influence Students' Motivation and Self-esteem

Few educators would argue with the premise that student motivation has an important influence on learning. Learner confidence, motivation and self-esteem have been regarded by many educators to be requisites for successful learning (Dweck, 1986; Dweck, 2000; James & Gipps, 1998). These variables were showcased in a collection of teacher accounts of teaching English to speakers of other languages (Samway & McKeon, 1993) that found common threads of successful practices across 11 countries. Each of the 11 accounts reveal that great pains were taken by teachers to structure the learning environment so that children could feel good about learning and about themselves. Then there are surveys that show that high academic engagement in effective schools is due in part to much of teaching being aimed at motivating students (e.g. Pressley, 2006; Pressley, Dolezal, Raphael, Mohan, Bogner & Roehrig, 2003).

Part of the reason for implementing the IESOL projects was to address the lack of students' motivation in English lessons – IESOL students did not have confidence in speaking English and their English learning had been a frustrating and difficult experience (Education Commission, 1999; Ng, 2001, Ng & Sullivan, 2001; SCOLAR, 2003). To counter the negative impact of disengagement, one main criterion considered in the selection of IESOL curriculum materials, activities, tasks and assessment was their impact on students' motivation and self-esteem. The IESOL projects aimed at raising student engagement by promoting a positive learning environment through such opportunities as the enjoyable Shared Reading read-alouds in the early grades and the open and animated book discussions in the higher grades. There was a conscious attempt to promote and celebrate student successes, supporting language learners to grow from where they are, rather than having them frustrated by some standard unattainable for them. Teachers were admonished to give frequent, early, positive feedback that supports students' beliefs that they can do well and to provide opportunities for students' success by assigning tasks that are neither too easy nor too difficult.

Notes such as the one below appear regularly in the printed guidelines for teachers.

> Note: Attempt to get every student to participate. Do not dismiss a student's answer because it is not the one you had in mind. Try to get some insight into the student's point of view. There may not be only one 'correct' answer, or the student may have the right idea but not be able to express it adequately.

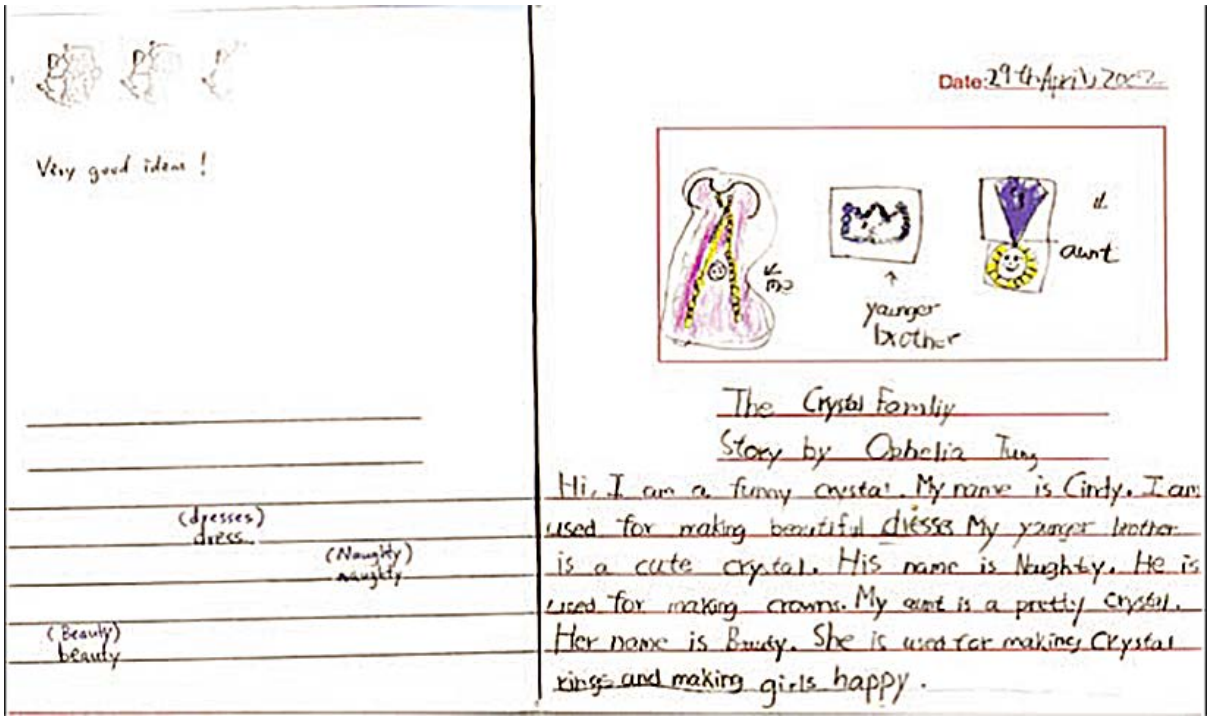
This is the basic principle that underlies all IESOL recommended teaching strategies – to get at students' thinking and to uncover the source of the confusions or differences in interpretation. To do this the teacher has to look at tasks and texts from the students' point of view, to get alongside them and to talk to them. While the above injunction relates to encouraging students' responses to teacher questions in shared reading sessions, it could apply equally to other situations such as when students contribute ideas to a piece of class writing. Resisting the urge to immediately edit students' contributions and instead, transcribing near exact matches, has the pedagogical value of helping students make the link between what he/she has said and what the teacher is writing – links that would strengthen understandings of letter-sound relationships and word identification skills. An equally important effect of the willingness to accept all students' contributions is that it encourages their willingness to contribute, and makes students feel that they are valued members of a learning community and that they have worthwhile responses and views.

The interaction in Transcript 6 above is a good example of a teacher's genuine interest in what her students have to say. The observation records noted that the true classroom dialogue occurred in a trusting and warm relationship between the teacher and the students, with even the very quiet ones offering their ideas. The discussion reflects the respect the teacher has for the students' contributions and editing suggestions. She rarely changed the students' wordings and she always sought the class's opinion before accepting an editing suggestion. Note at the end of the transcript that she did not put Rex's idea down but suggested that his contribution would be appropriate in a narrative text – thereby affirming his self-esteem but not missing the chance to draw attention to the differences between the two kinds of texts. Elaborating on Ken's editing of Rex's contribution would add to the students' development in thinking strategies, knowledge of text types and appropriate language use. One can see that such exchanges would be invaluable for fostering the students' verbal and thinking skills in authentic communication and building students' new knowledge. It shows a teacher skillfully moving between acts aimed at pedagogical and affective outcomes in the lesson.

The motivational principle is also applied to written feedback where the use of the dreaded red ink pen is discouraged. A recommended method for giving written feedback in IESOL projects in the early grades is shown in Table 2 below. It is clear from the instructions that the feedback concentrates on showing the way for improvement and giving a grade is not the main focus.

The students themselves are also encouraged when commenting on each others' written work (see next section on peer assessment), not to give grades, but instead to locate positive features of the work and features that require further attention.

Table 2. Recommended Method for Giving Written Feedback in IESOL Projects



Suggestions for marking

1. Teach your pupils to do their writing on the right-hand pages only of their exercise/jotter books, leaving the left-hand pages free for your comments and editing.
2. On the left-hand page, write the incorrect word(s)/structure(s) on the same line as the error on the right-hand page. This will provide feedback to the pupils as well as a record for you of their progress and areas that need to be re-taught or reinforced.
3. Ask the pupils if they know the correct form (often they will, once the error has been pointed out) and write it above the error on the left-hand page.
4. Teach the pupils to erase their errors in their writing and write in the correct form.

As a result of these measures, annual opinion surveys conducted in IESOL schools consistently found that nearly all school staff reported students' positive attitudes toward English classes and learning (LLELP, 2006; Ng, 2001; Ng & Sullivan, 2001). Interviews of IESOL students in Hong Kong (LLELP, 2006;) typically yielded responses such as these below.

Transcript 9

Ming (10-year-old girl):

When I was in Primary 1 to Primary 3, I don't want to go to school because I don't like English lesson and I don't want to speak to people in English. But now, I become not afraid about this, so sometimes I speak to my parents in English.

Transcript 10

Robert (11-year-old boy):

I feel very comfortable in English lesson. When I answer a question and it is correct, I feel very successful because I think before I think I am stupid, but now, not now… so I feel confident now.

Motivation is of particular importance for those who work with young children as there is research evidence indicating that failure in early literacy can limit school achievement. Most children start school enthusiastic and expecting to succeed. If the assessments show that they do not, the scenario is not very pleasant. By the end of their first year, the message to low progress children is that their initial expectations are not coming true. The continuing feedback they receive about their performance is damaging and as a result they begin to see school as punishing and degrading. Sadly, most children who start first year schooling in the bottom progress group stay in the bottom group throughout all of primary school (Clay, 2001; Juel, 1988; McGill-Frazen & Allington, 1991; Slavin, Karweit & Wasik, 1992). Better methods of teaching, assessment and the provision of informative feedback can change that debilitating cycle.

Having Students Own Their Learning

While most teachers are aware that students should be actively involved so that they own their learning, they find it difficult to put that maxim into practice. Traditionally it is the teacher who is more active, making decisions about the next steps in learning and then acting on them in the classroom. Harlen (1998) argues that it is the students who are taking the next steps and the more they are involved in the decision making process, the greater will be their understanding of how to extend their learning. Rather than being passive recipients of teachers' judgements of their work, he advocates that students take part in deciding next steps for learning including those for assessing that learning. The degree of involvement however, will have to depend on the level the

students are at, with teachers providing more support at the earlier levels and gradually withdrawing support as the students gain competence. Negotiating a good balance is a rather tricky affair and there should be constant vigilance for the guidance and support not to stifle students' involvement and ownership of their learning.

Students' Questions

An effective way of fostering student ownership of learning is to encourage them to ask questions (Commeyras, 2001). In the more flexible and relaxed IESOL project classrooms, teachers were not the only ones asking questions; the children also asked questions – even the very young beginner. The following questions were taken from transcripts of Hong Kong Year 1 IESOL children during shared readings of a story:

1. What does it say when we change the 'c' to 'r' in 'cat'?
2. If there are no more cats, why is 's' after 'cats' in the sentence 'no more cats'?
3. 'Baby' is not in the story – why is it on our spelling list?

These questions indicate engagement with different aspects of language and learning. The first two questions were connected to the stated lesson objectives; questions seeking clarification about graphophonic relationships and grammatical rules. The last question can be viewed as a challenge to the teacher's set curriculum. Although these seven-year-old IESOL children were not yet competent in English (only the first question was in English), when given the chance they seemed capable of complex and critical thinking about their English instruction in their dominant language.

In Baker's system of discourse analysis (1997), the last question can be classified as one that rocks the classroom boat. Such questions may be problematic because they deviate from the teacher set targets and plan for a particular lesson. The teacher's willingness to follow the children's lead and discuss their questions signals a shift in the balance of power and authority in the classroom with a move to a more democratic, in comparison to the usual autocratic, classroom environment. Such a radical shift demands changes in the teachers' fundamental beliefs regarding their roles as teachers and their expectations of the children and of the management of learning. Yet with the shift, the learners become genuinely involved in a process where they try to make sense of the lesson and relate it to their understanding of their worlds. In

the same way that teacher questions provide an opportunity for feedback on students' understandings, students' questions can also reveal directly to the teachers what is of great concern to the students, what is clear to them and what confuses them. Therefore using student generated questions to lead discussion could be an important step towards giving students more responsibility for their own learning.

The concept is somewhat similar to 'teachable moments'. In seizing spontaneous classroom events for instructional purposes, teachers must be ready not only to adjust but to abandon their planned activities, if those pre-planned activities are not as powerful in meeting students' needs. Making teaching decisions on the run is difficult for the teacher, but for students, knowledge gained this way is owned by them and it becomes part of their understanding – quite different from a collection of isolated facts or principles that have been memorised for an exam.

In view of the potential benefits of students' questions, a reading comprehension technique (Ogle, 1989) was introduced intentionally to stimulate students to ask questions in the upper primary levels of the IESOL projects. A key component of this technique is to have students generate their own questions for the text they are about to read, and after the reading, to check if their questions have been answered and to review the new information they have learnt from the text. This has the effect of having them set their own purposes for the reading, monitoring and assessing their own learning.

The Development of Self Assessment

When assessing, the teacher must make a judgement about a student's work (or performance) in relation to criteria of what constitutes quality. Self-assessment means that the student must learn what the teacher considers to constitute quality and to use that concept of quality to monitor continuously the quality of what is being produced.

Stated explicitly,

> "The learner has to (a) possess a concept of the standard (or goal/reference level) being aimed for, (b) compare the actual (or current) level of performance with the standard, and (c) engage in appropriate action which leads to some closure of the gap."
>
> (Sadler, 1989, p 121)

Much is still to be learnt about how best to promote self assessment, but self-corrections in oral reading, one type of self-assessment behaviour, have been found in the spontaneous behaviour of young children as they learn to read (Clay, 2001; Ng, 1988) and have been promoted as one of the effective strategies for helping children who have problems with reading (e.g. Clay, 2005; Goodman & Burke, 1973).

For the self assessment of written work, Black, Harrison, Lee, Marshall and Wiliam (2004) suggest that teachers provide students with a rubric written in student-friendly language, or develop the rubric with the class. Early trials found that our IESOL students seemed unable to use the rubrics to improve their work, even after the rubrics had been discussed and put into simple language. IESOL developers then began infusing elements of peer and self assessment in the different lesson components.

Assessment skills are nurtured from the outset in shared reading lessons when the teachers ask children after the reading what they liked about the story. A similar response is required of young readers after the independent reading of books in the Reading Centres (see Table 1). To show students the standards to aim for, books and stories used in the classroom have been selected to demonstrate concrete exemplars for quality writing. The discussion about what counts as quality text is carried through to the writing component. Drafts and completed pieces of writing generated by the whole class, by small groups of students and by individual students are discussed in attempts to discover features of quality writing. In all these discussions, the teacher models the editing process (assess, feedback and improve) to the whole class and then has students use those assessment steps when they are giving feedback to a peer about their work. Peer assessment greatly assists the development of self assessment.

Peer Assessment and Cooperative Learning

IESOL teachers have found that students, even at the emergent level in Primary 1, are better at spotting errors in other students' work than in their own work. Teachers who have used this technique with students as young as seven years old have been pleasantly surprised at how appropriate the students' comments are. The young students often communicate with one another more effectively than the teacher does. It appears that a child who is at the same level can provide feedback on how to go a little further as they also learn how to make positive comments, ask for clarification of ideas and suggest possible

alternatives for frequently repeated words. Looking at someone else's work is different from looking at your own work; it is less emotionally charged – a situation similar to a common adult practice of asking a colleague to critique one's draft.

Peer assessment can be easier developed in small groups, as talk should prove less daunting there than in a whole class situation, especially for young children. Transcript 11 below is taken from the talk of a group of four students given the task of writing a response to Joy Cowley, one of their favourite authors. The conversation is typical of IESOL small group discussions that can occur in any of the various writing stages: from the pre-writing stages of brainstorming for ideas, to the crafting of the language to express those ideas and into the editing processes.

Transcript 11

Bill:	This is too much, too much.
Sue:	But we do that in class writing.
Bill:	I know, but the class writing is not right...
Jill:	Be quick.
Bill:	But we say... our class writing is too much.
Sue:	But we need to write some...
Robert:	We want to write...
Sue:	Stop. Don't talk any more.
Bill:	We want to write some story... Write: We like her story.
Jill:	Should be 'your'...
Robert:	Your story. Your story is exciting and lovely.
Jill & Sue:	Now, we ask questions… Ask her about New Zealand... the class writing has questions
Bill:	OK, we don't write questions. The questions are too many...
Robert:	…are too many.
Bill:	We write funny things.
Jill:	Yes, yes.
Bill:	We say, Robert is a robber, ha ha.
All 4 students:	Ha ha ha ha.

This discussion proceeded through the offering of ideas, assessment of the contributions and negotiation of what to write before settling to a compromise of what to write – all part of the group's composing process. It also reveals that the budding writers were not afraid to deviate from the class writing – in fact the following excerpts from Transcript 11 issued open challenges to what the teacher wrote with contributions from the whole class.

Jill:	But we do that in class writing.
Bill:	I know, but the class writing is not right..
Jill & Sue:	Now, we ask questions.....Ask her about New Zealand... the class writing has questions
Bill:	OK, we don't write questions. The questions are too many...
Robert:	... are too many.

These children, led by Bill, decided eventually to do something different from the class writing – to inject humour into their writing. The teacher did not intervene in this particular discussion but IESOL teachers do normally circulate among the small groups to act as an observer, participate and/or assist talk and guide students in the composing task.

The research on collaborative learning, particularly in the work of Robert Slavin (Slavin, Hurley & Chamberlain, 2003), shows the powerful potential of students working together to maximise their own and each other's learning (Johnson & Johnson, 1994), provided two conditions are met. The first is that the learning situation has at its centre a requirement that all members work together to achieve a common goal, so that students are working as a group, rather than just working in a group (Baines, Blatchford & Kutnick, 2003). The second is individual accountability – that all in the group have to be 'pulling their weight' and making contributions and there can be no 'cruisers' or 'loafers' (Falloon, 2004).

The benefits of cooperative learning are many, chief of which is that it is particularly appropriate for complex tasks that learners may not be able to achieve on their own. It is known to facilitate student learning and motivation and to develop good interpersonal skills and relationships. An approach that encourages active engagement of all individuals and stimulates helping behaviours within groups must be desirable. In relation to assessment, it provides a scaffold for the development of valuable peer and self-assessment skills. It is also observed that the students providing the feedback benefit just as much as, and sometimes more than, the recipients, because they are forced to engage in understanding the rubric and to internalize the learning intentions and success criteria in the context of someone else's work. Using peer techniques also frees up teacher time for planning better instruction or working more intensively with small groups of students.

To work toward developing peer assessment skills, the learning intentions and success criteria must be accessible to the students (see preceding sections

on learning intentions) and the teacher must support the students as they learn how to help each other improve their work. The assessment should be done sensitively, whether by the teacher or student, so as not to embarrass anyone. For example, the individual piece of writing held for review could be anonymous. Students could also be asked to first focus on the positive features before giving a couple of suggestions for improvement. In line with research (see above), peer assessment should be focused on improvement, not on grading and students should not be asked to give a grade. Awarding grades involves another set of complex processes and should remain the teachers' domain.

Self Assessment

In addition to using peer assessment and feedback to scaffold the development of students' self assessment, there are writing conferences for individual children. This can be done on a one to one basis, focusing on what to aim for in quality writing. In these conferences, teachers can assure students that professional writers – those who wrote their stories and texts – ask themselves what they like/dislike about their own writing, what needs to be made clearer, the order of the ideas, and the correctness of their spelling, punctuation and grammar. A student can choose her/his best pieces and say what she/he likes most and least about those pieces. The teacher does the same, choosing key criteria, according to the piece of the work. Differences in the teacher's and the student's perception can be discussed and the student can be asked to generalise the knowledge and skills learnt to other pieces of work. If done on a rotation basis, all students in a class have opportunity for such sessions.

At the same time, students can also be taught how to use a simple set of questions to check themselves before asking a peer or the teacher to read it, whether they be about ideas, content and organisation, distinctiveness of the author's voice, word choice, or technicalities like sentence structure and grammar.

Writing checklist:

- Does this piece of writing make sense? Is it interesting to the reader? Do I communicate my ideas clearly?
- Are my ideas organised according to the purpose? Is there a smooth movement from on one idea to the next? Would organising the ideas in a different order make it better?
- Do I write in a way that expresses my ideas strongly and distinctively?

- Have I chosen words that help the reader grasp the main messages I'm trying to convey?
- Are my sentences well formed? Do I use different kinds of sentences to convey my ideas?
- Have I checked it for spelling, grammar and punctuation? Paragraphing?

The assessment processes modeled by the teacher will be picked up by some children who in turn can use them in peer assessment in small groups. At the beginning stages, the Primary 1 IESOL learners were only be able to offer simple remarks such as "I like this because it is funny/interesting' but those further down the track in Primary 5 were able to assess the appropriate use of language for a specific text type such as "You can't use 'I' in information reports because..." or to offer specific suggestions for using particular language devices to combine sentences for better effect (refer to Transcript 6). As observed by researchers such as Harlen (1998), through repeatedly using concrete examples in various settings and situations, students are helped to gradually internalise the standards, and assessment becomes a genuine part of their learning process.

CONCLUSION

While the practices promoted in this paper do not require large-scale interventions or incur as high costs as class size reduction and technologising classrooms, implementation nonetheless, is not easy. The average teacher has found that the IESOL teaching methods are a radical departure from existing practices (Ng, 2006; Ng, 1996). In addition to the usual whole class teaching, the IESOL teacher had to learn to manage a variety of groupings, and to generate more meaningful interactions between teachers and children and amongst children themselves. Instead of the usual closed questions requiring no more than yes/no answers in the classroom, they were to ask a range of questions that demanded greater resources of vocabulary and rich conversations. In place of the traditional ideal of a quiet and well-behaved class, the project teacher had to manage the active and noisy participation of children (sometimes more than 40). The application of IESOL and formative assessment principles led to a change in the interaction patterns and power relationships in the classroom, where the teacher was no longer seen as the sole purveyor of knowledge, children's voices and opinions were heard, and their contributions were encouraged and valued.

What we were asking teachers to do in these projects was to provide an experience for learning that was markedly and dramatically different from their own recollections of twenty and more years. Asian teachers who are part of bureaucratic systems with traditionally autocratic styles of management have not been nurtured in the liberal humanistic tradition of the West. Other constraints like exam pressures, large classroom sizes and inflexible time-tabling added to the difficulties inherent in developing liberal teaching styles.

Yet, those teachers who have embraced the exploration of formative assessment and feedback embedded in IESOL teaching strategies have reported on the rewards of those trials. The successful IESOL teachers have:

- organised their teaching programmes so that they could briefly but regularly observe their students' verbal and non-verbal responses.
- planned the kinds of tasks, activities and questions that they use with their students specifically to gather evidence of students' learning.
- observed who participated in the activities, how students talked in the classroom, how they interacted with other students and the kind of contributions they made to class projects.
- kept brief but regular records of those observations.
- monitored the progress of the high progress students at regular intervals and the low progress students more frequently.
- used teachable moments during class activities to provide information from their observations to show their students effective next steps for their learning.

Most of the IESOL schools are average schools with students coming from home backgrounds that many would deem unfavorable. The practices of the IESOL teachers discussed in this paper hold hope for the future – that some teachers despite the constraints, were able to make changes in their classrooms. The hope of students lies with the application of these principles. It may be that very few IESOL children will grow up as doctors or university lecturers, but most will leave the programme, at whatever achievement level they have obtained, as independent and engaged learners, enthusiastic to continue learning long after they finish school.

REFERENCES

Airasian, P. W. (2008). *Classroom assessment: Concepts and applications* (4th edn.). Boston: McGraw-Hill.

Baker, C. D. (1997). Literacy practices and classroom order. In A. Muspratt, A. Luke & P. Freebody (eds.), *Constructing critical literacies*. Creskill, NJ: Hampton Press.

Baines, E., Blatchford, P. & Kutnick, P. (2003). Changes in grouping practices over primary and secondary school. *International Journal of Educational Research* 39 (1-2), 9-34.

Bell, B. & Cowie, B. (2001). *Formative assessment and science education*. (pp 1-24). The Netherlands: Kluwer Academic Publisher.

Bennett, N. & Kell, J. (1989). *A good start? Four year olds in infant schools*. Oxford: Blackwell.

Black, P. J., & Wiliam, D. (1998a). *Inside the black box: raising standards through classroom assessment*. London, UK: King's College London School of Education.

Black, P., & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education,* 5 (1), 7-74.

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 8–21.

Brown, H. D. (1994). *Teaching by principles: an interactive approach to language pedagogy*. Englewood Cliffs, NJ: Prentice Hall Regents.

Brown, H. & Cambourne, B. (1987). *Read and retell*. Australia: Metheun Australia.

Burns, A. & Joyce, H. (1997). *Focus on speaking*. Sydney: National Center for English Language Teaching and Research.

Butler, R. (1988). Enhancing and undermining intrinsic motivation. *British Journal of Educational Psychology*, 58, 1–14.

Calfee, R. C. & Perfumo, P. (eds.). (1996). *Writing portfolios in the classroom*. Mahwah, NJ: Lawrence Erlbaum.

Carter, R., & McCarthy, M. (1995). Grammar and spoken language. *Applied Linguistics*, 16 (2), 141-158.

Cazden, C. B. (2001). *Classroom discourse: the language of teaching and learning*. Portsmouth, NH, USA: Heinemann.

Clay, M. M. (1985). *The early detection of reading difficulties*. Auckland, New Zealand: Heinemann Publishers.

Clay, M. M. (2005). *Literacy lessons designed for individuals. Part Two: Teaching Procedures*. Auckland, New Zealand: Heinemann Education.

Clay, M. M. (2001). *Change over time in children's literacy development*. Auckland, New Zealand: Heinemann Education.

Cohen, A. (1996). Developing the ability to perform speech acts. *Studies in Second Language Acquisition*, 18(2), 253- 267.

Commeyras, M. (2001). What can we learn from students' questions? *Theory Into Practice*, 34(2)101-106.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.

Davis, B. (1997). Listening for differences: An evolving conception of mathematics teaching. *Journal for Research in Mathematics Education*, 28(3), 355–376.

Dempster, F. N. (1992). Using tests to promote learning: a neglected classroom resource. *Journal of Research and Development in Education*, 25, 213-217.

Dickinson, D. K. & Tabors, P. O. (eds.). (2001). *Preparing for literacy at home and school: The critical role of language development in the preschool years*. Baltimore: Brookes.

Duckworth, E. (1981). Understanding children's understandings. Unpublished paper presented at the Ontario Institute for Studies in Education, Toronto.

Duffy, G. G. (2003). *Explaining reading: A resource for teaching concepts, skills, and strategies*. New York: The Guilford Press.

Duke, N. K. & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (eds.), *What research has to say about reading instruction* (3rd ed.; pp. 205-242). Newark, DE: International Reading Association.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.

Dweck, C. S. (2000). *Self-theories: their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.

Education Commission (1999). *Review of education system: Framework for education reform*. Hong Kong: Education Commission, Hong Kong Special Administrative Region.

Falloon, G. (2004). Cooperative groupings as an organisational system for classroom computer-use. *Computers in NZ Schools*, 16(1), 32-35.

Fuchs, L. S. (1993). Enhancing instructional programming and student achievement with curriculum-based measurement. In J. Kramer (ed.), *Curriculum-based measurement*. (pp. 65-103). Lincoln, NE: Buros Institute of Mental Measurements.

Fuchs, L. S. & Fuchs, D. (1986). Effects of systematic formative evaluation: a meta-analysis. *Exceptional Children*, 53, 199-208.

Gipps, C. (1994). *Beyond testing: towards a theory of educational assessment*. London: Falmer Press.

Goodman, K. S. & Burke, C. L. (1973, April). Theoretically based studies of patterns of miscues in oral reading performance (project No. 9-03375). Washington, DC: US Office of Education.

Harlen, W. (1998). Classroom assessment: A dimension of purposes and procedures. In K. Carr (ed.), *SAME papers* (pp. 75–87). Hamilton, New Zealand: Centre for Science, Mathematics and Technology Educational Research, University of Waikato.

Harlen, W. (1998). *Classroom assessment: A dimension of purposes and procedures*. Paper presented at the Annual Conference of the NZARE, Dunedin, December.

Hattie, J. (1999). *Influences on student learning*. Inaugural Lecture. Professor of Education, University of Auckland, New Zealand.

Hattie, J., Biggs, J. & Purdie, N. (1996). Effects of learning skills interventions on student learning: a meta-analysis. *Review of Educational Research*, 66, 99-136.

Hiebert, E. H. & Raphael, T. E. (1998). *Early literacy instruction*. Florida: Holt, Rinehart & Winston.

Hoffman, J. V., McCarthey, S. J., Elliot, B., Bayles, D. L., Price, D. P., Ferree, A. & Abbott, J. A. (1998). The literature-based basals in first grade classrooms: Savior, Satan, or same-old, same-old? *Reading Research Quarterly*, 33(2), 168-197

Hogan, K. & Pressley, M. (eds). (1997). *Scaffolding student learning: instructional approaches and issues*. Cambridge, MA: Brookline.

James, M. & Gipps, C. (1998). Broadening the basis of assessment to prevent the narrowing of learning. *The Curriculum Journal*, 9(3), 285-297.

Johnson, D. W. & Johnson, R. T. (1994). Professional development in CL: short-term popularity vs. long-term effectiveness. *Cooperative Learning*, 14, 52-54.

Johnston, (1992); Glazer & Brown, (1993); Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 59, 243-245.

Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.

LLELP (2006). *Report on the 'Learning to Learn' English Language Programme (LLELP) in Hong Kong Primary Schools (2004-2006)*. Final Report submitted to the Quality Education Fund. September, 2006.

McGill-Franzen, A. & Allington, R. L. (1991). Every child's right: literacy. *The Reading Teacher*, 45(2), 86-90.

Marshall, B. & Wiliam, D. (2006). *English inside the black box: assessment for learning in the English classroom.* London, UK: NFER-Nelson.

Martin, P. & Abdullah, K. (2003). English language teaching in Brunei Darussalam: Continuity and change. In W. K. Ho & R. Y. L. Wong (eds.), *English language teaching in East Asia today*. (pp 95-110). Singapore: Times Academic Press.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (National Institute of Health Pub. No. 00-4754). Washington, DC: National Institute of Child Health and Human Development.

Ng, S. M. (ed). (1988). *Research into children's language and reading development*. Singapore: Institute of Education.

Ng, S. M. (1996). Innovation, survival and processes of change in the bilingual classroom in Brunei Darussalam. *Journal of Multilingual and Multicultural Development*, 17(2-4), 149-162.

Ng, S. M. (2001). The Brunei Reading and Language Acquisition project. *International Journal of Educational Research*. 35(2), 169-179.

Ng, S. M. (2006). Literacy and diversity: challenges of change. In A. McKeough, L. M. Phillips, V. Timmons & J. L. Lupart (eds.), *Understanding literacy development: A global perspective*. (pp. 153-184). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. Chapter 9.

Ng, S. M. & Sullivan, C. (2001). The Singapore Reading and English Acquisition programme. *International Journal of Educational Research*, 35(2), 157-167.

Ng, S. M. & Sullivan, C. (2008). *Primary grade children learning English as speakers of other languages in Hong Kong*. Hong Kong: Hong Kong Council of Early Childhood and Services.

Ogle, D. M. (1989). The know, want to know, learn strategy. In K. D. Muth (ed.), *Children's comprehension of text*. New York, Del: International Reading Association.

Pressley, M. (2006). *Reading instruction that works: The case for balanced teaching* (3rd ed.). New York: The Guilford Press.

Pressley, M., Dolezal, S. E., Raphael, L. M., Mohan, L., Bogner, K. & Roehrig, A. D. (2003). *Motivating primary-grade students*. New York: Guilford.

Pressley, M., Rankin, J. & Yokoi, L. (1996). A survey of instructional practices of primary teachers nominated as effective in promoting literacy. *The Elementary School Journal*, 96(4), 363-383.

Pryor, J. & Torrance, H. (1996). Teacher-pupil interaction in formative assessment: assessing the work or protecting the child? *The Curriculum Journal*, 7, 205-226.

Rivers, W. M. (1981). *Teaching foreign language skills (2nd ed.)*. Chicago: University of Chicago Press.

Roskos, K. A., Tabors, P. O. & Lenhart, L. A. (2004). *Oral language and early literacy in preschool*. Newark, DE: International Reading Association.

Samway, K. D. & McKeon, D. (eds.). (1993). *Common threads of practice*. USA: Teachers of English to Speakers of other Languages, Inc.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5(1), 77-84.

SCOLAR (Standing Committee on Language and Research). (2003). *Action plan to raise language standards in Hong Kong: consultation document*, January 2003. Hong Kong Special Administrative Region Government.

Siobhan Leahy, Christine Lyon, Marnie Thompson and Dylan Wiliam (2005). Classroom Assessment: Minute by Minute, Day by Day. *Educational Leadership*, 63(3), 19-24

Slavin, R. E., Hurley, E. A. & Chamberlain, A. M. (2003). Cooperative learning and achievement. In W. M. Reynolds & G. J. Miller (Eds.), *Handbook of Psychology* volume 7: educational psychology (pp. 177-198). Hoboken, NJ: Wiley.

Slavin, R. E., Karweit, N. L. & Wasik, B. A. (1993). Preventing early school failure: What works? *Educational Leadership*, Dec. 1992/Jan. 1993, 10-17.

Smith, A. B. (1999). Joint attention: Learning to “know their minds”. *Early Childhood Folio*, 4, 13-15.

Stiggins, R. J. (2005). From formative assessment to assessment for learning: a path to success in standards-based schools. *Phi Delta Kappa*, 87(4), 324 – 328

Wiliam, D., Lee, C., Harrison, C. & Black, P. J. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education: Principles, Policy and Practice*. 11(1), 49-65.

Wiliam, D. (2007). Keeping learning on track: formative assessment and the regulation of learning. In F. K. Lester Jr. (ed.), *Second handbook of mathematics teaching and learning*. Greenwich, CT: Information Age Publishing.

Wylie, C. & Thompson, J. (2003). The long-term contribution of early childhood education to children’s performance – evidence from New Zealand. *International Journal of Early Years Education*, 11(1), 67-78.

Author’s e-mail: dawn@igrin.co.nz

Using Vocabulary Profiling Assessment Software to Promote Independent Process Writing

Arthur McNeill
The Hong Kong University of Science and Technology

Introduction

The extent to which assessment tools can make a positive contribution to student learning is likely to remain a controversial topic in education in spite of the recent interest in exploring the learning potential of some approaches to assessment and the pedagogical uses of assessment data. In the minds of many educators, assessment and teaching are activities that need to be distinguished clearly from each other. For example, in examination-dominated education systems, the widespread use of examination papers as teaching materials has been criticized on the grounds that the practice promotes test-taking skills at the expense of developing students' understanding of the subject matter. However, the current interest in Assessment *as* Learning has encouraged new perspectives in language assessment, including a reappraisal of some established language tests. This paper explores the potential of a customized version of the Lexical Frequency Profile (LFP) assessment software (Laufer & Nation 1995) as a learning tool.

The original purpose of LFP was to measure the lexical richness of texts. The classic version of the software provides a profile of a text's lexical content by grouping the vocabulary items into four categories: (a) words from the first 1000 frequency level, (b) words from the second 1000 frequency level, (c) words included in the Academic Wordlist (AWL) (Coxhead 2000), and (d) words not included in the previous three categories, i.e. "off-list". The words included in the first and second 1000 frequency levels are basically the items in the General Service List (GSL) (West 1953). A later version of LFP divides a text's lexis into twenty levels based on the British National Corpus (BNC)

(Leech et al. 2002), plus “off-list”. LFP has typically been used by teachers to assess texts written by students. It has also been used to scrutinize texts when considering their suitability for use in examinations and textbooks. The LFP performs a quick analysis of a text’s lexical content by showing the frequency level of the words used and providing some basic descriptive statistics about the text based on word frequency. The profile shows the extent to which a text consists of low frequency and academic words. Although LFP has proved to be a popular tool among researchers, other approaches to measuring lexical richness have been adopted (e.g. Meara 2001, Meara & Bell 2001, Miralpeix & Celaya 2002, Bell 2003, Šišková 2012).

Lexical richness has been identified as one of the most important features of second language writing and can determine the quality of a text (Koda 1993). Many advanced learners of English, in particular undergraduate and postgraduate students studying through the medium of English, also benefit from LFP and use it to check the vocabulary content of course assignments and academic papers before submitting them. Unfortunately, however, this use of the software for checking drafts is currently less accessible to intermediate level learners because the wordlists used in the original LFP are based on the frequency of occurrence of words within the language in general, rather than a particular L2 vocabulary syllabus where the prescribed lexical content may vary considerably from frequency lists such as GSL and BNC, which comprise the databases of LFP.

This paper describes an attempt to produce a version of LFP that is of particular relevance to learners in Hong Kong, by replacing the frequency-based wordlists (GSL and BNC) with the four wordlists (Key Stages 1 – 4) that were developed for Hong Kong schools (Curriculum Development Council 2009, 2012; McNeill 2011). The paper also shows how institutional wordlists, such as the prescribed words of a university English course can be included within a customized LFP. It is argued that students will benefit more from using LFP as a learning tool if its lexical database corresponds to the vocabulary curriculum they are expected to study. In its original form, LFP is driven by wordlists that consist of lexical items selected because of their frequency of occurrence in English generally.

Hong Kong English Wordlists and a Customized LFP

Increasing evidence has emerged to suggest that many Hong Kong students leave school with an inadequate English vocabulary. For example, a study conducted with first-year undergraduates at the Chinese University of Hong Kong concluded that the majority of school leavers entering university know fewer than 3000 English words (Chui 2005). International research into the English language proficiency of students studying degree courses through the medium of English has suggested that a vocabulary size of 5000 words is necessary for students to cope with the demands of reading academic texts in English (Laufer 1989, 1992). In order to promote higher English vocabulary targets for Hong Kong school leavers, the Education Bureau commissioned a study of the vocabulary needs of Hong Kong primary and secondary students, with a view to developing an English vocabulary curriculum for primary and secondary education. The first stage of this curriculum project involved the creation of four wordlists, consisting of the vocabulary items that students could be expected to know at different stages of their education.

It was agreed at the outset of the curriculum project that a sensible starting point in selecting words for the lists would be corpus data showing the frequency of occurrence of English words. It was assumed that frequency of occurrence would provide a useful provisional ranking of the potential words. This initial ranking would then be subjected to scrutiny by a number of stakeholders, including teachers familiar with the student population for whom each list was intended, before arriving at a final selection of target lexis. However, when analyzing the patterns of response from teachers, it soon became apparent that the number of words rejected was far greater than expected (McNeill 2011). The use of the BNC corpus was, at first sight, attractive to the research team because of its spoken component. The Hong Kong word lists were intended to reflect vocabulary used in spoken as well as written English. However, because BNC is representative of contemporary usage within UK, a large number of the words were judged to be inappropriate in the S.E. Asian context. Many of the items were considered to be restricted to users of colloquial British English or relied on familiarity with contemporary British culture.

Examples of high-frequency BNC words judged to be restricted to the UK context were *bobby, dodgy, dole, heck, lass, plonk, posh, innit, shilling, tuppence, kiddy, owt.* BNC words considered too colloquial for a L2 school curriculum included *baffle, boo, blob, buck, clobber, fiddle, eve, grumble, grotty, hassle, fuss, tumble, wobble, potter, dodgy, swap, poke, cop, buzz, whack.* The high number of rejected items from the 4000 level of BNC illustrated the enormous difference that appears to exist between the high-frequency vocabulary in the English used by British people for everyday communication among themselves and the kind of vocabulary which ESL learners might be expected to learn for the purpose of education and employment in their own countries.

Probably because of their very high frequency within English, the words in GSL presented fewer cultural barriers to being accepted into the Hong Kong lists than the comparatively 'newer' vocabulary of BNC. In spite of the development of enormous corpora in recent years, there remains little controversy over the words that constitute the first 2000 words of English. However, when GSL words were rejected in the teacher decision-making tasks of KS1 and KS2, it was usually because the items were judged to be inappropriate for young learners. The majority of the items were eventually included in the lists for KS3 and KS4. Therefore, unlike the original LFP, the customized Hong Kong version does not assume that the first 2000 words of English, by themselves, should be the starting point of an L2 vocabulary curriculum for school children.

Customizing the LFP Software

At first sight, the task of replacing one word list with another in the LFP software looks straightforward. However, one of the challenges for LFP, particularly when serving as a learning tool, is to assign a text's lexis to the prescribed lists (in our case the vocabulary of the four Key Stages) as accurately as possible. Basically, the software needs to anticipate and accommodate all morphological variations of the words in each list. While the vocabulary target for Hong Kong students by the end of secondary education is 5000 English

words, the number of word forms which students actually know is much larger than 5000. The lists contain 5000 word families (Bauer & Nation 1993), which means all the derived forms of each headword need to be included in the database so that the LFP will recognize them and place them in the correct category when profiling students' texts. If a word form is not recognized, LFP places it in the "Off-List Words" category. Since all 570 of the AWL words are included across the four KS lists, there is no separate AWL list. The basic LFP report format is shown in Figure 1.

	Tokens	**Percentage**
KS1 Words		
KS2 Words		
KS3 Words		
KS4 Words		
KS1+KS2+KS3+KS4		
Off-List Words		
Total		

Figure 1. Categories used to profile lexis based on Hong Kong's local wordlists

Post-Secondary Customization

The main appeal of producing a customized version of an assessment software program such as LFP is the opportunity to focus learners' attention on the language they are expected to know according to their local curriculum, rather than being judged according to knowledge of words that may not be relevant to their own situation and needs. Preparing a version of LFP based on the Hong Kong primary and secondary school vocabulary lists can make the tool available and accessible to the thousands of students who study English up to Diploma of Secondary Education (DSE) level. However, for students who proceed to post-secondary education, it is possible to include additional lists, such as the prescribed words of particular English courses and lists of discipline-specific lexis for subjects such as business, engineering, science, medicine, law, architecture, etc.

This section provides an illustration of an additional wordlist based on a first-year university English course. The course, “English for University Studies”, is taken by the majority of undergraduates at the Hong Kong University of Science and Technology (HKUST) and focuses on improving overall English proficiency and developing academic literacy. The course has a prescribed wordlist, derived mainly from the course reading materials, which students are expected to master.

A wordlist consisting of the prescribed words for “English for University Studies” (e-Core Words) was added to the Hong Kong version of LFP and students were encouraged to make use of the tool by submitting drafts of their course assignments into LFP and making revisions to their texts in the light of the lexical profile. Many university writing courses adopt a process approach, whereby a student submits a draft of a writing assignment to the instructor, who then provides formative feedback on the written text so that the student can re-draft and improve it. The pedagogical advantages of process writing are undisputed and it has long been established that re-drafting a text in the light of feedback results in improved writing quality (e.g. Raimes 1983, White & Arndt 1991). However, as every writing instructor knows, giving feedback on drafts is enormously time-consuming on the instructor’s part. In some respects, LFP allows students to learn to write in a process mode by providing them with an individual lexical profile of a draft. Students can then decide whether they need to make changes to the content in order to demonstrate that they are able to use the words that have been prescribed for study on a particular course.

Experience of using the LFP with course-prescribed words at HKUST reveals that students typically produce first drafts with an over-reliance on high-frequency items, then make a serious effort to monitor their lexical choices and improve the sophistication of the vocabulary in the light of their LFP profile. Fears that students might insert and substitute words in a reckless manner just to improve their profile were largely unfounded. However, to test the system, a text was prepared which deliberately used as many of the course-prescribed words as possible. The results are shown in Figure 2.

Vocab-Profile:

	Tokens	Percent
KS1 Words:	58	40.28%
KS2 Words:	10	6.94%
KS3 Words:	7	4.86%
KS4 Words:	13	9.03%
KS1 + KS2 + KS3 + KS4	(88)	(61.11%)
eCore Words:	44	30.56%
Off-List Words:	12	8.33%
Total:	144	100%

intuition is the antithesis of rational thinking and can be radically differentiated from coherent models of cognitive deduction intuitive thinkers may be individualistic spontaneous self-reliant sophisticated and prone to self-indulgence paradoxically however they are not immune to criteria such as randomness and consensus their deviation from socioeconomic and ethnic trends can imply a transitional immersion in dogma yet we should be reluctant to generalize about intuition on the basis of anecdotes and superficial brainstorming although we may sympathize with and even idolize the hallmarks of unique engagement in diverse ventures texts such as this assert that lexical richness is a relevant variable of tech-savvy writing for resourceful learners however some instructors may be skeptical about its stark contribution to cohesive anything

Figure 2. LFP analysis of a text written to display mastery of a prescribed university wordlist

Although the above text represents an extreme lexically-oriented approach to writing in which the author deliberately included words from the course prescribed wordlist, the resulting text suggests that high-quality academic writing is heavily dependent on lexical choice. In the above example, more than 30% of the words were taken from the course-prescribed wordlist! Nevertheless,

the practice of encouraging students to reflect on the vocabulary content of their draft assignments and allowing them the opportunity to re-draft their texts through lexical substitution is clearly worthwhile educationally.

The use of different colours to represent the words of the various lists provides students with an immediate visual analysis of their text. Once students learn which colours are highly valued in terms of the profiling, they generally find the lexical editing of their work engaging and enjoyable.

Conclusion

This paper has discussed the potential of the LFP software as a learning tool to promote second language writing. As explained above, LFP was developed as an assessment instrument to measure the lexical richness of written texts and has, in fact, been used by learners, usually advanced level adult learners, to check drafts as part of a revision process. Therefore, LFP has always been an assessment tool that can help language learners to improve the quality of their written texts. However, as explained in the paper, LFP's potential as a learning tool can be greatly enhanced when used in contexts where there is a local vocabulary curriculum whose wordlists can replace the frequency-based lists of the original LFP. Such customization of the tool can allow it to be used by students of all proficiency levels, including young learners, since the vocabulary databases used correspond completely to the students' local English curriculum.

References

Bauer, L. and Nation, I.S.P (1993). Word Families. *International Journal of Lexicography*, Vol. 6, No. 3, 253-279.

Bell, H. (2003). Using Frequency Lists to Assess L2 Texts. University of Wales Swansea: Unpublished Thesis.

Chui, S.Y. (2005). *An Investigation of the English Vocabulary Knowledge of University Students in Hong Kong*. Unpublished M.Phil. thesis. Hong Kong: The Chinese University of Hong Kong.

Coxhead, A. (2000). A new Academic Word List. *TESOL Quarterly*, Vol. 34, No. 2, 213-238.

Curriculum Development Council (2009). *Enhancing English Vocabulary Learning and Teaching at Primary Level*. Hong Kong: Government Logistics Department.

Curriculum Development Council (2012). *Enhancing English Vocabulary Learning and Teaching at Secondary Level*. Hong Kong: Government Logistics Department.

Koda, K. (1993). Task-induced variability in FL compositions: Language-specific perspectives. *Foreign Language Annals*, 26 (3), 332-346.

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In Lauren, C. and Nordman, M. (eds.) *Special Language: From Humans Thinking to Thinking Machines*. Clevedon: Multilingual Matters.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In Arnaud, P.J.L. and Béjoint, H. (eds.) *Vocabulary and Applied Linguistics*. London: Macmillan 126-132.

Laufer, B. and Nation, I.S.P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, Vol. 16, No. 3, 307-322.

Leech, G., Rayson, P., and Wilson, A. (2002). Companion website for word frequencies in written and spoken English: based on the British National Corpus. Available at http://ucrel.lancs.ac.uk/bncfreq

McNeill, A. (2011). From corpus to language curriculum: a data-based or data-driven exercise? *Proceedings of EDULEARN11*, International Association for Technology, Education and Development. ISBN: 978-84-615-0441-1.

Meara, P. (2001). PLEX v1.1. University of Wales Swansea. Available at http://www.swan.ac.uk/cals/calsres

Meara, P. and Bell, H. (2001). "P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts", in *Prospect*, 16, 3: 323-337.

Miralpeix, I. and Celaya, M.L. (2002). "The Use of P_Lex to Assess Lexical Richness in Compositions Written by Learners of English as an L3", in I. Palacios (ed.), *Actas del XXVI Congreso de AEDEA*. Santiago de Compostela/ University of Santiago de Compostela Press, 399- 406.

Raimes, A. (1983). *Techniques in Teaching Writing*. Oxford: Oxford University Press.

Šišková, Z. (2012). Lexical richness in EFL students' narratives, *University of Reading Language Studies Working Papers*, Vol. 4, 26-36.

West, M. (1953). *A General Service List of English Words*. London: Longman.

White, R. and Arndt, V. (1991). *Process Writing*. London: Longman.

Author's e-mail:　lcamcneill@ust.hk

*Stop Being Composition Slaves!

Icy Lee
The Chinese University of Hong Kong

Do English teachers in Hong Kong want to be “composition slaves” (Hairston 1986)? I am sure the answer is “no”. In reality, however, many teachers find themselves slaving over student compositions, burning the midnight oil to mark student writing. In my article ‘A new look at an old problem: How teachers can liberate themselves from the drudgery of marking student writing’ published in *Prospect: An Australian Journal of Teaching/Teachers of English to Speakers of Other Languages (TESOL)* (Lee 2009), I challenge teachers’ existing assumptions about feedback and outline ten perspectives to help teachers re-examine their feedback practices from a new vantage point. They are summarized as follows:

1. Students have every right to write in the way they do

Spandel (2005) reminds us of the importance of respecting the rights of student writers, including their rights to go off-topic and to write badly. To emancipate themselves from the drudgery of marking student writing, the first thing teachers can do is to look at student writing using a new lens, treating it as an artefact created and owned by the student writer (rather than the teacher) and showing it greater respect. Rather than labeling it as poor writing or covering it with red ink, teachers could find out what student writing tells them about students’ personal perspectives, what they are able (or not able) to do in writing, their world views, their idiosyncrasies, and above all, who they are as persons and as writers (Murray 1985). Rather than speaking for their students by re-writing student texts (and changing their meaning) on the assumption that they are extraordinarily devoted and hardworking teachers, they could talk to students to find out what exactly they want to say. They could also avoid derogatory commentary like ‘you write badly’, which are unlikely to change

students' ability to write. If students have every right to write in the way they do, then in the first place teachers have to learn to respect this very right.

2. Why 100 percent accuracy?

As teachers respond to every single error in student writing, they are sending a message that they expect students to produce error-free writing. Errors, however, are a natural part of language learning. Even teachers' best efforts at error correction do not produce 100 percent accurate writing (Ferris 2008). So why should teachers act as 'error hunters' (Hairston 1986: 122) and impose an unattainable goal on students and themselves? More importantly, when teachers keep their attention on surface-level concerns, they lose sight of other important dimensions of writing, such as ideas, rhetorical features, style, and voice. Writing is not a mere vehicle for language practice. To rid themselves of 'tunnel-vision', teachers have to bear in mind that there is much more to good writing than grammatical accuracy. Sometimes student writing can be accurate but unnatural because of its non-idiomatic expressions, or it does not read fluently, or it fails to meet the readers' expectations in terms of rhetorical conventions. As teachers change their attitude to surface-level errors, they also broaden their perspectives and let other important aspects of student writing inform their feedback.

3. Tell them and they will forget

While the argument in favor of marking all written errors emanates from the felt need to inform students of the mistakes they have made, teachers have to understand that for the majority of students, their attitude to teacher feedback is one of 'tell-them-and-they-will-forget'. This is particularly so when they have a large number of errors to attend to. There is feedback research which shows that selective error feedback is preferable to comprehensive error feedback (Ferris 2003), with the former being more focused (also referred to as focused corrective feedback – see Ellis 2009) less threatening and more manageable for L2 learners. As teachers work hard at locating, categorising, and providing correct answers for students, or even writing almost the entire paragraph or the

whole piece for students, they are usurping students' right to learn for themselves, also depriving them of the opportunity to develop self-editing skills. Learning in the writing classroom can take place only if teachers are willing to let go of their teacher-dominated role as an editor, let students self-edit or peer-edit, and train them to do so. Tell them but they will forget – involve them, and they will understand. Teachers have to allow students to take greater control of their learning.

4. Change the rule of the game

Currently, the rule of the game is that once students have finished their writing that is the end of their responsibility. They pass their papers to teachers and wash their hands of them. Then it is the teachers' turn to show their effort by responding to the papers laboriously. Teachers become key players of the game. But why is this the case? Who should be doing the error correction and editing? And who should be learning? It is the students not the teachers. To salvage the situation, teachers need to change the rule of the game. When students finish writing, their responsibility is not over. They should be held accountable for their own writing. For example, they should be given opportunities to tell teachers what they want to get from teacher feedback; they can help each other review their writing and improve it; they can also set themselves some short-term and long-term goals and monitor their own writing development. More importantly, they should be given opportunities to act on teacher feedback to bridge their gaps in writing. If effective learning is to take place, students have to be the key players in the game.

5. No more double standard – stop being so harsh to developing writers

Writing is a painstaking process. Whether it is L1 or L2 writing, in real life it takes time for ideas to incubate and for students to get started, to draft and redraft, and to polish their writing. Real-life writing is never a simple process. Teachers know it. While teachers themselves may write with full awareness of their constraints as writers, in writing classrooms they may apply a different

standard to their students. They may not give students enough time to gather ideas, to let ideas develop, to draft and redraft, and to edit their writing. In addition, they may not explicitly teach students how to do so. Worse still, they may put students in exam-like situations where they are given a topic and have to write within a certain time limit. If even teachers themselves reckon that writing is a difficult process, why are they so harsh to their students, who are but developing writers? The fact that teachers have to spend so much time reading less than satisfactory student writing is partly because they have not given students sufficient time to develop a good piece of work.

6. More is not better

Teachers exhaust themselves through marking student writing perhaps because they think more is better – more red marks on the page, more errors corrected, more error codes, more comments, as well as more compositions from students. In some contexts (like EFL school contexts), it is not uncommon for teachers to collect a great number of compositions from students in one academic year (e.g. ranging from 10 to 14), all treated as terminal drafts performed in a relatively exam-like environment (i.e. timed and with minimal help from the teacher). Teachers may also think that the more they assess, the better teachers they are. However, more error corrections, more codes, and more compositions do not necessarily make students better writers. Instead of teachers marking all student texts, isn't it the case that some of them can be read by students for peer or self-assessment? Instead of enforcing more exam practice, can't teachers adopt a process approach, assign fewer writing tasks but give students more time to produce a piece of work through multiple drafting? Instead of writing lots of comments on student papers, would it be possible for teachers to give fewer but more focused comments, especially those they believe students are able to act on and benefit from? As teachers start to think that more is not necessarily better, they will work smarter.

7. Feedback as well as feedforward

Feedback should also serve the purpose of feedforward (Carless et al. 2006) – that is, students using feedback information to improve their writing. Given this idea, teachers' focus should not be on feedback per se (i.e., getting their job of delivering feedback done) but more importantly on how feedback can be utilised to help students improve their writing. A paper filled with red ink suggests that there are probably far too many things for a student to attend to, while feedback that addresses the major problems in student papers can help students focus on specific areas, and this is likely to be more manageable for students. So why mark student writing so feverishly?

8. Don't be a coroner – be a diagnostician

As teachers toil away at student writing, they tend to play the role of a judge – specifically that of a coroner, declaring the 'death' of student writing and commenting on it retrospectively. If feedback is to fully realise its potential as feedforward, teachers have to play the role of a diagnostician, helping students identify the most critical problems in their writing (Murray 1985). Rather than let student compositions overpower and overwhelm them, as diagnosticians teachers take control, work on student papers, and tease out the most serious problems for students to act on.

9. Teach not just test

Why do teachers think that students don't write well? Why does student writing exhibit multiple problems? It is because students are under-prepared. Teachers have to spend so much time marking student writing because there is a missing link between teaching and assessment. Apart from diagnosing student writing, teachers have to be coaches. They need to put a greater emphasis on teaching, and specifically teaching that informs assessment, and spend less time on testing. If teachers teach what they assess and assess what they teach, then marking student writing will be a much easier job to do.

10. The hidden agenda of feedback

As institutions or schools set up expectations for teachers to mark student writing in certain ways, teachers have to comply because they are held accountable. Failure to do so may result in unsatisfactory evaluation by students and even negative appraisal by school administrators (see Lee 2008). How much do teachers want, through their meticulous feedback, to show their students, parents, colleagues and superiors that they are hardworking, dedicated and competent teachers? Given such a hidden agenda of feedback, if teachers are to fully harness the potential of feedback and emancipate themselves from the drudgery of feedback, they probably need to undertake a feedback revolution, which will require them to negotiate with school administrators, talk to colleagues, students and parents, and initiate a whole-school approach to change.

If teachers can challenge some of the taken-for-granted assumptions about feedback outlined in the ten perspectives above, and if they can think outside the box and explore alternatives together with their colleagues, responding to student writing will become a more rewarding and productive experience.

*The original article is published in *Prospect: An Australian Journal of Teaching/Teachers of English to Speakers of Other Languages (TESOL)*.

References

Carless, D., Joughin, G., Liu, N. F., & Associates. (2006). *How assessment supports learning: Learning-oriented assessment in action*. Hong Kong: Hong Kong University Press.

Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal, 63*(2), 97–107.

Ferris, D. R. (2003). *Response to student writing: Implications for second language students.* Mahwah, NJ: Lawrence Erlbaum Associates.

Ferris, D. R. (2008). Students must learn to correct all their writing errors. In J. Reid (Ed.), *Writing myths: Applying second language research to classroom teaching* (pp. 90–111). Ann Arbor, MI: the University of Michigan Press.

Hairston, M. (1986). On not being a composition slave. In C. W. Bridges (Ed.), *Training the new teacher of college composition* (pp. 117–124). Urbana, Il: NCTE.

Lee, I. (2009). A new look at an old problem: How teachers can liberate themselves from the drudgery of marking student writing. Prospect: *An Australian Journal of Teaching/Teachers of English to Speakers of Other Languages (TESOL), 24*(2), 34-41.

Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second Language Writing, 17*(2), 69–85.

Murray, D. (1985). *A writer teaches writing* (2nd ed.). Boston, MA: Houghton Mifflin Company.

Spandel, V. (2005). *The nine rights of every writer: A guide for teachers*. Portsmouth, NH: Heinneman.

Author's email: icylee@cuhk.edu.hk

Writing ePlatform: A Corpus-based Resource to Support Learning and Assessment in Writing

Sean McMinn
The Hong Kong University of Science and Technology
Flora Fung Yin Leung
Education Bureau, Hong Kong

Abstract

Based on the findings of a recent study conducted by Education Bureau (EDB) and Hong Kong Examinations and Assessment Authority (HKEAA) on using assessment data to enhance the learning and teaching of Speaking and Writing, an online corpus-based resource tool that provides instant informing feedback has been developed to assist Key Stage Three students writing. This paper discusses the rationale and development of the tool, the Writing ePlatform.

Introduction

In a study conducted by Education Bureau and Hong Kong Examinations and Assessment Authority aiming to identify possible problems in the learning and teaching of Speaking and Writing in the English Language at Secondary Three, Cheung and Leung (2012) observed that students' performances in speaking and writing were highly correlated. In the sub-construct of "vocabulary and language patterns", they also saw evidence of transfer between spoken and written. However, transfer from writing to speaking seems more likely. These imply that it is worthwhile to address students' difficulties in writing, which would then benefit students' speaking. A qualitative analysis of the speaking and writing performance of the same students shows that different error types were observed among students of varied ability. It is also found that "grammar accuracy index" is one of the strongest predictors of students' writing

performance. As Harmer (1983: p.35) pointed out, “An error is the result of incorrect rule learning; language has been stored in the brain incorrectly”. Hence, students need to learn about common error types, how they arise, and how to avoid them. The findings moreover indicate that students’ self-ratings bear little relationship to the ratings which their performances received, suggesting that students do not clearly know where they are in their learning process. Therefore, it would also be useful to step up e-learning resources that can give students instant feedback, and provide them with advice on different learning strategies and metacognitive skills, with the intent to help them become more autonomous learners.

When findings from the study and recommendations were presented to the participating schools, most teachers shared the vision of developing a computer corpus-based identification and classification system for students’ errors. It was decided that developing an online system to enhance writing skills was desirable given the ready availability of information technology (IT) facilities in schools. The Center for Language Education (CLE) at the Hong Kong University of Science and Technology (HKUST) was commissioned to design, produce and test an online system as the first stage of developing online tools that could assist or enhance the teaching and learning of spoken and written English. This paper discusses the background, development and piloting of the system.

The Aim and Objectives

It should be noted that the creation of this interactive resource is part of a larger research project, which has its own aim and objectives. Because of space, this paper focuses only on the construction of the ePlatform, which provides students with interactive feedback that is geared towards their ability level and given immediately after they have submitted their writing to the online system. The feedback focuses on common writing problems, especially for low achievers. One distinctive feature of this system is that it could be incorporated into process writing.

The platform aims to: (1) assist with constructing the lexico-grammatical and discoursal/rhetorical knowledge of the target language and the skills required to access and apply that language; (2) encourage reflection and metacognition, where students are encouraged in independent learning and self-confidence; and (3) develop 'cognitive apprenticeship', where coaching and modeling occur, and where scaffolding is provided to support language learning. Two key features of the system are given as follows:

(1) **the eLab**: an interactive web-based platform where students can submit writing and receive instant feedback and suggestions for how to improve their written English. The eLab is designed to be flexible according to student proficiency levels and needs. Additionally, the eTutor, student-oriented concordancers, vocabulary tools, and a vocabulary profile that refers to Key Stage 1, 2 & 3 vocabulary lists are accessible features of the eLab that assist students with developing their writing and provide teachers with diagnostic information for face-to-face lessons; and

(2) **the eTutor**: a web-based portal that contains English learning materials and interactive learning objects based on common errors made by Hong Kong students. The eTutor provides guidance on how to address language learning issues relating to vocabulary, grammar and mechanics in response to common errors made by students in their writing.

To the best of our knowledge, no similar corpus-based system has been developed in the market for the Hong Kong, or Chinese, context and student age range.

Theoretical Underpinnings of the ePlatform

The Writing ePlatform is a corpus-based resource. Corpora have become research tools for exploring the lexical features of learners' language production. Meyer (2002) and Nesselhauf (2005) create corpora from student-written essays collected in classes and Kuo (2005) explores employing test takers' writing passages to create a learner corpus. Coniam (1997) investigates the extent to

which it is possible to produce tests using corpus word frequency data with currently available computer technology. Coniam (1999) further identifies word frequency as an indicator of language proficiency in the written English of Grade 13 learners of English in Hong Kong. The study extends Laufer and Nation's (1995) work involving the Lexical Frequency Profile (LFP), in which vocabulary profiles were extracted from student writing on the basis of the frequency of the words. With regard to vocabulary, corpus data analysis and corpus tools have been causing a significant refocusing of views concerning the nature of English language assessment and the making of important pedagogical decisions.

The system is also grounded in cognitivist theories of Second Language Acquisition (SLA), from the interlanguage concepts of Selinker (1972), to the emphasis on lexical forms by Felix (1981) and Hoey (2005). This approach also meets the requirements laid out for corpus-based language learning technology in Ghadessy et al. (2001) and Römer (2006); it takes advantage of advances in computational linguistics and has been implemented according to the latest developments of human language technology. The system incorporates techniques that can help KS3 students acquire accuracy and fluency in written English and develop life-long writing habits in learning English. As suggested by Milton (2006; 2011), we are taking advantage of online resources to help KS3 students and teachers shift from a machine- or teacher-centered pedagogy to one that puts the KS3 students at the center of the writing process by making the learner accountable, and ultimately more confident and independent.

Furthermore, based on and adapted from Bates' (2007) e-learning rationale, the system increases access to learners' resources, enhances teaching and learning, better prepares students for communication skills required in the international setting of Hong Kong, develops independent learning skills through online programming and mobile learning, and better accommodates the differing styles and background of students.

This rationale is also related to the term 'human-assisted' in that we still require expert analysis of potential 'errors' and/or problems Hong Kong students

make in their writing to assist us with writing instant prompts. The 'human-assisted' is also in relation to the expertise and experiences from Hong Kong teachers – by their use of data collected through the ePlatform of students' writing to diagnose individual problems. Sample writings from KS3 students were analysed to develop a better understanding of what type (and scope) of feedback is required for students when using the program; this allows for accurate and meaningful feedback prompts (and vocabulary profiles) for both students and teachers to use. *Figures 1.1 and 1.2* outline how we envisioned the Writing ePlatform to be used within a process writing approach to teaching and learning. The affordances of the Writing ePlatform vary and are not restricted to classroom teaching, nor encouraged, to be used only by teachers as a teaching tool in the classroom. In fact, the design of the ePlatform takes into consideration methodologies that include: independent learning, blended learning, and assessment as/for learning. For example, *Table 1* outlines an adapted version of Earl's (2003) assessment roles and goals, where both the teacher and the Writing ePlatform have shared roles and goals in a blended learning approach.

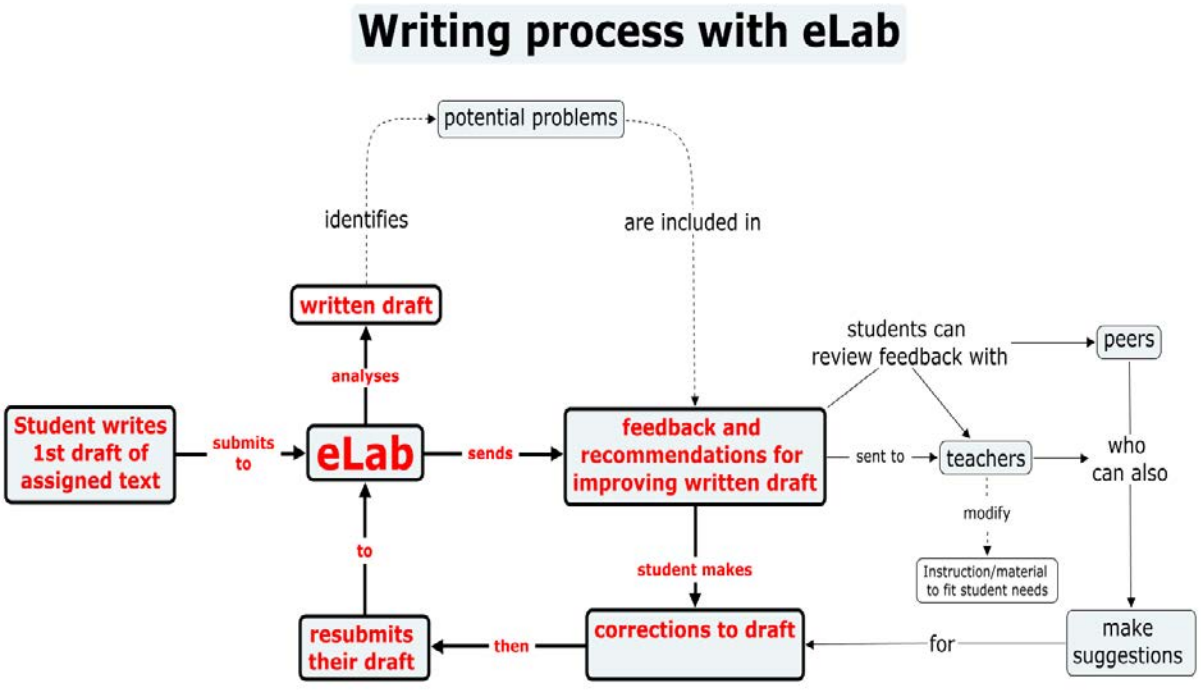


Figure 1.1 Suggested process writing approach with the Writing ePlatform

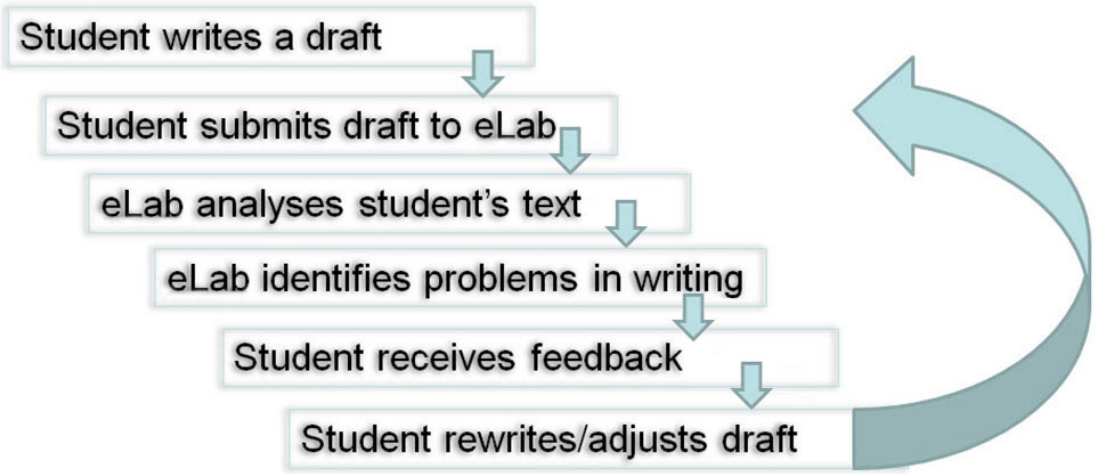


Figure 1.2 How Writing ePlatform assists in process writing

Table 1. Earl's Assessment Roles and Goals

Assessment Roles and Goals

Role	*Goal*
Teacher + **ePlatform** as mentor	Provide feedback and support to each student in **both formal (classroom) and informal (home) environments.**
Teacher + **ePlatform** as guide	Gather diagnostic information to lead the group **and individual students** through work at hand **and/or process.**
Teacher + **ePlatform** as accountant	Maintain records of student progress and achievement, **allowing review of entire processes and accumulation of learner corpora.**
Teacher + **ePlatform** as reporter	Report to parents, students, and the school administration about the student progress and achievement, **showing entire learning process.**
Teacher + **ePlatform** program director	Make adjustments and revisions to instruction practices **that focus on individual student needs based on empirical data.**

Source: Adapted from *Assessment As Learning* (Earl, 2003).

Developing and Improving Error Rules for the System

The establishment of error rules and the improvement of the reliability and validity of these rules are two key tasks in establishing this online system. These rules were compiled based on previous research at HKUST and the research data provided by EDB. Initially, a total of 1,800 written essays on 10 different topics (180 essays per topic) were collected from either Secondary 3 (S3) TSA or S3 students from six secondary schools. These were analysed and a list of common errors produced by local Hong Kong students was drawn up based on specialists' expertise. These common language problems were interpreted into error patterns that could be used to analyse text heuristically. Additional rules were added based on the work of Milton (2006, 2010, 2011) and through analysis of the corpus. To enrich the feedback given by the system, a repository of information about errors and reference tools was created. Students submitted written assignments set in class to the system for analysis. Their scripts were saved by the system and this allowed the organic growth of the corpus of Hong Kong S3 student essays. In this way, the accuracy of testing rules was continually improved and this feature will enable further modification and improvement of the system in the future. The following is an example of the process of identifying an error, writing rules, and designing the instant prompts for that error.

Common Error: Although + but in a sentence

Although the food was not great but I felt very happy that I was able to cook the meal by myself.

Although the food was not great, I felt very happy that I was able to cook the meal by myself.

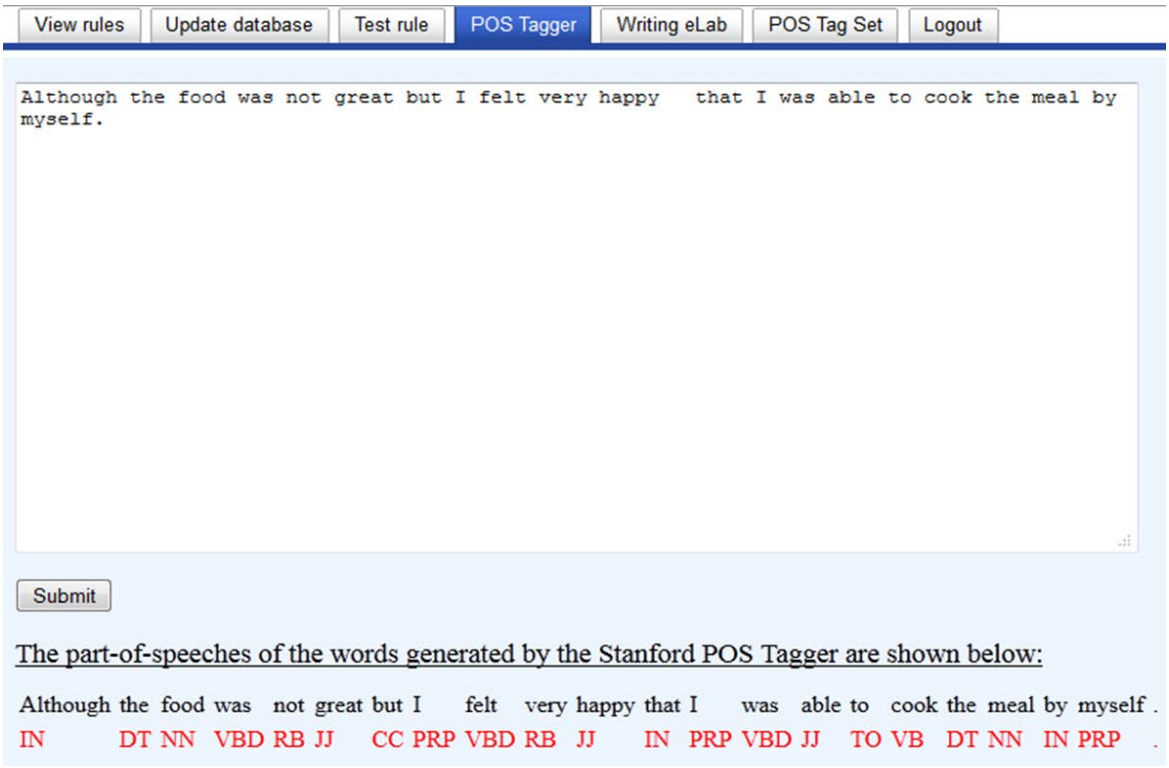
These rules were then converted to a formula for the ePlatform to identify patterns in students' writing.

Error formula:

!{well|such|known}{as|since|although}<{(CC)|(IN)}> *<0-6> (CC)<{but}>

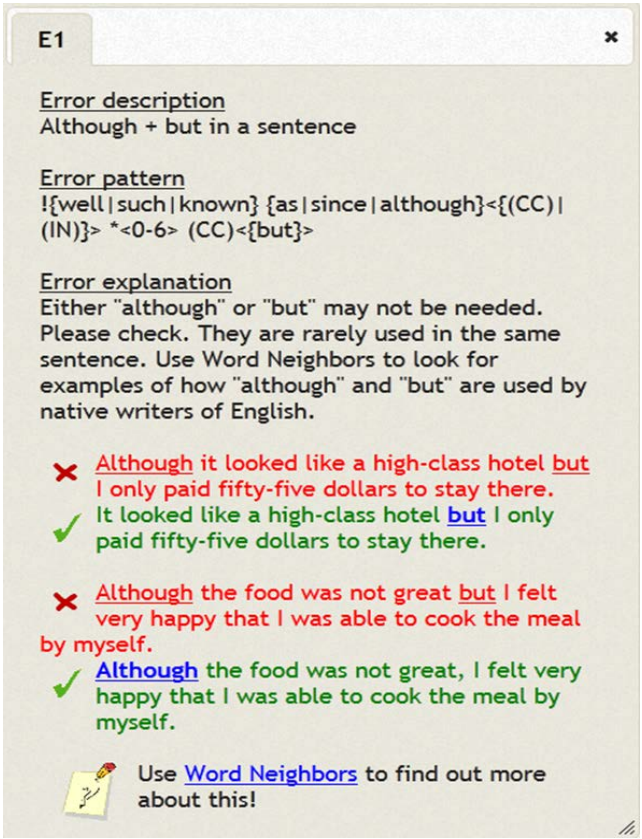
By using the Stanford Natural Language Processing Group's parts of speech tagger, the rules were tested using a learner corpus.

Example:



Then instant prompts were written and allocated to relevant rules.

Example:



We launched and completed two pilots with the same schools and students: the first in October 2012; the second in January 2013. Four topics were assigned to seven schools in October, and the remaining six topics were assigned to the schools in January. Each school was given two topics. Additionally, we collected feedback from schools on the use of the ePlatform at the end of each pilot. This included an online survey completed by 336 students; observation reports from teachers, providing feedback on any user or technical problems that have occurred, and suggestions on how to improve the program.

To inform the team on what rules needed adjustment or deletion, numerous batch tests to analyse, review and evaluate rules used in the Writing ePlatform were conducted during the project pilot. This included measuring the effectiveness and hit rate of rules. The batch tests include: (a) an analysis of the Learner Corpora on all 10 topics provided by EDB of students from previous years, and (b) an analysis of the Learner Corpora created from the October-December and January-May pilots. Results and analysis also included comparing the two learner corpora. Below is an example of findings from the batch test analysis:

Issues: Developmental challenges.
Example: Error identifies too many false positives.
Other error type: Wrong error identified.
False positives: Rule flags correct use of English as an error.
Correct identification: Rule identifies correct error.

Error Rule 1: Although + but in a sentence (start with a capital, end in a full stop)
Hits: 22
Issues: None.
Other error type: None.
False positives: None.
Correct Identification (samples):
{Although they were so old but }they were still very strong.

{Although it looked like a high-class hotel but }I paid fifthty-five dollars only.
{Although it is quite expensive, but }it is a fun way to learn English as we can know more about the culture of the place at the same time.
{Although we can't said Japan's language, but }we can said english with they.
{Although, this trip was very short, but }we were enjoyed for this tirp.
{Although, I know I will be fat, but }I have bought it!
{Although the food was not yummy but }I felt so happy that I can cook the meal by myself!
{although I had hurt, but }I really enjoy this trip.

Error Rule 2: A sentence / clause has more than 2 verbs
Hits: 23874
Issues:
This rule is very problematic – too many false positives. Detection issues.
Other error type: None.
False positives:
During the past summer holiday, my family and I {went to a village in the Mainland China to visit }my grandparents.
When we arrived, grandpa and grandma {were in excitement since we had}n't visit them for a long period of time.
They {took out a watermelon from the fridge to greet }us in ecstasy.
Our sweats from the travel {had gone }away immediately.
Because of summer holiday so there {were lots of children who living }near to me playing with me I was very joyful when I was playing with them.
Because of summer holiday so there were lots of children who living near to me {playing with me I was }very joyful when I was playing with them.
Because of summer holiday so there were lots of children who living near to me playing with me I was very joyful when I {was playing }with them.
On the second day, I {went to the farm to help }the farmer with my family.
I {tried to be }a farmer and started farming.

It can be seen in this example that the system was able to flag when Error Rule 2 required amendment. The use of the rule, “A sentence / clause has more than 2 verbs”, was problematic since it resulted in numerous false positives. (A “false positive” refers to a correct sentence being mistakenly judged as erroneous by the system. An online error correction system must seek to avoid false positives.) The many occurrences of false positives also correlate with other findings from student survey’s and teacher interviews, where frustration was expressed due to correct sentence patterns being identified as incorrect causing confusion among students. The second of the two examples suggests that the Although-but rule was applied correctly by the system.

Findings from the batch tests were then categorized to determine whether there was a relationship between error type and topic. Establishing the link between errors and writing topics led to us indexing the online material by both error type and topic in the eTutor, which teachers can use to inform teaching. For example, when teachers teach a new topic they can take into consideration errors identified as common in that topic or metalinguistic explanations available there for their students. The following table shows an example of rules and the hit rate in students writing according to topic. Topics were categorized into three main text types: describing, giving advice, and giving information.

Table 2. Occurrences of Error Rules Applied by Topic and by Text Type

Error Rule	Topic					Total	Text Type			Total
	1. An Accident	2. Hong Kong Sightseeing	3. Information about Schools in Hong Kong	4. Information and Advice about Going to S Secondary Schools	5. Sports Day		Describing (Topic 1 + Topic 2 + Topic 5)	Giving Advice (Topic 4)	Giving Information (Topic 3 + Topic 4)	
Go/ going/ went + shopping	5	5	0	0	0	10	10	0	0	10
Many + Singular	11	59	25	30	29	154	99	30	55	184
Modal verbs (will / can/ would) + adjective	4	6	3	12	5	80	15	12	15	42
Modal verbs (will/ can/ would) + noun	0	55	7	8	0	70	55	8	15	78

Further analyses on tallies such as those given in Table 2 together with other information such as the differing writing abilities of students, can provide useful information to the ePlatform team on more efficient ways to individualise feedback. The tallies would also have pedagogical implications for teachers. For example, teachers may know what grammatical points to emphasise in their pre-teaching of the writing tasks (i.e. the emphasis on teaching modal verbs before students are being assigned to write Sightseeing in Hong Kong).

Writing ePlatform: Features and Functions

Based on our pilot and batch test findings, our team finalized and developed the following key eLab and eTutor functions:

(1) A basic student user interface (*Figure 2.1*): This is main user interface for the eLab where students will submit their text for feedback. Students are given the option to choose the topic of their text, proficiency level to determine the type of instant feedback they will receive, and links to additional tools to assist them.

(2) Instant prompts (*Figure 2.2*): After a student has submitted their written work for analysis, any problematic text will be identified and highlighted in the student user interface. Clicking on the highlight text will show an instant prompt providing feedback on how a student can improve their text. Feedback in the instant prompts will also direct students to learning materials in the eTutor.

(3) Word Tag (*Figure 2.3*): This will give teachers and students a visual analysis of vocabulary frequency and type.

(4) Vocab-Profile (*Figure 2.4*): This will provide students and teachers with an analysis outlining vocabulary frequency based on Key Stage 1, 2, & 3 word lists (available on the website of Education Bureau).

(5) Web-based tools (*Figure 2.5 & 2.6*): These are web-based tools that operate within the eLab. These tools are available to users for discovering common vocabulary usage, collocations, and frequency.

(6) eTutor (*Figure 2.7, 2.8 & 2.9*): This web-based portal containing interactive learning objects based on common errors is organized by topic, error type, and additional input (videos). The eTutor provides guidance on how to reflect on common errors made by students in their writing.

These features allow the ePlatform to function as a tool within assessment as/for learning methodologies. For example, the ePlatform: (1) provides feedback and support to each student in both formal (classroom) and informal (home) environments; and (2) gathers diagnostic information to lead the group and individual students through work at hand and/or process.

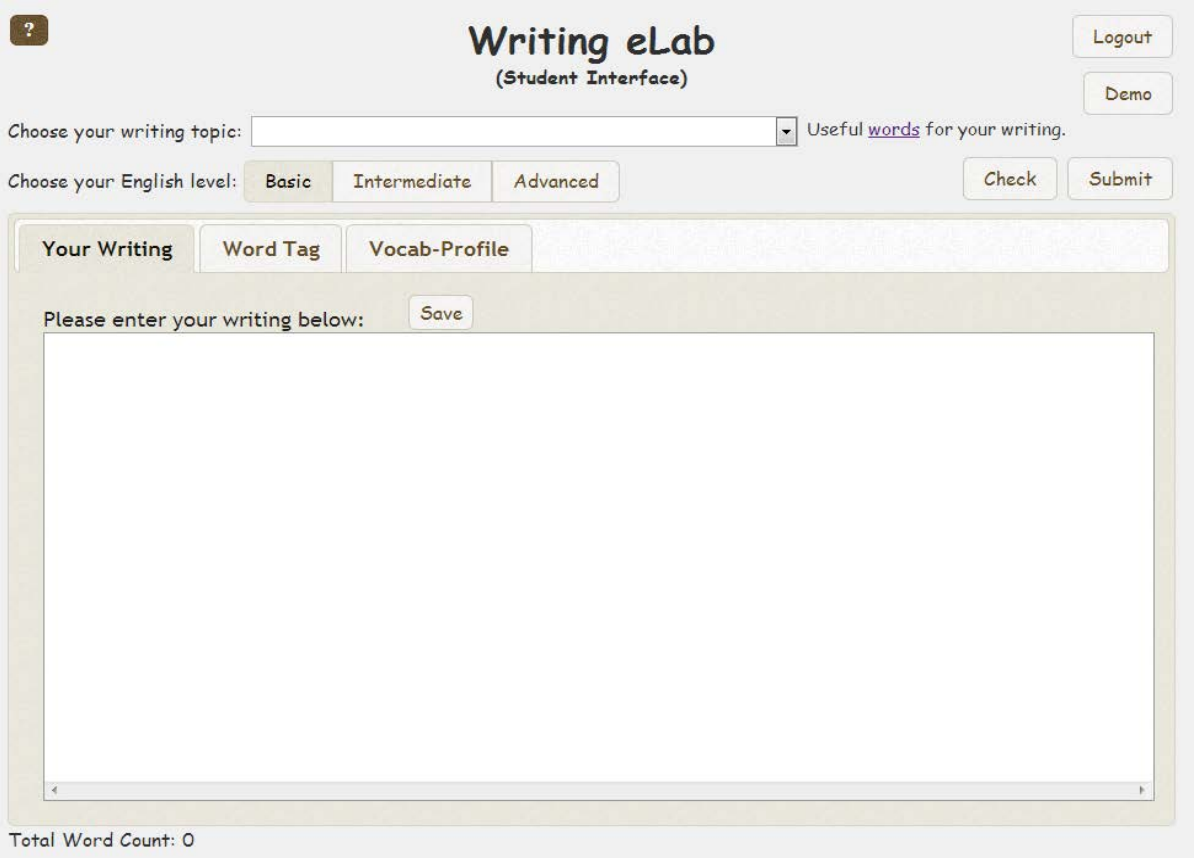


Figure 2.1 Writing ePlatform student user interface

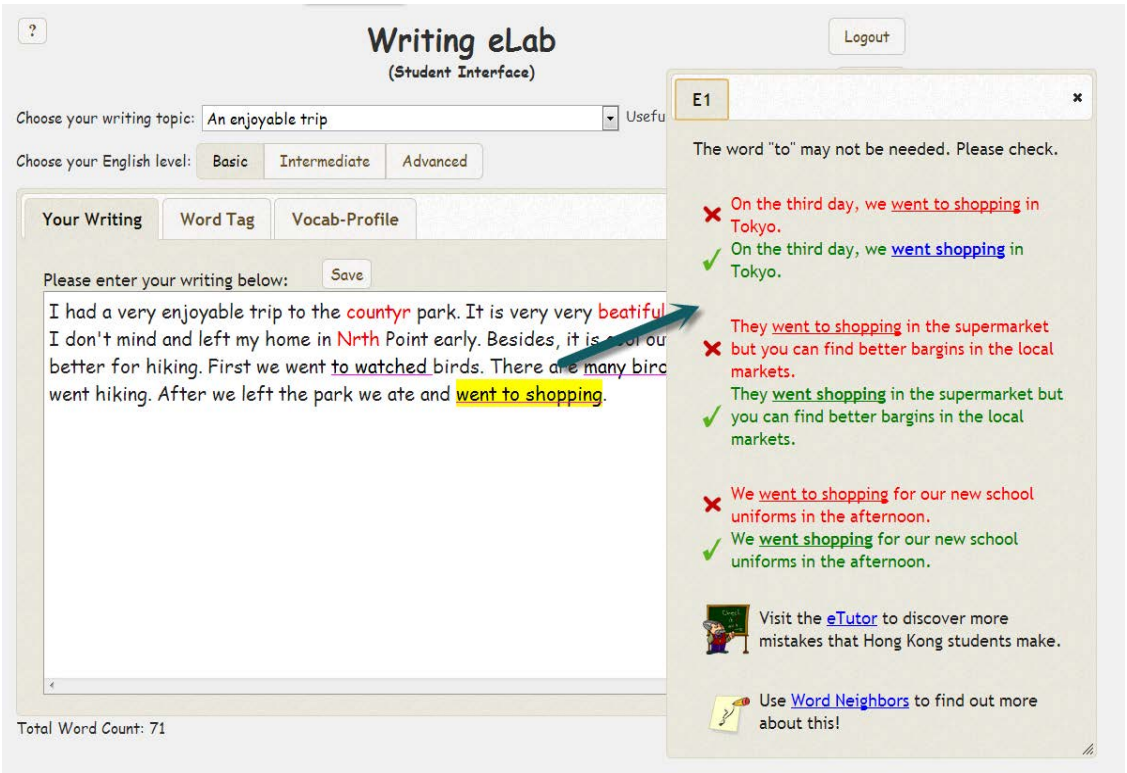


Figure 2.2 Instant prompt feedback

Figure 2.3 Word Tag

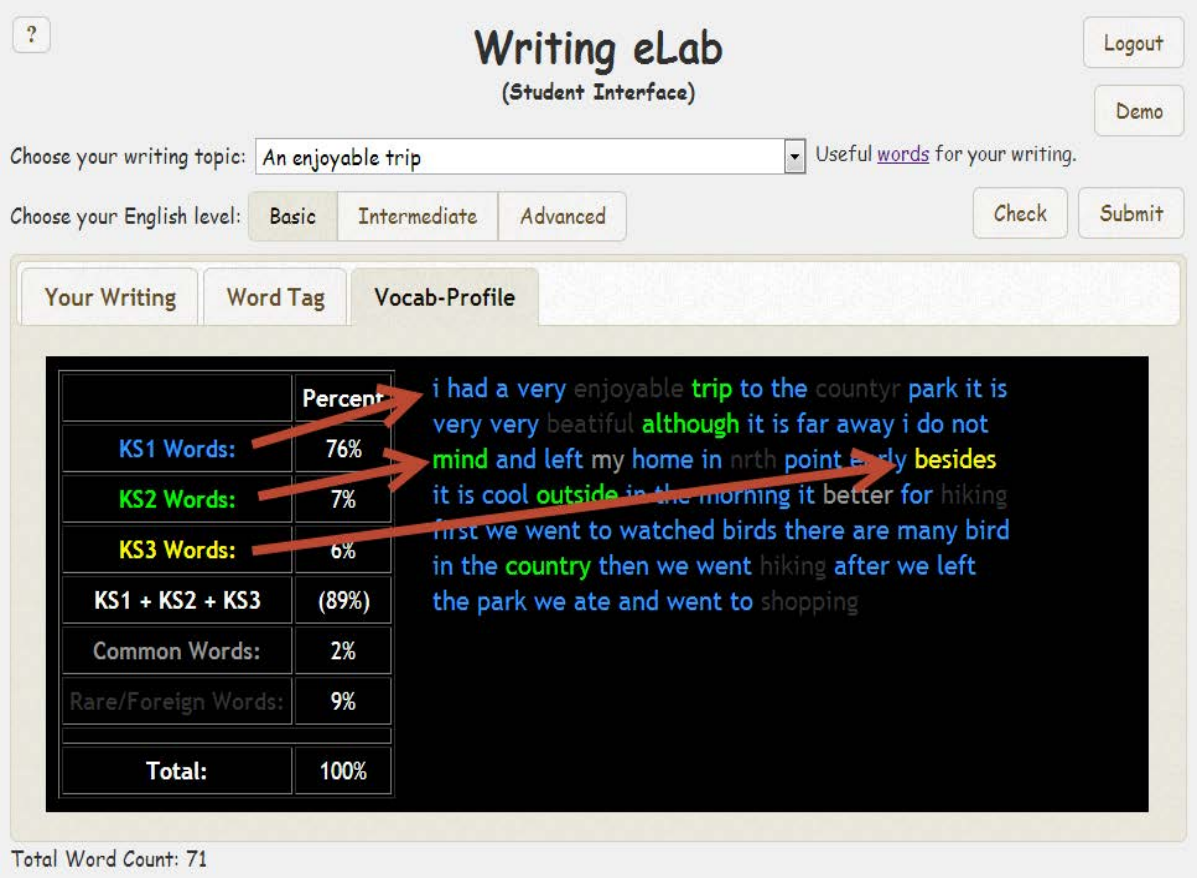


Figure 2.4 Vocab-Profile

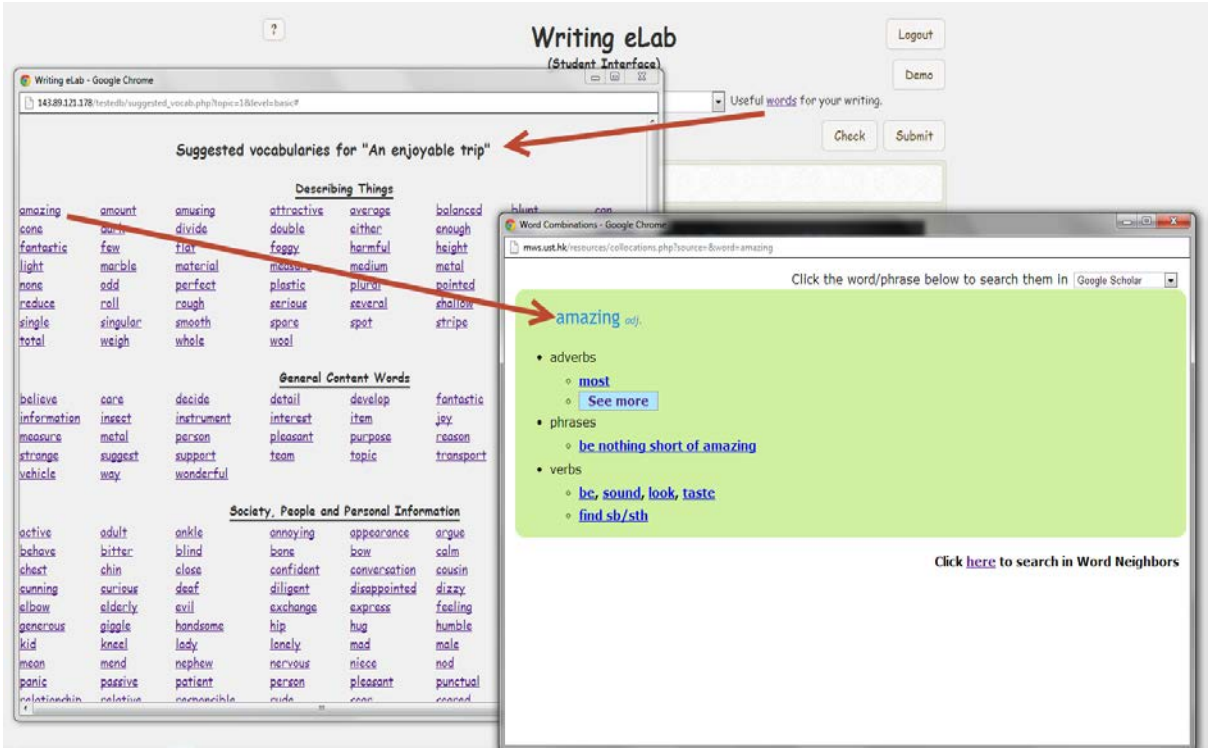


Figure 2.5 Useful Words for Your Writing

Figure 2.6 Word Neighbors (concordancer)

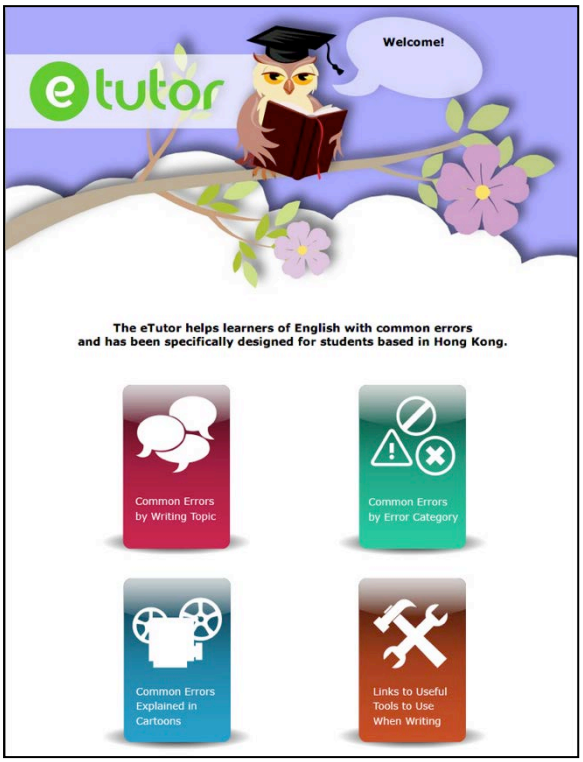


Figure 2.7 eTutor landing page

Figure 2.8 Category menu

Figure 2.9 eTutor error categorization

Conclusion

Great efforts have been made by the team to improve the validity and reliability of the online feedback system. Pilots with participating secondary schools were conducted, and the rich data which we have collected would assist us in improving continuously the Writing ePlatform. The learner corpus generated from the project also enables our team to analyse students' writing, all of which are useful for designing feedback and support material and also future projects and research. We believe that the final product will be a positive addition for the Hong Kong secondary school community, and with proper training and planning, the Writing ePlatform can enrich and enhance the writing process for English language teaching and learning. By engaging individual students with feedback and feedforward, the ePlatform can potentially benefit students and teachers in the writing classroom and beyond.

References

Anderson, T. (2004). Introduction: Teaching in an online learning context. *Theory and practice of online learning.* Alberta: Athabasca University.

Bates, A. (2007). Strategic planning for e-Learning in a polytechnic. In M. Bullen and D. Janes (Eds.), *Making the transition into e-learning strategies and issues.* Hershey, PA: Information Science Publishing.

Cheung, A. K. M. & Leung, F. F. Y. (2012). Using assessment data to investigate the relationship between speaking and writing. *Assessment & Learning, 1,* pp. 140-156.

Chickering, A. and Ehrmann, S. (1996). Implementing the seven principles: Technology as lever. *AAHE Bulletin*, October, pp. 3-6.

Coniam, D. (1997). A preliminary inquiry into using corpus word frequency data in the automatic generation of English cloze tests. *CALICO Journal, Vol. 16*, Nos. 2-4, pp. 15-33.

Coniam, D. (1999). Subjects' reactions to computer-based tests. *Journal of Educational Technology Systems, Vol. 23*, No. 3, pp. 195-206.

Earl, L. (2003). *Assessment as learning*. Thousand Oaks, CA: Corwin Press.

Felix, S. (1981). On the (in)applicability of Piagetian thought to language learning. *Studies in Second Language Acquisition, 3*(2), pp. 201-220.

Ghadessy, M., Henry, A. & Roseberry, R. (Eds.) (2001). *Small corpus studies and ELT: Theory and practice*. Amsterdam: John Benjamins.

Harmer, J. (1983). *The practice of English language teaching*. London: Longman.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.

Kuo, Grace (2005). A preliminary corpus study on EFL test takers' writing proficiency. Proceedings of the 8th Academic Forum on English Language Testing in Asia: Assessment for Learning, Hong Kong.

Meyer, C. F. (2002). *English corpus linguistics*. Cambridge: Cambridge University Press.

Milton, J. (2006). Resource-rich web-based feedback: Helping learners become independent writers. In K. Hyland and F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 123-138). Cambridge: Cambridge University Press.

Milton, J. (2010). A toolkit to assist L2 learners become independent writers. In Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics and Writing (pp. 33–41), Los Angeles, California.

Milton, J. (2011). A case study of a corpus-based, IT-enabled approach to improving English proficiency. Paper presented at ICTATLL2010 Kyoto Conference, Japan.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Co.

Römer, Ute. (2006). Where the computer meets language, literature, and pedagogy: Corpus analysis in English Studies. In A. Gerbig & A. Müller-Wood (Eds.), *How globalization affects the teaching of English: Studying culture through texts* (pp. 81-109). Lampeter: E. Mellen Press.

Selinker, L. (1972). Interlanguage. *IRAL, 10*, pp. 209-231.

Authors' email: Sean McMinn lcmcminn@ust.hk

Flora Leung scdoase@edb.gov.hk

Using SP Xpress as an Analysis Tool to Enhance the Learning and Teaching of Reading in English Language

Zoe So Yee Chan
Lingnan University Alumni Association (H.K.) Primary School

Introduction

Assessment for learning is an effective method of raising students' learning performance. It is based on the principle that students will most improve if they understand the objective of their learning and how they can achieve this objective.

Over the past two years, our school has aimed to develop assessment-for-learning practices. By investing in assessment data analysis, it was found that teachers can ensure that learning is meaningful for all students, teaching is effective and attainment outcomes are improved. This article shares our experiences of using SP Xpress as an analysis tool for assessment for learning in the key learning area of English Language Education.

The Aim of Assessment for Learning Strategy

The purpose of assessment for learning is to collect information about students' progress and achievements in relation to the learning objectives at the corresponding level. Teachers review their expectations of students' learning, the content of learning, and their teaching strategies to enhance learning and teaching through assessment, thereby providing *quality feedback* on how to improve performance. This is assessment for learning (CDC, 2004).

Our strategy is to make assessment for learning more systematic and consistent. Since this was the first stage of developing assessment for learning over the past two years, our aim was that *every teacher* has (1) equipped to make well-founded judgments about students' attainment, especially in reading,

(2) understood the concepts and principles of progression, and (3) have been able to know how to use their assessment judgments to plan ahead, particularly for students who are not achieving basic levels of competence.

Using SP Xpress® (Version 2.2) as an Analysis Tool

As mentioned in the previous section, assessment helps teachers understand students' learning processes, strengths and weaknesses, and the effectiveness of the related teaching strategies. Based on assessment data, teachers can adjust their own teaching objectives in order to fulfill students' learning need more effectively.

SP Xpress® (Version 2.2) is an assessment data analysis tool which helps teachers to implement assessment for learning. It analyzes students' performance in assessments, thereby generating statistical data and graphs to help teachers diagnose the strengths and weaknesses of each student. SP Xpress® (Version 2.2) also helps teachers further improve the quality of assessing items in order to transcend learning and teaching.

SP Xpress® (Version 2.2) reveals students' performance in assessment and thereby allows teachers to understand the relationship between students' learning ability and the facility level of the assessment items. It categorises students into 4 types according their performance in the assessment. These 4 types are:

Type A: high learning ability; high learning stability

Type B: high learning ability; low learning stability

Type C: low learning ability; low learning stability

Type D: low learning ability; high learning stability

Type A Students are able to follow the learning process and grasp the ideas of the assessing items; Type B Students are those who get the correct answers by luck due to their unstable performance in the assessment; Type C Students are the careless students who may lack assessment skills or misunderstand the

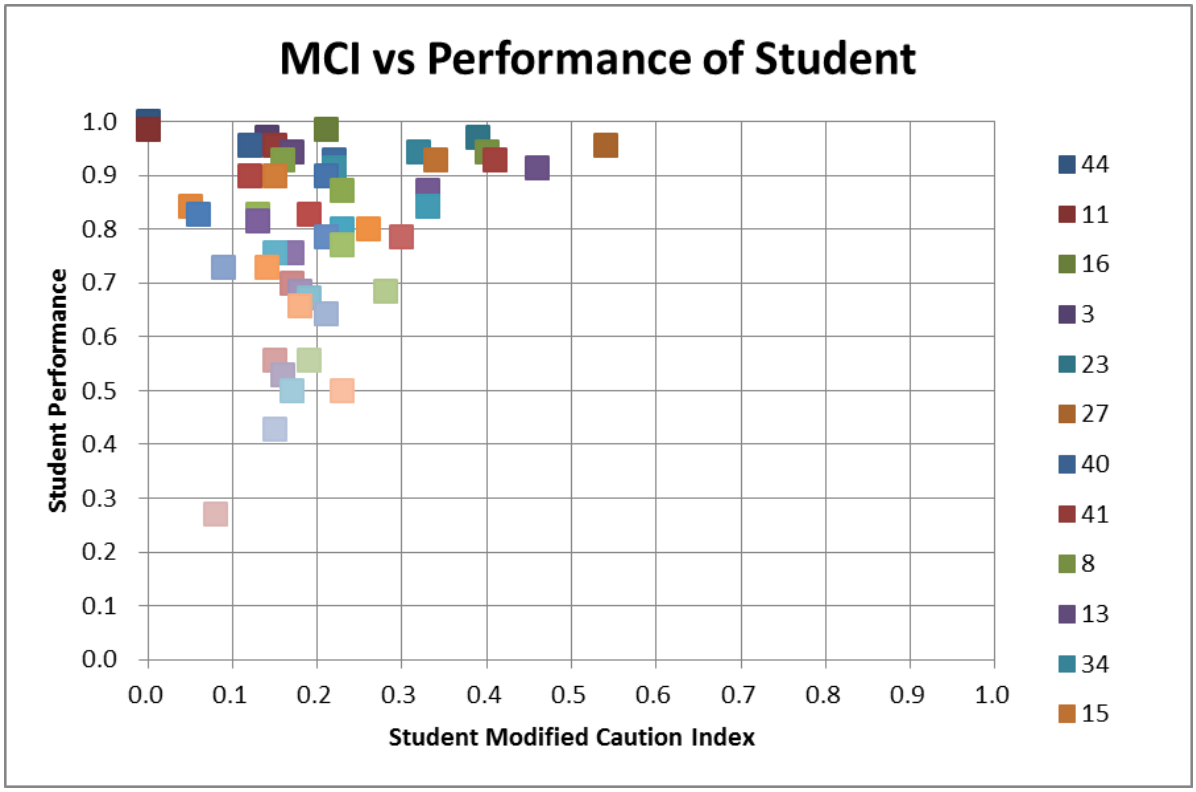
assessing items; Type D Students are remedial students. After knowing their students' situation, teachers can make plans with appropriate learning and teaching strategies for effective improvement (莫慕貞 et al., 2011).

The Use of Assessment Information

Using Summative Assessment Data for Groupings

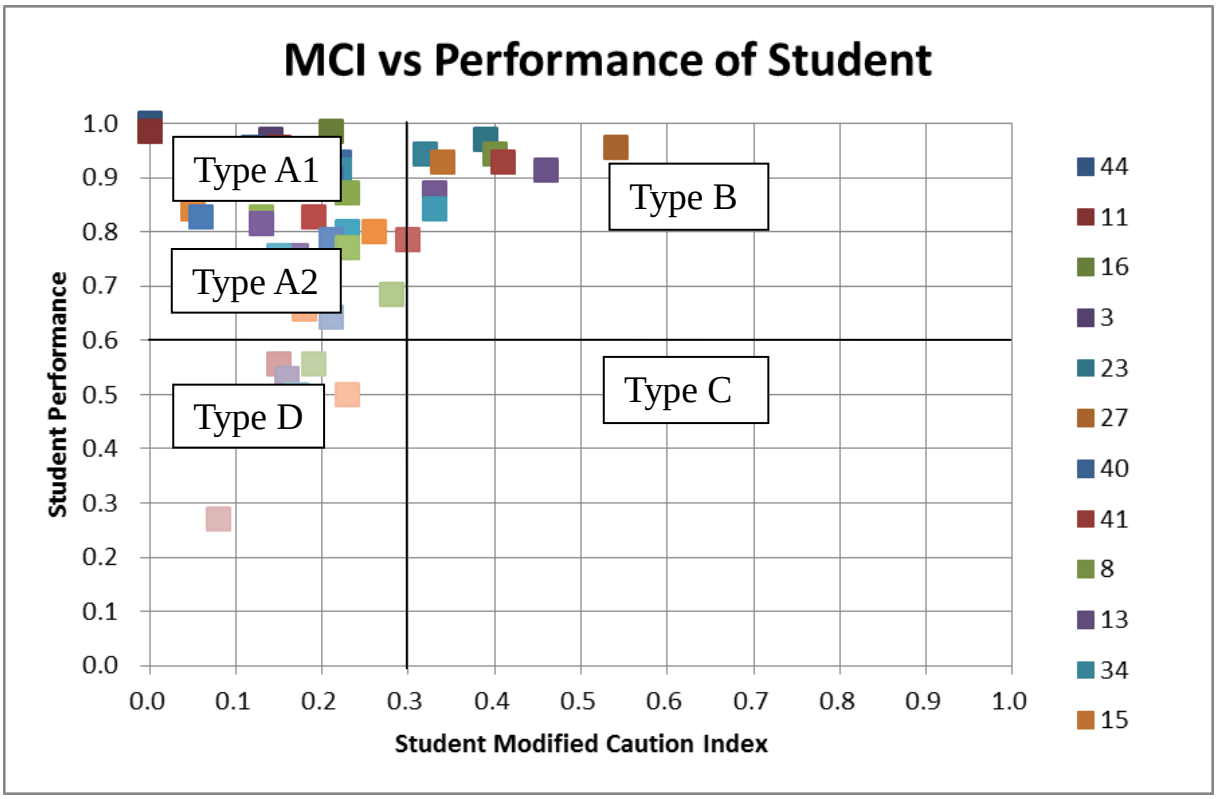
In the study of students' self-regulated learning, it was found that students with lower abilities used fewer self-regulated strategies than those with higher ability (Salili and Lai, 2003). In order to better cater for learning diversity, we placed students into 2-3 groups according to their English learning ability in Primary 3 to Primary 6. Thus, the first step of our strategic approach was to have *parallel English lessons* in each form level so that we can put students into the ability groups.

Summative assessment data of the previous school term was used when forming groups. We looked into the chart of "MCI vs Performance of Students" provided by SP Xpress which shows students' accuracy rate and their modified cautions index.



There are two ability groups in Primary 3 to Primary 5. All students of Type A are placed into RED GROUP and those who are Type D are placed into BLUE GROUP. For Type B and C Students, the form level English teachers are required to look into their performance in the assessment and assign suitable groups for those students.

For Primary 6, students are divided into 3 groups (RED GROUP, BLUE GROUP and YELLOW GROUP). As mentioned above, we look into the information from the chart of "MCI vs Performance of Students" but categorize students into 5 types instead of 4.



All students of Type A1 are placed in RED GROUP (higher ability group) while students of Type A2 are assigned to BLUE GROUP. YELLOW GROUP is the remedial group, which are Type D Students with lower learning abilities in English. As same as Primary 3 to Primary 5, other form level English teachers are required to look into the assessment performance of Type B & C Students and assign suitable groups for those students.

Using Summative Assessment Data for Locating Students' Learning Difficulties

For the first stage of the development of assessment over the past two years, we have focused on students' reading ability and 21 school-based reading strategies have been developed. Questions raised in reading lessons and asked in school-based reading worksheets are all based on these reading strategies according to our school-based reading curriculum framework. These reading strategies are then assessed in both formative and summative assessments in order to understand students' learning processes in reading. Using SP Xpress, the accuracy rate of each question is shown and we make good use of this data when planning reading lessons.

Question	Q.1	Q.2	Q.3	Q.4	Q.5	Q.6	Q.7	Average
Reading Strategies	Recognizing the format of text types	Analysis skills	Constructing meaning from the text	Dictionary skills	Identifying main ideas	Constructing meaning from the text	Relating personal experiences to reading	
5Red accuracy rate	92.59%	**14.81%**	*66.67%*	**37.04%**	**14.81%**	*66.67%*	100%	**56.08%**
5Blue accuracy rate	*61.54%*	**15.38%**	**26.92%**	**26.92%**	**15.38%**	**42.31%**	*65.38%*	**36.26%**
Total accuracy rate	77.36%	**15.09%**	**47.17%**	**32.08%**	**15.09%**	**54.72%**	83.02%	**46.36%**

The above shows an example of one of the comprehension sections in a summative assessment. According to the information, generally, students did not do well in constructing meaning from the text, identifying main ideas, applying analysis skills and dictionary skills. This implies that students at this form level should have further practice with these four reading strategies during the next school year and these strategies will come to be the core reading strategies at this form level next school year. Teachers are required to think of more questions with those reading strategies when planning lessons and designing reading worksheets. Questions with those reading strategies will be assessed with reading comprehension in both formative and summative assessment to understand students' learning progress.

Since students are formed into ability groups, questions with different learning levels are provided. When planning units of work, teachers are asked to think of questions with different reading strategies according to students' learning ability. That means teachers have to plan individual units of work of reading sections for RED GROUP, BLUE GROUP and YELLOW GROUP. When teaching the same reading strategy, questions may be different in terms of wording, cues, ways of delivery, etc.

Using Formative Assessment Data to Detect Students' Learning Difficulties

There are 2 formative assessments for reading comprehension in each term. The taught reading strategies, including the core reading strategies, are assessed in the paper. The reading strategy is shown next to the comprehension question to let students know about their reading performance and which kind of reading strategies they are weak in. Similar to the summative assessment, SP Xpress is also used as data analysis. As mentioned before, we emphasize on reading at the first stage.

Question	Q.1	Q.2	Q.3	Q.4	Q.5	Average
Reading Strategies	Recognising the format and language features of some common text types	Identifying main ideas	Locating specific information	Locating specific information	Analysis skills	
3RED accuracy rate	100%	92.59%	85.19%	*77.78%*	100%	91.11%
3BLUE accuracy rate	95.83%	95.83%	**29.17%**	*62.5%*	**54.17%**	*67.50%*
Total accuracy rate	97.92%	94.21%	**57.18%**	70.14%	77.09%	*69.22%*

Using the summative assessment information of the previous term, we were able to locate the core reading strategies for each level. For the formative assessment information, we could also decode students' existing/potential learning difficulties in reading. For example, RED GROUP students performed well in the above formative assessment but generally, only 77.78% of the students got the correct answer in Question 4 with the reading strategy of locating specific information. And BLUE GROUP students also had an

unsatisfactory performance in the application of the same reading strategy. The data reveals that maybe there was a lack of practice on locating specific information in the previous weeks or the questions within this reading strategy were too difficult for the students. For the form level English teachers, they were asked to think of some questions with this strategy in the following lessons. For BLUE GROUP students, they need to have training of analyze the passage as the assessment data shows that only 54.17% of the students got the accurate answer when applying analysis skills.

Outcomes of Using Assessment Data Analysis

With the development of 21 school-based reading strategies and school-based curriculum framework, teachers, as well as students and parents, know what students are learning and what is going to be assessed. By having better understanding of assessing knowledge, teachers know more about students' learning difficulties and with the development of structured and systematic assessment systems, teachers find it easier to track students' progress, especially in reading. Each teacher is able to recognize the weakest reading strategies that his/her students have and develop ways to improve the strategies by looking at the units of work and lesson plans. After two years, it has been observed that our students' reading level is improving when having benchmark level reading tests and the TSA. Although it is quite time-consuming to insert assessment data, teachers have revealed that it is worthwhile to spend time on the data analysis and they would like to have further analysis on other language skills such as writing and speaking.

Further Development

Self-regulated Learning

The development of assessment of students is still continuing and we aim to have further development over the next few years. Having built on the work of our teachers, our aims for the next stage are that:

- every student is aware of their progress, and understands what is required to improve and how to achieve this;
- every parent is aware of how their child is doing, what they need to do to improve and how they can support their child and their teachers.

The long-term goal of assessment analysis is to let students become 'metacognitively wise' (Galton, 2007). Students should learn to self-regulate their own learning by not only acquiring appropriate reading strategies provided by teachers but also be able to apply the strategies autonomously and recognize when a chosen strategy is inappropriate. A key to developing this kind of self-regulated learning is the use of processing feedback. In the future, we plan to get students involved more in their own learning by letting them identify their own mistakes, questioning them about the suitability of their chosen reading strategies and calling for suggestions as to how to solve the same problems when reading.

References

Curriculum Development Council (2004). *English Language Education Key Learning Area: English Language Curriculum Guide (Primary 1-6).* Hong Kong: The Education and Manpower Bureau.

Galton, M. (2007). *Learning and teaching in the primary school*. London: Sage Publications.

Lee, J. C. K., Yin, H., & Zhang, Z. (2009). Exploring the influence of the classroom environment on students' motivation and self-regulated learning in Hong Kong. *The Asia-Pacific Education Researcher, 18*(2), 219-232.

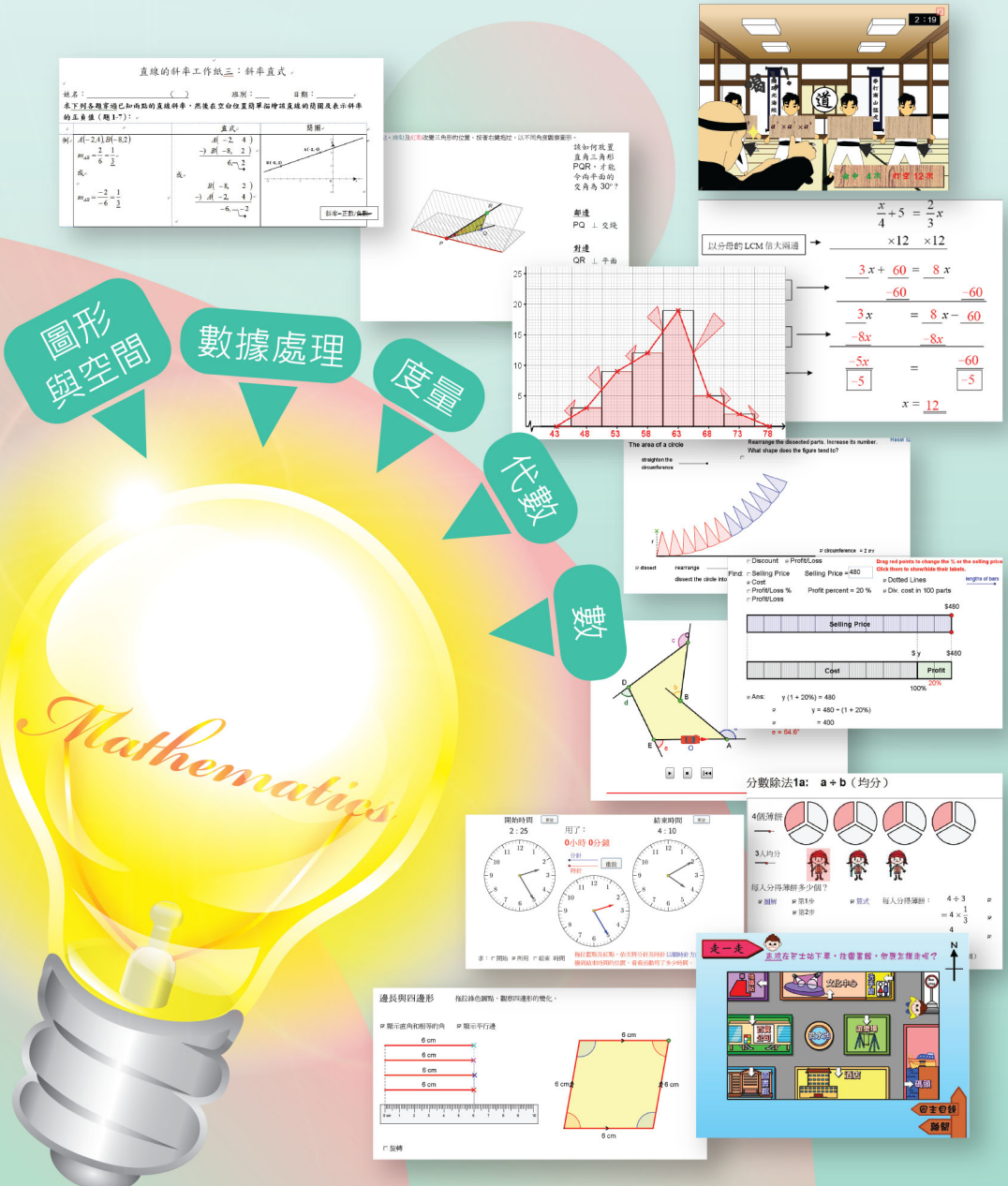
莫慕貞、丁彥銓、何昊璇、黃英華、謝棹南、徐坤、姚靜靜 (2011)。《優化學習導向評估之SP Xpress 2.2》。香港：PACE Publishing Ltd。

Author's e-mail:　chanmakmak@gmail.com

WLTS | 促進學習評估資源庫
網上學與教支援

wlts.edb.hkedcity.net

- ✓ 針對學習難點發展的互動課業
- ✓ 實踐「促進學習的評估」的理念



教學簡報 · GeoGebra/Cabri 3D · 網上遊戲

研究報告 · 課堂工作紙 · 教學計劃

Investigation in Learning and Teaching of Graphs of Linear Equations in Two Unknowns

Arthur Man Sang Lee
The University of Hong Kong
Anthony Chi Ming Or
Education Infrastructure Division, Education Bureau

Introduction

This article reports on a study conducted in 2011 for better understanding of students' learning difficulties on graphs of linear equations in two unknowns so as to improve the learning and teaching of this topic. It was the second part of a series of studies for improving teaching and learning of secondary mathematics based on in-depth analysis of assessment and pedagogy in selected topics. The study included testing and interviews with students, lesson design and tryout with teachers. Diagnostic tests and preliminary findings on students' misconceptions were informed by previous TSA results. Teaching and assessment materials were developed throughout the study and converted into teaching packages for online dissemination in Education Bureau's Web-based Learning and Teaching Support website (http://wlts.edb.hkedcity.net) after the study.

Curriculum and Basic Competencies

Knowledge about linear equations in 2 unknowns (or variables) and their graphs are essential for algebra learning after basic understanding and mastery of algebraic manipulations and equation solving. The concept is crucial for algebra learning in the later stage regarding solving systems of equations, functions and relations, coordinate geometry. In the current curriculum guide for Key Stage 3, this topic is subsumed under the unit of Linear Equations in Two Unknowns from the Number and Algebra dimension. The learning

objective is described as “plot and explore the graphs of linear equations in 2 unknowns”.

In the common textbook treatments this objective may easily be interpreted as a pre-requisite of solving simultaneous equations with graphical methods, that is, considered mainly as a preparation for learning the techniques of solving equations. When the emphasis is put on the tools for solving equations, crucial concepts above equations in 2 unknowns and their graphical representations may not be properly developed. We find that the introduction to graphing of individual linear equations is provided in the beginning of the chapter on simultaneous equations. This section is usually followed by another on graphical method of solving equations where graphs of 2 equations will be plotted to determine their intersection. The usual practice on graphing is focused on setting simple tables of 3 columns, which is considered as an essential skill of getting appropriate points to determine a straight line on the graph paper manually.

More detailed specifications of the learning outcomes related to the learning objective are given by the following Basic Competency (BC) descriptors for the Territory-wide System Assessment (TSA):

NA13-1 Student can plot graphs of linear equations in 2 unknowns

NA13-2 Student can demonstrate recognition that graphs of equations of the form $ax + by + c = 0$ are straight lines

NA13-3 Student can determine whether a point lies on a straight line with a given equation

These learning outcomes can be considered as elaboration of the objective in the curriculum guide regarding plotting and exploring graphs of linear equations in 2 unknowns. These detailed specifications may help to shift the attention of the overall task of plotting 2 equations from getting the intersection and extracting the solution to both equations. It is now more likely to observe what the students can or cannot do when simply plotting one equation and interpret any point on line as possible solution to an equation in 2 unknowns.

Performance in TSA

The BC "NA13-1: Student can plot graphs of linear equations in 2 unknowns" seems to expect a very simple skill of graphing a single equation (compared with coordinating graphs of two equations and getting the intersection). However, students' performances over the years are consistently low. In 2008 and 2009, there were parallel questions in separate papers asking students to plot graphs with and without providing a table. Results in these two years obviously show that the questions without table are answered less as well (Figure 1).

2008 TSA Paper 1 Q.34 *Facility: 34.8%*

Draw the graph of $x + y = 1$ on the given rectangular coordinate plane in the ANSWER BOOKLET.

2008 TSA Paper 2 Q.49 & Paper 3 Q.48 *Facility: 76.2% (table); 68.4% (graph)*

Complete the following table for the equation $2y = x + 4$ in the ANSWER BOOKLET:

x	−4	0	4
y			4

Draw the graph of this equation on the rectangular coordinate plane given in the ANSWER BOOKLET.

2009 TSA Paper 1 Q.28 *Facility: 30.9%*

Draw the graph of $2y = x + 1$ on the rectangular coordinate plane given in the ANSWER BOOKLET.

2009 TSA Paper 2 Q.43 & Paper 3 Q.43 *Facility: 71.6% (table); 61.6% (graph)*

Complete the following table for the equation $2y = x + 1$ in the ANSWER BOOKLET.

x	− 3	0	3
y			2

Draw the graph of this equation on the rectangular coordinate plane given in the ANSWER BOOKLET.

Figure 1

Evaluation of random sample answer scripts on these questions also reveals qualitatively different answers to these questions. That means removing the table from the questions may not simply increase the procedural or computational difficulty of the tasks and result in more mistakes. It may also suggest that students can interpret the task in another way and show their lack of fundamental understanding of the graphing activity. For example, the answer in Figure 2 is found in 6 answer scripts among 100 randomly selected scripts containing 2009 Paper 1 Q28. In this figure, exactly one point (1,2) is marked but it is not clear how this is related to the given equation. Another 16 answers, out of those 100 sample scripts, are also found to be similarly incomprehensible. However, from sample scripts of other papers containing the parallel question with table, no such kind of answers is found.

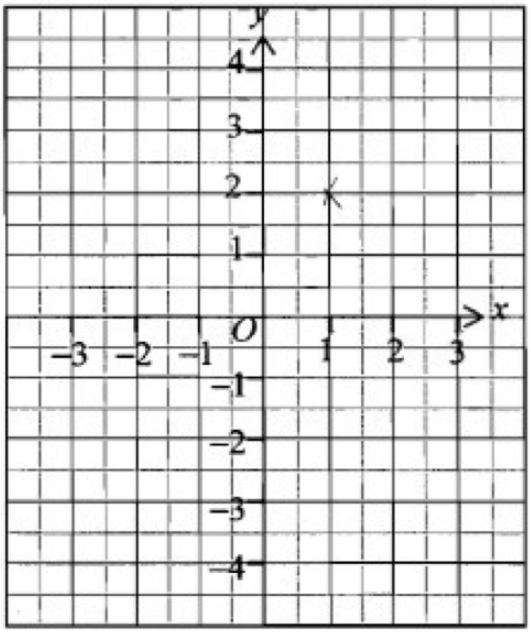


Figure 2

On the other hand, another kind of rather common responses are found among those wrong answers associated with tables. Since mistakes in calculation there may lead to 3 non-collinear points, students may join the 3 points to make a triangle in such case (like the example show in Figure 3). It is noteworthy that these cases suggest clearly students' misinterpretation of the procedure of joining the points as creating a geometric figure rather than producing a representation of infinite solutions to an equation.

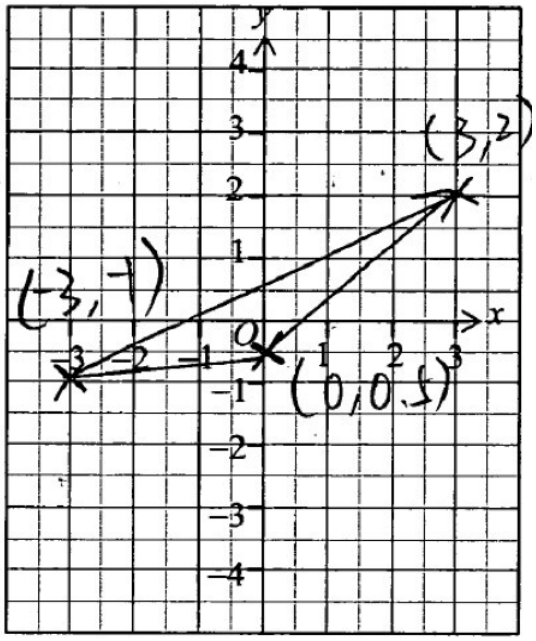


Figure 3

The BC "NA13-2: Student can demonstrate recognition that graphs of equations of the form ax + by + c = 0 are straight lines" focuses on recognition of the distinction between linear and non-linear graphs in terms of their appearance as a straight line or other forms. Such conceptual understanding was not explicitly emphasized in the previous curriculum. This characteristic of linear graphs should be accepted by students if they have to handle confidently the graphical method of solving equations. However, little is known about how they acquire this concept from their limited experience working with equations and graphs. On the other hand, it is also not clear to what extent and how this concept should be explained to students. In fact, a satisfactory explanation may involve deeper understanding of graphs and functions, which is usually not expected in this stage.

While identifying this subtle difficulty in teaching and learning, questions on this learning outcome in the papers so far can barely assess students' understanding in this area. The form of questions is very similar in all these papers. Students are asked to identify from 4 given graphs the one for a given linear equation. They need not pay attention to the details of the graphs and the major difference among the options is the linearity. The following item in 2006 illustrates the type of questions asked similarly in the subsequent years.

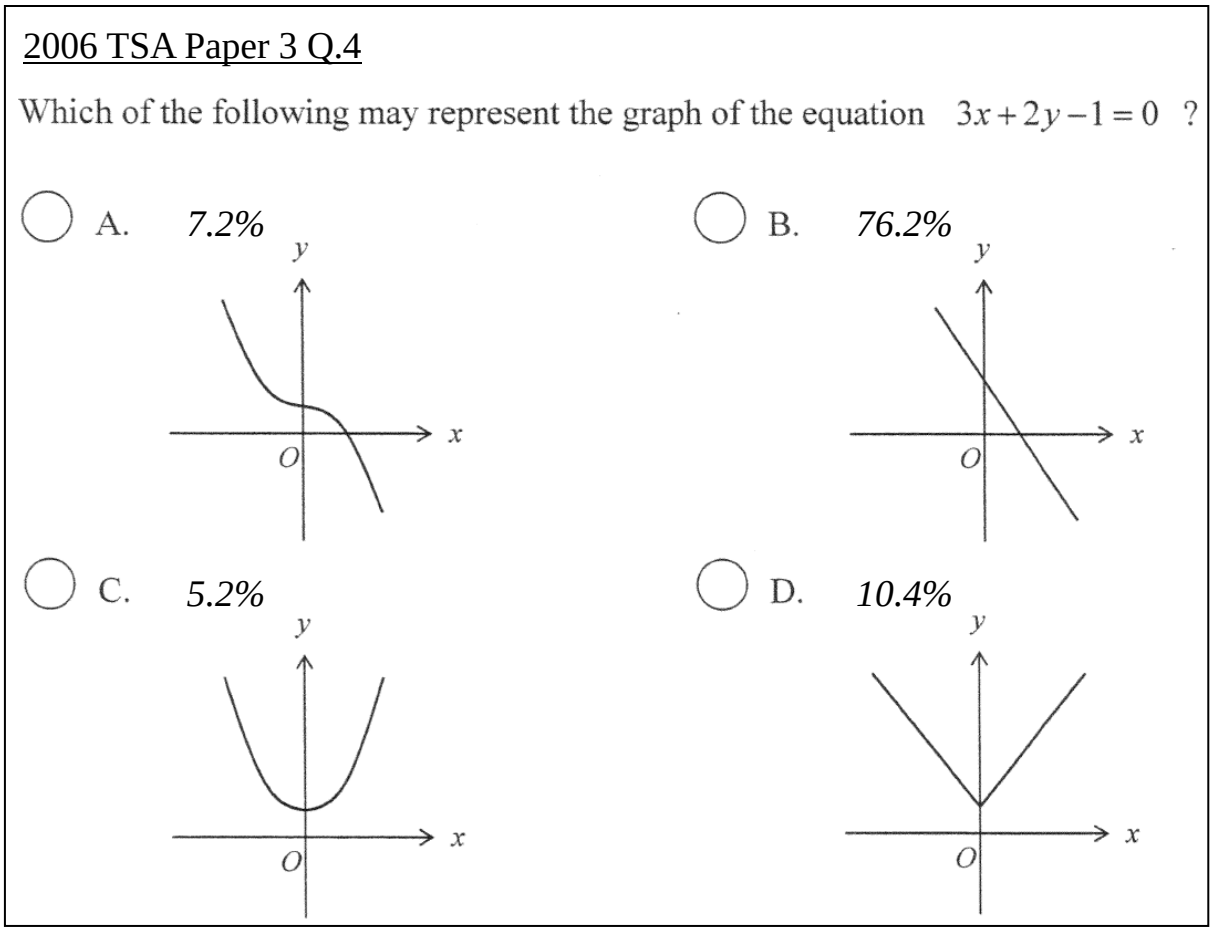


Figure 4

The facilities of these questions from 2006 to 2010 are respectively 76.2%, 74.2%, 76.3%, 74.1% and 69.9%. There is obviously limitation on this way of asking. The given equation must be linear since they are not expected to recognize or have experience on other types of equations. Among all the options, all they have to choose is a linear graph and most probably they have not worked with other types of graphs of equations. Despite this repeated form of questions, students' performance remains in a relatively stable level of 75% facility.

As suggested from these results, we have some crucial questions about teaching and learning.

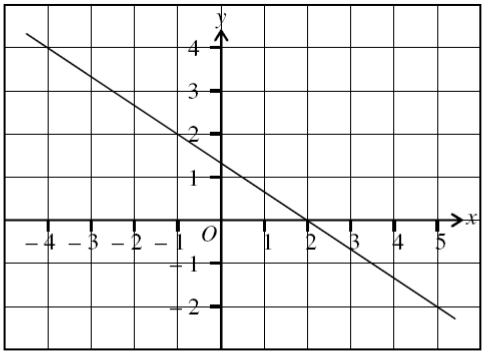
- Why do students fail to recognize this apparently obvious characteristic of linear graph?
- Does the failure to recognize this characteristic interfere with the general work of graphing and solving equations?

- Is this inability to distinguish linear graphs linked with other fundamental conceptual obstacles of graphing a general or only a specific issue about linear equations?
- For students who answer correctly, will they merely answer by eliminating unfamiliar examples or relying on superficial clues?

Some of these become guiding questions in the subsequent investigation.

The BC "NA13-3: Student can determine whether a point lies on a straight line with a given equation" is also closely related to students' understanding of the meaning of solutions to an equation of 2 unknowns, whether verified algebraically or represented graphically. There are two main types of questions. More often, a linear equation and its graph are given while students are asked to decide whether given points are solutions to the equation or lying on the graph. This can be done partially by quick inspection on the graph and further verified by substitution in the equation. In other questions, the graph is not given and students can only verify a solution by algebraic substitution (Figure 5). Facilities of these questions in 2006 to 2010 range from 41.8% to 67.4%. It is quite clear that the performance remains in a relatively low level.

2006 TSA Paper 3 Q.4 *Facility: 50.9%*



The figure above shows the graph of $2x+3y=4$. Which of the following points lie on the straight line? (There may be more than one answer.)

$$P(-3\ ,\ 3),\quad Q(-1\ ,\ 2),\quad R(0\ ,1),\quad S(4\ ,-\frac{4}{3})$$

Figure 5

Pretest and Diagnosis

A learning study was carried out in a secondary school. Before the study, a diagnostic test was conducted in the school. The test covers some key learning objectives related to linear graphs, partly based on the findings explained in the previous section. It is intended to explore understanding of secondary 3 students who have just studied the topic of simultaneous equations, including the knowledge of graphical methods. It is also considered as a pretest to the following learning study. Due to the special arrangement of the school in this year, this topic will also be taught to secondary 2 students in the second term. Findings from the test can therefore inform our design of the research lesson on the same topic.

The test starts with several questions directly copied from the TSA. These include questions for plotting graphs with or without a table given (the 2009 TSA question in Figure 1). It also includes multiple-choice questions for distinguishing linear from non-linear graphs (e.g. the 2006 TSA question in Figure 2). The rest of the test is further developed from similar questions in order to gather more detailed information about students' understanding in this topic. Several of these questions require students to compare algebraic relations expressed in the form of symbolic equations, tables and graphs (Figure 6 and 7). They are also required to distinguish points that belong to an equation or not while presenting these values of x and y in different ways (Figure 8). Another type of questions require students to distinguish linear and non-linear equations (Figure 9).

Which of the following equations is satisfied by the data in the table?

x	−2	−1	0	1
y	−5	−4	−3	−2

☐ A. $y = x - 3$
☐ B. $y = x - 1$
☐ C. $y = 2x - 1$
☐ D. $y = 3x + 1$

Figure 6

Which of the following sets of data can be used to draw a straight line on the rectangular coordinate plane?　(You can choose more than one answer.)

☐

x	y
−1	18
0	20
1	22
2	24

☐

x	y
−1	1
0	3
1	7
2	11

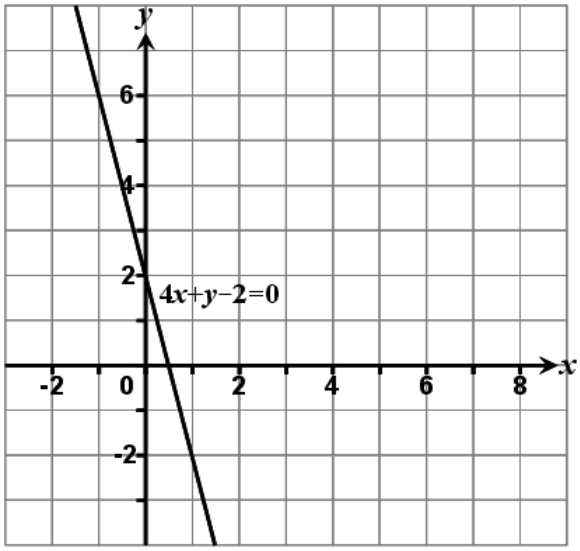
☐

x	y
−1	2
0	3
1	8
2	16

☐

x	y
−1	50
0	45
1	40
2	35

Figure 7



The figure above shows the graph of $4x + y - 2 = 0$. (Put a tick (✓) into the box of the answer.)

Point	Does the graph of $4x + y - 2 = 0$ pass through the point?
(−1, 6)	☐ Yes ☐ No
(1, 1)	☐ Yes ☐ No
(−2, 10)	☐ Yes ☐ No
(−5, 18)	☐ Yes ☐ No
(−2, −6)	☐ Yes ☐ No
$(a, 2-4a)$	☐ Yes ☐ No

Figure 8

Put ticks (✓) into the boxes of the following equations whose graphs are straight lines.

☐ $y = 2x$	☐ $\frac{1}{2}x = 3y$
☐ $x = 2 - 3y$	☐ $y = \frac{1}{1+x^2}$
☐ $2x + 3y - 5 = 0$	☐ $y = -1$
☐ $y = x^2$	☐ $y = \frac{1}{x}$
☐ $xy = 4$	☐ $y = \frac{1}{2-x}$

Figure 9

The test was done by 5 classes in secondary 3. The following are some major findings and questions about students' learning from analysis of their answers.

- Similar to the results of TSA, when a table is not provided, performance on graph plotting is significantly poorer.
- Some students merely join the points in a graph with line segments only, instead of a line extended to the edges of the graphing region. Do they notice any difference between lines and line segments for representing an equation?
- Most students can carry out substitution for finding values in an equation. However, this skill may not lead to understanding of the relation between equations and graphs.
- When matching a graph with an equation, some students may evaluate only one pair of values. Do they know that checking one ordered pair is not sufficient?
- When they are given tables of values but without equation, it is not easy to identify linear relations between the variables.
- When they are asked to suggest equations containing a specific point, most can provide only one such equation. Do they know that many different equations can have the same point lying on their graphs?

The diagnostic test is followed by task-based interviews. Students are selected based on their responses to the test. During the interviews, students are asked to repeat some questions from the test and explain to interviewers their reasoning. An additional group interview is also arranged to test students' responses when they are working together on a learning task that may be done in normal classroom setting.

The individual interviews show that the students are generally capable of checking individual points on a graph by substitution in an equation. Therefore, they basically understand the relation between a point on a graph and a solution for an equation. However, such knowledge may not sufficiently support their work and reasoning on various unfamiliar tasks relating equations, tables and graphs. One non-trivial task for them is to distinguish between linear and non-linear equations or graphs. When they are given some graphs with various forms for matching with a linear equation, they may not immediately recognize that the graph should be a line. Instead, they may refer to the basic principle by testing sample points or rely on a standard plotting procedure to find out which graph is appropriate (Figure 10). This shows that these students cannot readily accept this assumption that linear equation should generate a line graph and in fact this assumption may not be thoroughly explained due to limitation from the curriculum and common classroom practice.

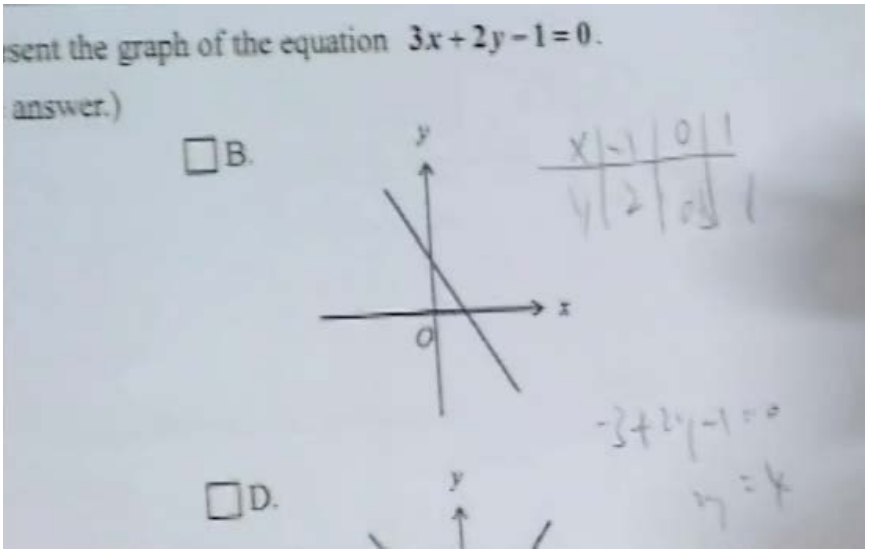
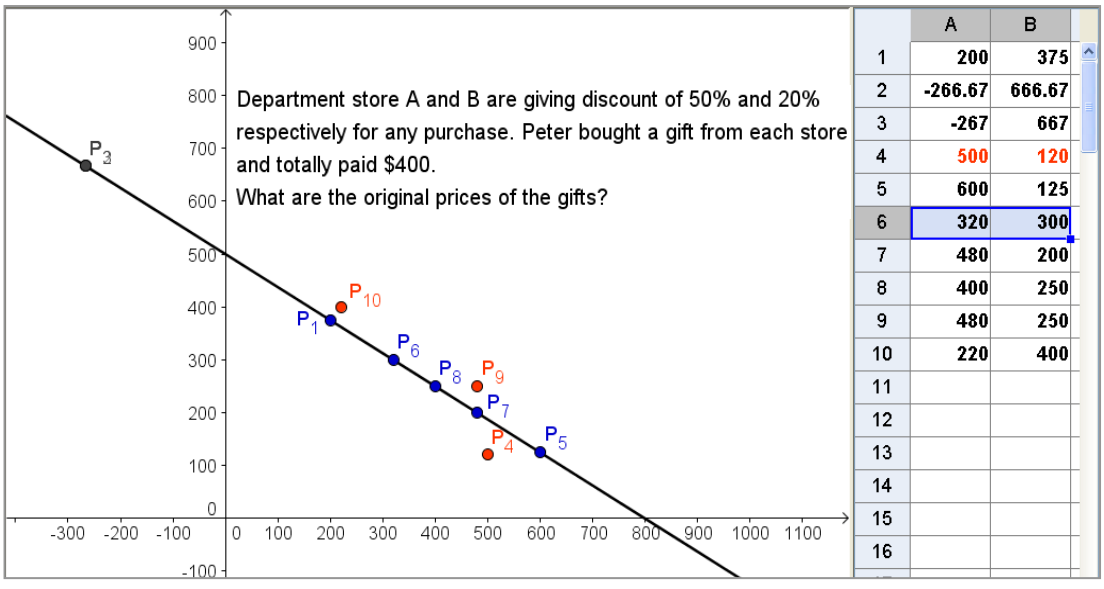


Figure 10

The group interview was carried out with 3 students working on an extended task. They were given a daily-life problem about shopping, leading to an equation of two variables representing possible prices of two gifts (see Figure 11). We would like to know whether the students can recognize a linear equation in an unfamiliar setting and whether the graphical representation makes sense to them in this context. Findings from this interview are useful for further developing this task to be used in the coming research lesson.

In the first part of the interview, students mainly focused on formulating an equation that can describe the unknown or variable quantities. They also managed to suggest pairs of possible values satisfying the conditions. These pairs of values suggested by students were then pooled together and listed in a spreadsheet projected on a white board. While they agreed that these were solutions for the equations and the points were plotted on the graphing area, they seemed not to recognize that there should be a linear pattern among the points. This was further verified when they were asked to distinguish additional points that may lie on or off the graph. They were not certain whether a point is solution to the formulated equation even when it was plotted and obviously collinear or non-collinear with the previously verified points. The behavior of the students suggests that even if they master the skills of graph plotting and understand the meaning of individual points on the graph, they may not fully understand the overall linear pattern of these points and how a line can represent all possible solutions.



	A	B
1	200	375
2	-266.67	666.67
3	-267	667
4	500	120
5	600	125
6	320	300
7	480	200
8	400	250
9	480	250
10	220	400
11		
12		
13		
14		
15		
16		

Figure 11

Research Lessons

A series of research lessons were designed and tried out in the second term in all secondary 2 classes. These lessons were supposed to be introduction to the concept of linear equation in 2 unknowns and their graphs. There were meetings with mainly the secondary 2 teachers for reviewing the pretest and interviews findings, identifying objects of learning and developing appropriate tasks for these beginning lessons of the chapter. The objects of learning and critical aspects were formulated based on previous findings of students' difficulties and conception. The teaching and learning tasks were designed by analyzing these objects of learning and considering potential of dynamic graphing environment.

As an introduction to the concept of linear equations and graphing, we spend most of these beginning lessons on developing the notions of solutions to an equation in two variables and graphical means to visualize a collection of solutions. Another key idea in the later part is the distinction between linear and non-linear equations. There is also attempt to explain why graphs of linear equations should be straight lines. These emphases are normally not included in this chapter but the school teachers agree that this learning study provides a good opportunity for adding these important elements to the lessons.

Research Lesson 1

Object of learning: recognizing graph of an equation in 2 unknowns in the coordinate plane

Critical aspects:
- Infinitely many points satisfying the equation
- ALL these points form a line or a curve in the coordinate plane
- Equation can be used to determine whether a point lies on the graph

Outline of lesson:

1. Given the point (2,4), ask students to propose different equations satisfied by this point.
2. Find more points satisfying y=x+2 and plot them in GeoGebra.
3. Trace other points satisfying the equation in GeoGebra.
4. Define graph as the set of ALL points satisfying the equation.
5. Check whether a point lies on the graph using the equation.

The first lesson starts with a simple revision on the coordinate plane. From examples of points and coordinates, which are worked out by the students, the concept of a relation between coordinates from a point is introduced and explored. Although the students should have experience in plotting and reading points, the idea of relating coordinates from the same point is probably new to them. For example, when focusing on the point (2, 4), we may say that the *y*-coordinate is twice the *x*-coordinate. Similarly, students are encouraged to suggest other possible relations to describe this pair of coordinates. When they are comfortable with talking about these relations, a 'short form' can be introduced by teacher to facilitate communication: express them as equations in *x* and *y*.

Each of these equations suggested by the students or teacher can give rise to multiple solutions. Choosing $y = x + 2$ as an example, (2, 4) is not the only point where *y*-coordinate is 2 more than the *x*-coordinate. The task is then led to search for other possible solutions or points sharing the same property. The process gradually leads to formulation of equations in two variables, solutions and graphs.

In this activity, the use of prepared dynamic graphing tool is crucial. A template is set up so that teacher can work directly on a spreadsheet and graphing area (Figure 12). Points and equations suggested during the activity can be easily recorded and amended on the screen. Various questions are designed to prompt students to identify different points on the grid belonging (or not belonging) to a specific equation, such as points with fractional coordinates, negative coordinates, or lying outside the graphing region that can be easily revealed with adjustment tools built in the software. Finally, automatic marking of suitable positions as a point being freely dragged on the screen gives a vivid holistic picture of the graph as a collection of points satisfying an equation. The dynamic tool specifically designed for the activity greatly enhances the students' participation and teacher's explanation. The teachers admit that the tool is new to them and appreciate its use.

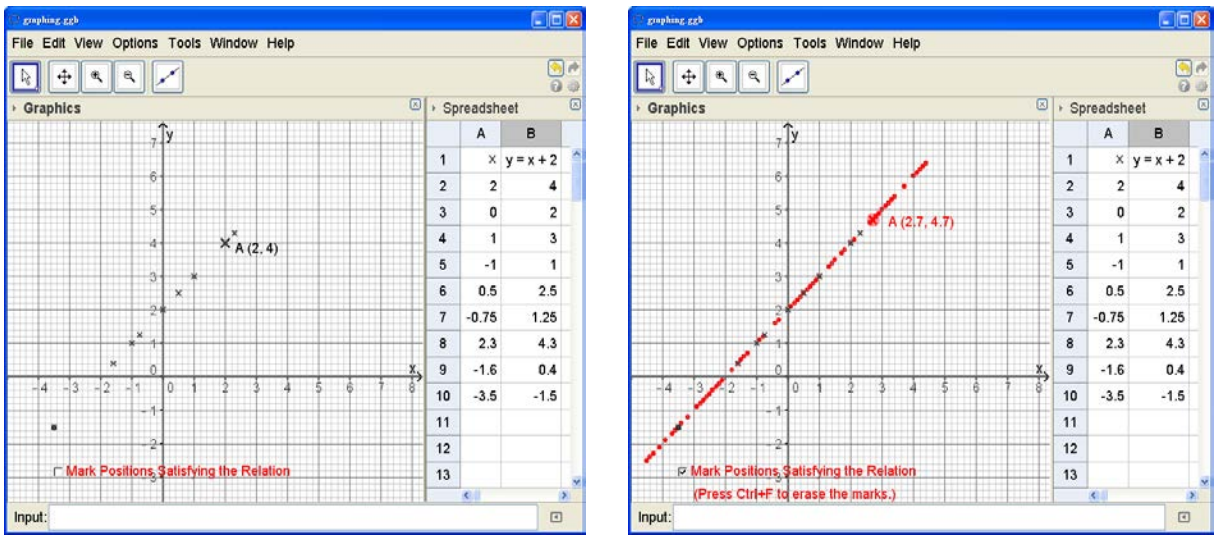


Figure 12

Research Lesson 2

Object of learning: plotting graphs of equations in 2 unknowns

Critical aspects:
- y in terms of x
- Equidistant points of integral coordinates (equal increment of x with equal increment of y)

Outline of lesson:

1. Equations are first given in the form $y = f(x)$.
2. Find values of x which give integer values of y.
3. Observe the "equidistance" between values of x and y.
4. Observe distance between values of x and the "denominator".
5. Plot the points in GeoGebra.
6. Express y in terms of x for equations in general form.

The tasks in this lesson focus on more systematic methods of graphing. Building on the previous notion of infinitely possible solutions to an equation and the possibility of representing them in a graph, the students are led to some systematic search of solutions to linear equations and observation of pattern generated by these solutions in the table as well as the coordinate plane. They are suggested to turn a linear equation into the explicit form $y = f(x)$, with their recent skills of change of subject. Once putting in such form, samples of x and y may be easily computed and recorded in table or coordinate plane (Figure 13).

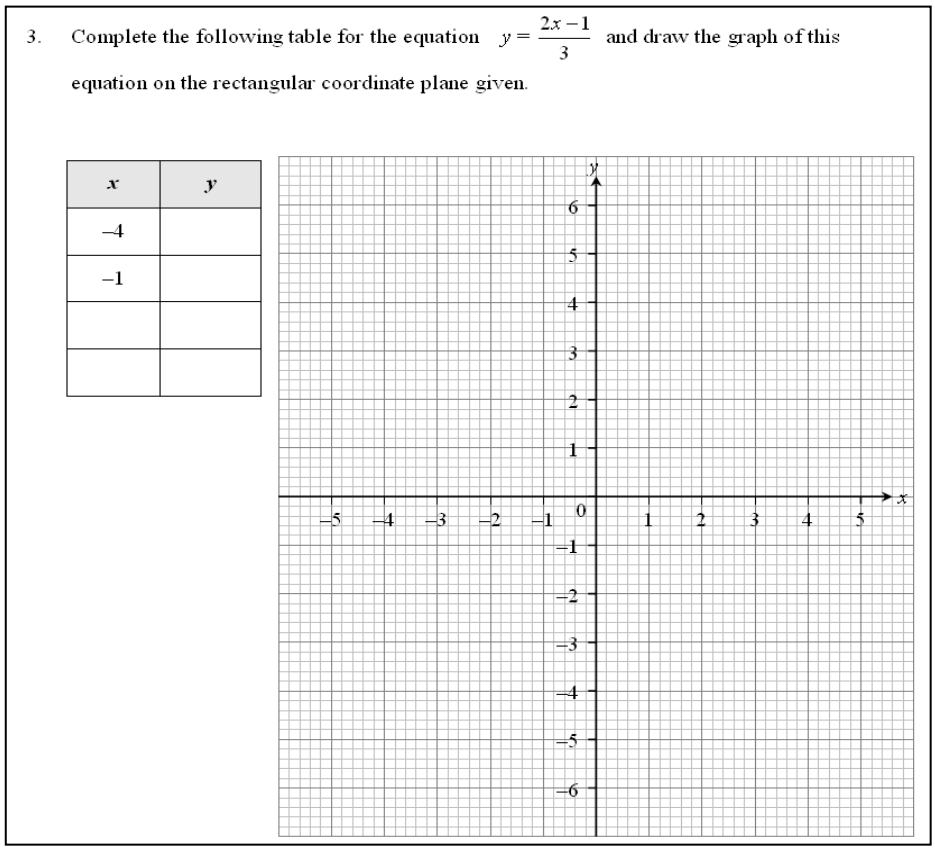


Figure 13

This approach does not merely provide an efficient and reliable method of generating graphs, which is usually emphasized in teaching this unit. Instead, the emphasis is put on noticing spatial and numerical patterns resulting from deliberate choice of equally spaced *x* values. This is intended to suggest why a line graph should result from a linear equation by relating progressions in *x* and *y* values generated in this way. In other words, we attempt to explain linearity of graphs and equations.

Research Lesson 3

Object of learning: linearity of $ax + by + c = 0$
Critical aspects: equation of the graph of a line can be written in the form $ax + by + c = 0$

Outline of lesson:

1. Ask students to guess whether the equation has a linear graph or not without plotting.
2. Check with students by plotting the graph in GeoGebra.
3. Summarize with students the characteristics of equations with linear graphs ($ax + by + c = 0$).

The task in this research lesson helps students to distinguish linear and non-linear graphs and equations through examples and non-examples, some of which generated by students. They are encouraged to articulate forms of linear equations from examples experienced. Once again, the dynamic graphing tool naturally provides a means to quickly and flexibly test and check graphs of any equations imagined by the students or assigned by teacher (Figure 14).

	Equation	The graph of the equation is:	
		a straight line.	not a straight line.
(a)	$2y = x + 1$	✓	
(b)	$x + y = 6$	✓	
(c)	$2x + 3y - 5 = 0$	✓	
(d)	$y = x^2$		✓
(e)	$xy = 1$		✓
(f)	$x^2 + y^2 = 1$		

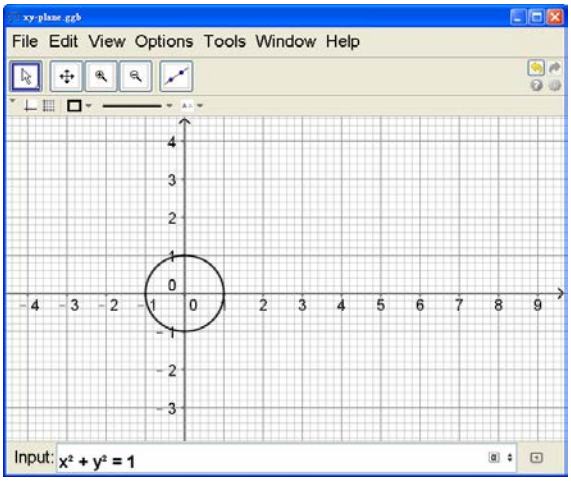


Figure 14

Research Lesson 4

Object of learning: write the equation of a straight line from its graph ($y = mx + c$, m and c are integers)

Critical aspects: relationship between x and y coordinates of points on the graph

Outline of lesson:

1. Identify integral points from the graph.
2. List their x, y coordinates in a vertical table.
3. Observe number patterns from $y = mx$ to $y = mx + c$.

The task in the last research lesson is an enrichment task for those classes satisfactorily completing the previous lessons. It requires students to suggest equation for set of coordinates equally spaced along a line by looking at the pattern of values of x and y in a vertical table form (Figure 15). This task aims at helping the students to gain a sense of linearity from another way, that is, formulating equations from tables of values. We do not aim at systematically finding equations of straight lines which should be covered in Key Stage 4.

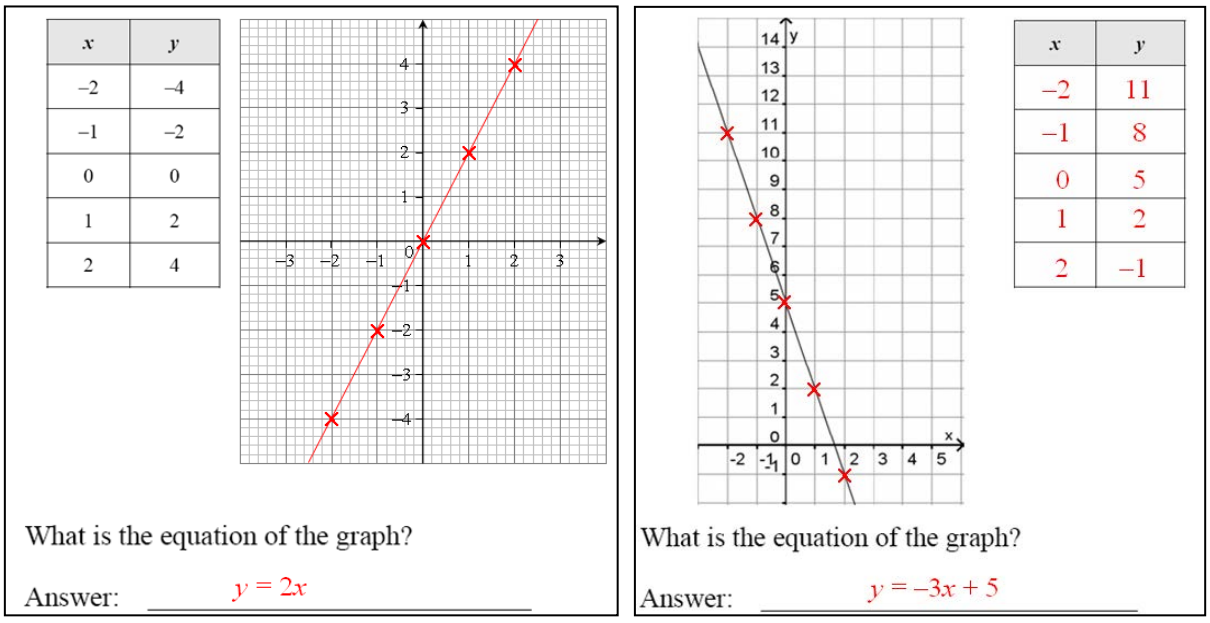


Figure 15

Evaluation and Reflection

The research lessons were carried out in all secondary two classes in the second semester of the school year 2010-2011. Most of the teachers involved and advisers observed the lessons and conducted post-lesson evaluations. Although there were comments on lesson design and implementation during evaluation after each lesson, the teachers preferred following more or less the same plan throughout the study, without any significant modification.

Most of the tasks could be completed satisfactorily although students' interaction varied from class to class. Teachers found the dynamic graphing tool useful and manageable. They appreciated the new approach, supported by technology, to vividly illustrate and discuss about various aspects of graphing and equations.

There were evaluation of students' performance and understanding through lesson observation, post-lesson written tests for all students and task-based interviews of selected students. It is found that the students could generally acquire the concept of equations and graphs, and demonstrate good skills in various tasks of graphing and solving equations. The basic concept and technique of matching points/solutions and equations were adequately

developed for most students. Regarding understanding of linearity, a range of interesting responses could be identified from the interviews and post-tests. Here are some observations.

- Most students could perceive linearity as “equidistance” of integer points in the graph.
- Expressing y in terms of x, the vertical table form and the “pattern” of integer points help students to find ordered pairs.
- Many students could determine whether the graph is linear or not from the equation, although they could not use the term “degree”, nor could explain why.
- Some weak students avoided to express y in terms of x.
- Some students still could not determine linearity of graphs from equations, but expressing y in terms of x may help students to recognize linearity of data.
- When asking whether the graph or the equation $y = 2x$ is linear or not, one student said no by writing it as $y/x = 2$, but then said yes by writing it as $2x – y = 0$.

Overall, although some students cannot explain how they determine linearity of equations and graphs and may rely on superficial clues in the forms of equations for recognizing linearity, the explicit form of equation, spatial and numerical patterns in square grids and tables are promising means to develop this sense of linearity. It is hoped that the case reported in this article could let us understand more on students’ learning difficulties, and could provide some useful insights into some effective pedagogical practices on this topic.

Authors’ e-mail: Arthur Man Sang Lee amslee@hku.hk
Anthony Chi Ming Or anthonyor@edb.gov.hk

Designing Tasks for Visualization and Reasoning in Dynamic Geometry Environment[1]

Anthony Chi Ming Or
Education Infrastructure Division, Education Bureau

Introduction

One of the important aims of the Hong Kong Secondary Mathematics Curriculum is to develop students' ability to conceptualize, inquire, reason and communicate (Curriculum Development Council, 1999, p.4). Hence the terms "explore" and "justify" appear in the learning objectives of many topics. For example, it is expected that students could "explore the formula for the area of circles" (ibid., p.20), or could "explore and justify the methods of constructing centres of a triangle such as in-centre, circumcentre, orthocentre, centroids, etc." (ibid., p.23).

However, it seems that the learning tasks proposed in Hong Kong textbooks could not always fulfill this aim. For instance, when constructing the circumcircle of a triangle, students are told directly to first construct the perpendicular bisectors of the three sides using rulers and compasses, and then use their intersection as a centre to draw a circle passing through one of the vertices, and finally see that this circle also passes through the other two vertices (Figure 1). In this task students are neither provided the opportunity to explore nor asked to justify how the circumcircle could be constructed, but instead just to verify the correctness of the procedures to construct the circumcircle given by the textbooks.

---

[1] This article is a slight revision of a paper presented in the International Commission on Mathematical Instruction Study Conference 22: Task Design in Mathematical Education in Oxford, UK in July 2013. The paper was published online in the conference proceedings at http://hal.archives-ouvertes.fr/hal-00834054.

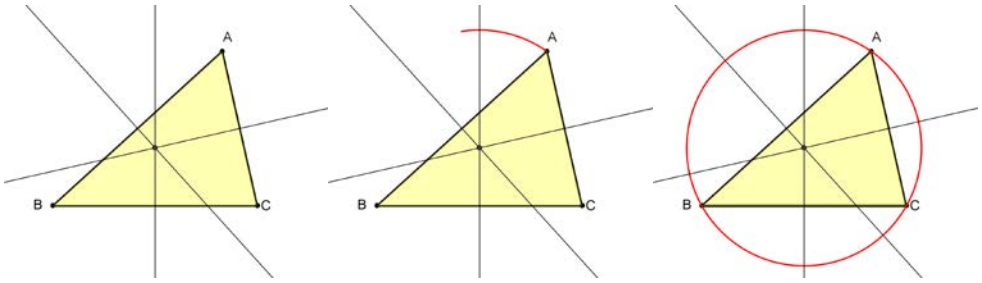


Figure 1

In this article, I am going to present some explorative tasks designed in GeoGebra, a powerful dynamic geometry (DG) freeware, that could facilitate justification through visualization. I shall propose a framework on task design in dynamic geometry environment (DGE) to facilitate visualization and reasoning based on Duval's model of the role of visualization in the development of geometrical reasoning (Duval, 1998). In particular, I suggest that the use of soft constructions (Healy, 2000) is an effective approach for designing tasks to foster operative apprehension for visualization and reasoning in DGE.

Duval's Model of Geometrical Reasoning

Duval (1998) suggests that geometry involves three kinds of closely connected cognitive processes fulfilling specific epistemological functions, namely, *visualization*, *construction* and *reasoning*. Their epistemological functions and connections are represented by Figure 2 below.

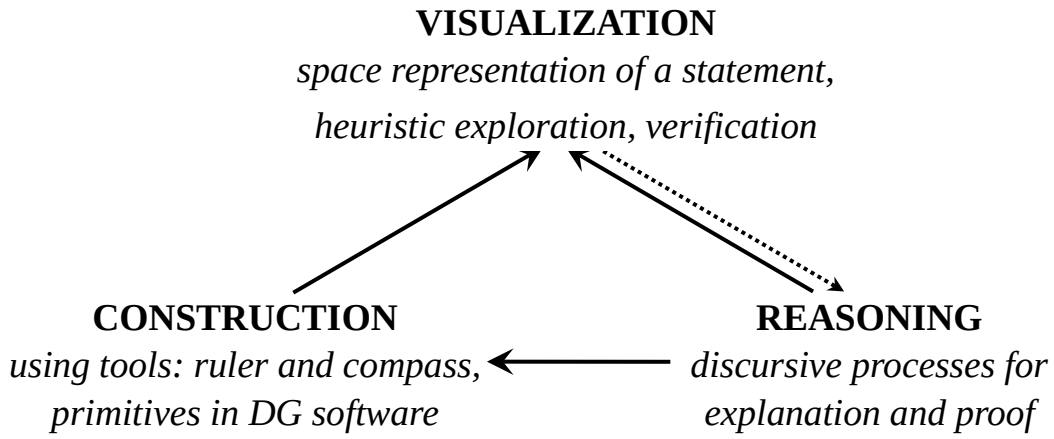


Figure 2

In the figure each arrow represents the way a kind of cognitive process can support another kind in a task. The dotted arrow suggests that visualization does not always help reasoning. For example, visualization can be misleading if our visualized image is a special case. Duval states that these three kinds of cognitive processes are quite different and must be developed separately, and the significance of the teaching of geometry is to develop visual representation and reasoning abilities and to favour the synergy of these processes.

To facilitate visualization as well as reasoning, Duval suggests the necessity of a kind of apprehension of geometric figures called *operative apprehension*, which means operations on the figure or its subfigure, either mentally or physically, that gives insight into the solution of a problem. He emphasizes that operative apprehension is crucial and teachers have to identify factors triggering or inhibiting it so as to make visualization possible and gives rise to various transfers.

With regard to the use of dynamic geometry software (DGS), Duval states that DGS provides enormous possibilities of visualization through the introduction of the aspect of movement, and allows manipulations of geometric objects and hence true explorations of geometrical situations. However, the construction-centered design of DGS does not develop all functions of visualization, in particular the operative apprehension.

Duval's theory emphasizes the importance of operative apprehension to facilitate visualization and also reasoning in the teaching and learning of geometry. In view of his comments on the uses and limitation of DGS in visualization, I shall discuss how to design tasks in DGE to foster operative apprehension for visualization and reasoning. I would first define what operative apprehension means in the DGE, and how the use of soft constructions proposed by Healy (2000) could be an effective approach to designing task to foster operative apprehension.

Operative Apprehension in DGE

A task is a set of pre-designed, environmentally situated materials aiming to engage learners in activities that could transform the ways they see and do mathematics (Leung, 2011). A task has to be pre-designed in the way that through these pre-designed means, learners are guided to construct insights and the meaning of the mathematics knowledge. A task is also environmentally situated, in the sense that the qualities or tools of the environment have been made use of to empower learners with extended or amplified abilities to acquire knowledge which could not be acquired in the same ways as in other environments (Leung, 2011). In what follows, I shall discuss how to design tasks situated in DGE to foster operative apprehension for visualization and reasoning in Duval's framework. In particular, I will focus on pre-designed DG figures and interpret Duval's operative apprehension as the following:

> Operative apprehension of a mathematical concept or problem in DGE is the insights into the concept or the solution of the problem revealed by operating on a pre-designed figure in the environment through dragging.

Let me illustrate the significance of operative apprehension in DGE using a task I designed in GeoGebra. This task originates from the following problem in a textbook.

> *A quadrilateral is dissected by a line joining the mid-points of one pair of opposite sides, and the perpendiculars to this line from the mid-points of the other pair of opposite sides.* (See Figure 3(a).) *What shape can you get from this dissection?*

A task is designed in GeoGebra to help learners to explore this problem (http://www.geogebratube.org/student/m3459). In this task, a quadrilateral is dissected into four pieces as described in the problem. Each piece can be rotated through dragging the red point at the vertex. In this way we can see that how the four pieces could form a rectangle (Figure 3). Also, the operation gives us the insights to reason why this dissection gives a rectangle, by, for

instance, thinking about why the four angles at the vertices give a sum 360° (Figure 3(d)).

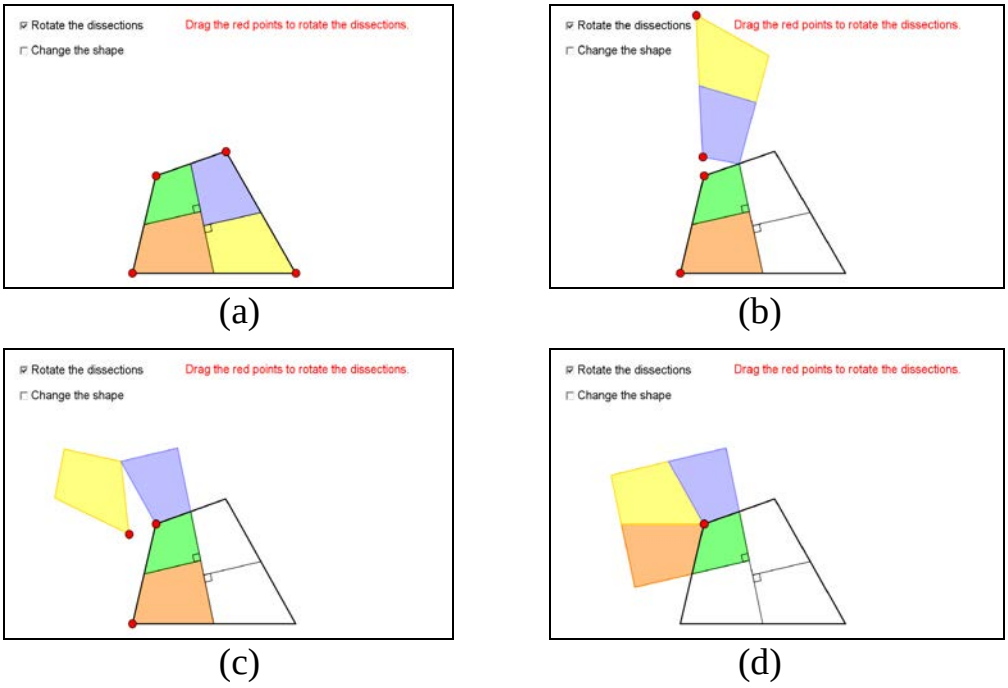


Figure 3

Besides rotating the four pieces, learners can also operate on the shape of the quadrilateral. After checking the "Change the shape" box, four green points appear at the vertices of the quadrilateral and the shape of the quadrilateral could be changed by dragging them. Through dragging the vertices, I see that the dissection would give a square for some shapes of the quadrilateral (Figure 4).

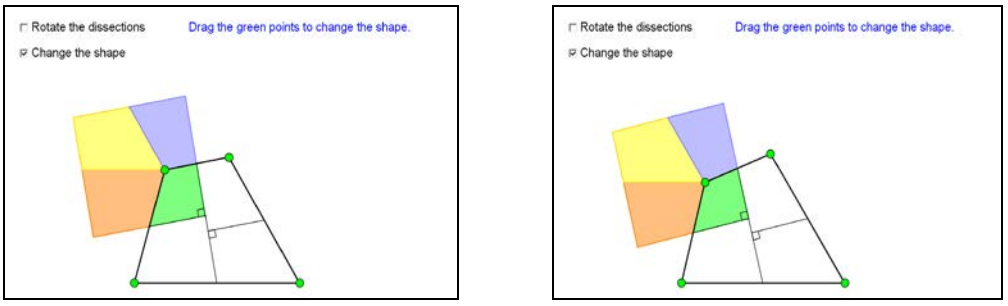


Figure 4

This problem reminds me the famous Haberdasher Puzzle composed by English mathematician Dudeney (Dudeney, 1907). This puzzle shows how an

equilateral triangle could be dissected into a square (Figure 5). Although I have known this puzzle for a long time, I never understand how Dudeney could think of this method of dissection, nor have any idea how to generalize his method to dissect an arbitrary triangle into a square.

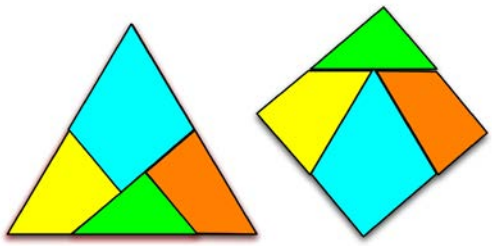
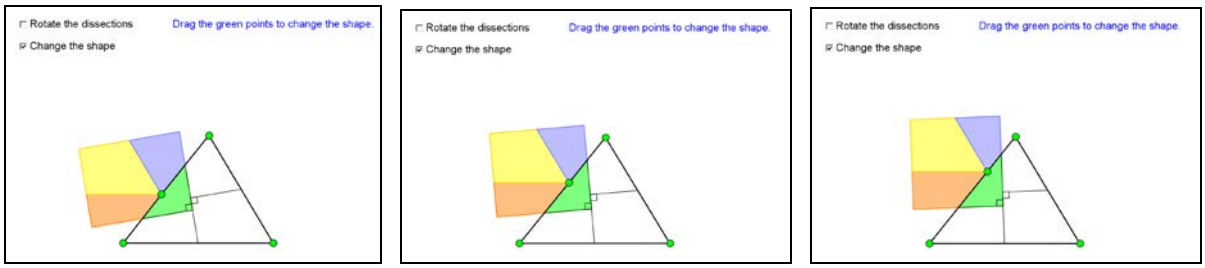


Figure 5

When I try to compare the quadrilateral problem with Dudeney's puzzle, I suddenly realize that if I drag a vertex, say the upper-left one, to a position at which it is collinear with the other two adjacent vertices (Figure 6(a)), the quadrilateral would be degenerated into a triangle which is dissected into a rectangle. Furthermore, if I drag this vertex along the side of the triangle (Figure 6(b)), the shape of this triangle is unchanged and the area of the rectangle is kept constant while its length and width are decreasing and increasing respectively through dragging. Hence I should get a square somewhere on this side (Figure 6(c)).



(a)　(b)　(c)

Figure 6

After the above exploration I see how an arbitrary triangle could be dissected into rectangles of various sizes, and there should be a particular dissection that gives a square. Through operating on the shape of the quadrilateral, I get important insights of comprehending how a general triangle could be dissected into a rectangle, and investigating when the dissection would give a square.

This example illustrates the advantage of fostering operative apprehension in DGE. If we use a paper quadrilateral, although we could cut it to see how it could be dissected into a rectangle, it is impossible for us to operate on its shape. In DGE we can operate on the shape of the quadrilateral so that we can degenerate it to a triangle to get the insights of how a triangle could be dissected into a rectangle.

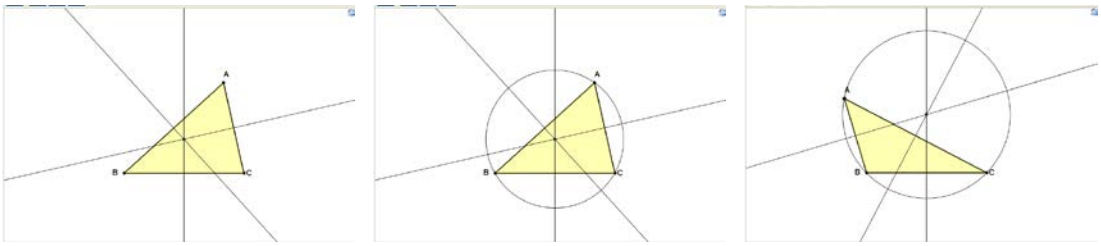
Operative Apprehension for Visualization and Reasoning: Soft Construction

At the beginning of research in dynamic geometry, tasks in robust constructions, i.e. constructions preserve relationships upon dragging, were recognized as promoting for the learning of geometry. However, Healy (2000) discovered through observation that, rather than robust constructions, students preferred to investigate constructions "in which one of the chosen properties is purposely constructed by eye, allowing the locus of permissible figures to be built up in an empirical manner under the control of the student". Healy called these constructions *soft constructions*.

Healy differentiates the roles of dragging in robust and soft constructions. In a robust construction, dragging provides a visual verification of the validity of the construction through dragging. In a soft construction, dragging is not verification but part of the construction itself. Through dragging, the general can emerge from the specific by searching empirically for the locus of figures fulfilling the given conditions. Soft constructions offer a transition from an empirical approach to a theoretical approach in solving a geometry problem.

In the lens of Duval's model of geometrical reasoning, tasks in robust and soft constructions can be considered as operative apprehension on figures serving different functions of visualization: a robust construction provides a verification of the construction, while a soft construction provides heuristics or insights through an empirically searched locus which mediates reasoning. I shall illustrate this point with two GeoGebra tasks of drawing the circumcircle of a triangle, one in robust construction and one in soft construction.

In the robust construction task, perpendicular bisectors of the three sides are first constructed using the "Perpendicular Bisector" tool  . A circle centred at their intersection (found by the "Intersection" tool ) and passing through either one vertex (say *A*) is constructed using the "Circle" tool  , and it could be seen that this circle also passes through the other two vertices (*B* and *C*). By dragging the vertices of the triangle, learners can check the validity of the construction by seeing that the circle always passes through the vertices. They can also see that the circumcentre lies outside the triangle when the triangle is obtuse (Figure 7).



(a)　(b)　(c)

Figure 7

In the soft construction task (http://www.geogebratube.org/student/m3958), learners are first given the triangle and a circle which can be moved by dragging its centre (in red) and a blue point on its circumference (Figure 8(a)). Learners first drag the blue point to either one vertex, say *A*, and a dotted line joining *A* and the centre would then be shown (Figure 8(b)). Then they drag the red centre to different positions at which the circle also passes through another vertex *B*, and when this happens a dotted line joining *B* and the centre would be shown. These positions of the centre are marked in red, and learners can see that the locus of the centres of circles passing through *A* and *B* is a straight line (Figure 8(c)). Learners can then be asked what this line of locus should be, and the two dotted lines from the centre to *A* and *B* providing hints for them to reason that this line is the perpendicular bisector of *AB* (through looking at two congruent triangles). Once they recognize that the locus of the centres should be the perpendicular bisector, they can find empirically the loci of the centres when the circle passes through *A*, *C* (Figure 8(d)) and *B*, *C* (Figure 8(e)), and finally visualize that the circumcircle should centre at the intersection of the

three loci, i.e. the intersection of the perpendicular bisectors of the three sides (Figure 8(f)).

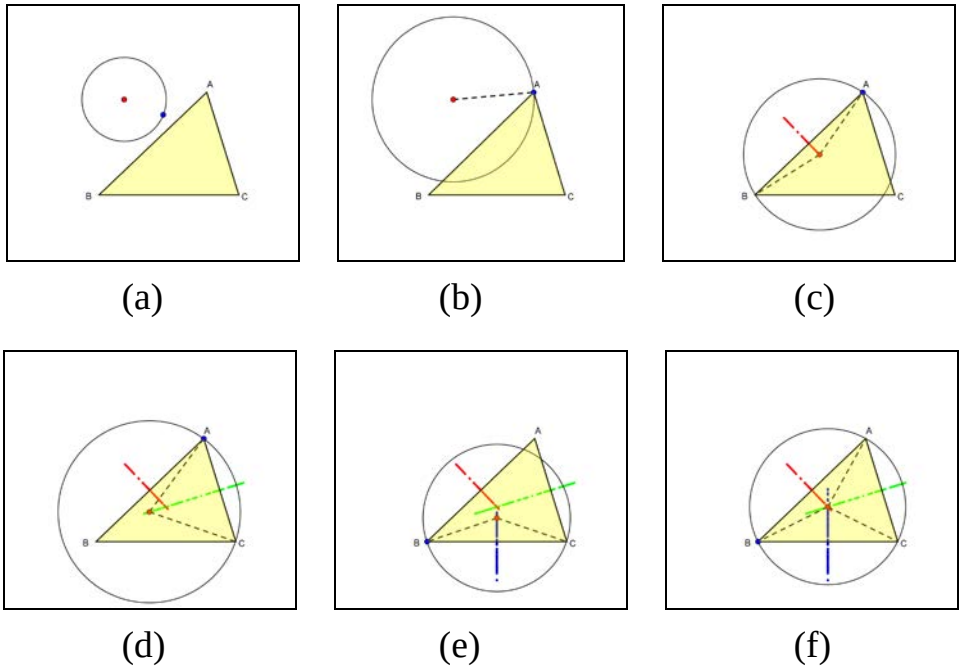


Figure 8

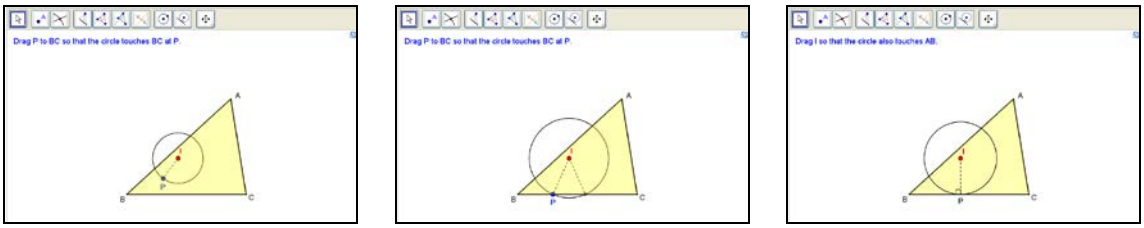
The above example illustrates how a task in soft construction could foster operative apprehension by recording the loci of positions at which the eye construction satisfies the given conditions. These loci of positions provide insights to solve the problem, and also mediate the reasoning of why the problem could be solved in this way. I now propose the following principle of using soft constructions to design task fostering operative apprehension for visualization and reasoning in DGE.

Principle of using soft constructions to foster operative apprehension

> Learners are provided opportunities to perform soft (eye) construction by dragging. The loci of the dragging satisfying the given conditions, together with the other elements supporting their visualization and reasoning, would be shown to the learners so that theoretical elements could emerge from the empirical evidences.

I further elaborate the above principle using a more sophisticated task of finding the incircle of a triangle (http://www.geogebratube.org/student/m4363).

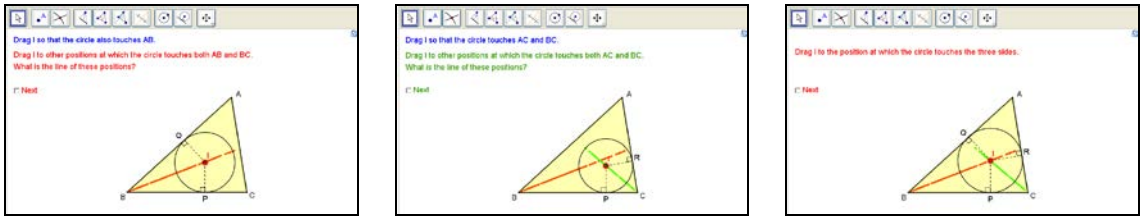
In this task the triangle and a circle of centre *I* and passing through *P* are given, and the radius *IP* is also shown. Learners are first asked to drag *P* to the side *BC*, then another dotted line would be shown to indicate that there are two intersections (Figure 9(b)). By dragging *P* towards the other intersection learners would visualize that for the circle to touch *BC*, the two dotted radii should overlap to form one radius *IP* perpendicular to *BC* (Figure 9(c)). I also anticipate that this process of dragging, together with the overlapping of the two radii, would help learners to reason why the tangent of a circle should be perpendicular to the radius.



(a) (b) (c)

Figure 9

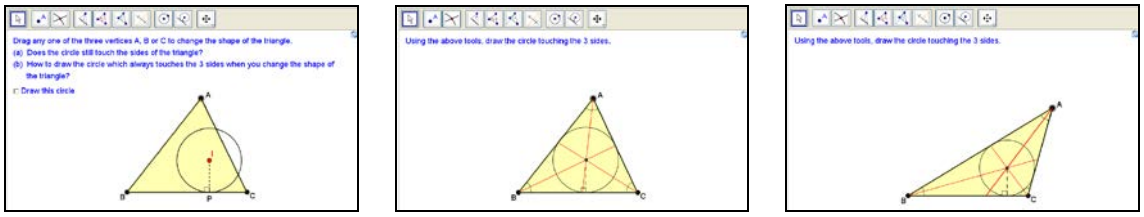
Once the circle touches *BC*, *P* can no longer be dragged and learners are asked to drag the centre *I* to different positions so that the circle would also touch *AB*, and the locus of *I* is marked in red (Figure 10(a)). Learners are prompted to identify this line of locus as the angle bisector at *B*, and could explain this by looking at the congruent triangles *IBP* and *IBQ*. Similarly learners identify the locus of *I* at which the circle touches *BC* and *AC* as another angle bisector at *C* (Figure 10(b)), and see that the circle would touch the three sides when *I* is at the intersection of the angle bisectors (Figure 10(c)).



(a) (b) (c)

Figure 10

Finally, the three vertices of the triangle are made draggable to the learners and they are asked to drag the vertex *A* to change the shape of the triangle, and see that the original circle no longer touches the three sides (Figure 11(a)). They are then asked to perform robust construct of the incircle by constructing the suitable lines in a triangle using the given tools (median  , angle bisector  , altitude  and perpendicular bisector ) and the touching circle tool  (Figure 11(b)). They can then check the validity of their construction by dragging the vertices (Figure 11(c)).



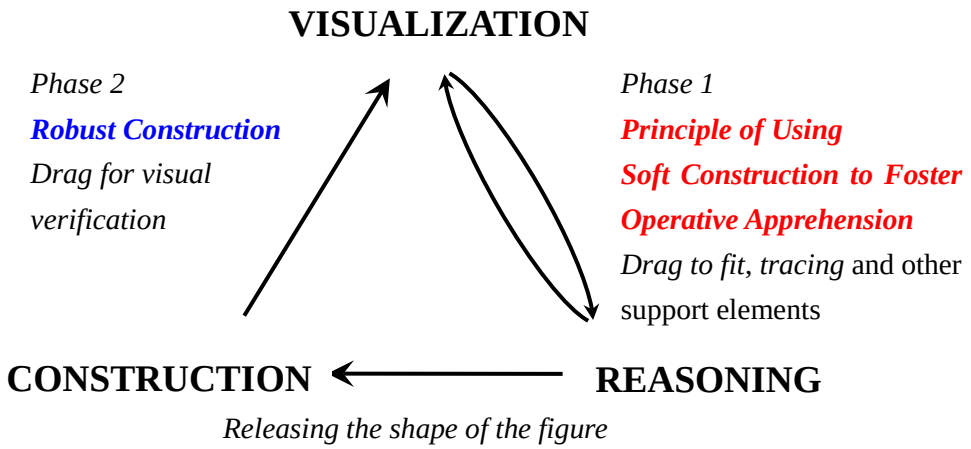
(a) (b) (c)

Figure 11

Discussions and Implications

Based on the above illustrations, I propose a model of task design in DGE to foster operative apprehension for visualization and reasoning by modifying Duval's model of geometrical reasoning as follows:

Task Design Model in DGE for
Visualization and Reasoning through Dragging



Task design in this model consists of two phases. In Phase 1, the *Principle of Using Soft Construction to Foster Operative Apprehension* is applied so as to foster students' operative apprehension through *soft construction*, i.e. to use the *drag to fit* strategy to find solutions satisfying the given conditions. In the process of soft construction, the *trace* of the *locus of validity* (Leung and Lopez-Real, 2002) and other support elements that mediate reasoning would be shown. Use my in-centre task as an example (p.7), the dotted radii and their overlapping through dragging (Figure 9) are the support elements which are shown to students to mediate the insight and reasoning of perpendicularity of the radius and the side when the circle touches it. Similarly, the traces and the radii shown by the software in Figure 10 support the reasoning that the traces are the angle bisectors of the triangle and that in-centre lies on their intersection.

In this phase dragging and tracing are the *cognitive tools* (Leung, 2011) to start a recursive cycle between visualization and reasoning until a solution and its justification is reached. In the design of the in-circle task, students are guided to first visualize through dragging that the radius has to be perpendicular to the side when the circle touches it. Then with this property students are further guided to visualize through dragging and tracing that centre of the circle must lie on a certain line when the circle touches two sides of the triangle. They are then guided to reason, using the trace and the dotted radii, and explain that this line is in fact the angle bisector. Finally they further visualize that when the centre lies on the intersection of the two angle bisectors, the circle would touch all the three sides and at this stage they should be able to explain why this happens.

When the solution and its explanation are reached in Phase 1, the task is then transited to Phase 2 in which students are required to use the construction tools given by the software to do a *robust* construction to verify the solution and explanations they obtained in Phase 1. This is done by releasing the shape of the figure in the problem so that students observe that the soft construction in Phase 1 no longer works when the shape of the figure is changed (Figure 11(a)). Students are then asked to use the tools of the software to construct a robust in-circle that always touch the three sides (Figure 11(b)(c)). In this phase dragging is a tool for visual verification of the construction.

This model shows how the different roles of robust and soft constructions could foster operative apprehension, through which the synergy of visualization, reasoning and construction can be facilitated. If we agree with Duval that developing visualization and reasoning abilities to favour the synergy of the three cognitive processes is of crucial importance for the teaching of geometry, designing tasks to foster operative apprehension for visualization and reasoning in DGE effectively would then be very promising to promote the teaching of geometry. This is also a great challenge to all teachers, educators and researchers. It is hoped that the principle and the model of task design in DGE proposed in this article could provide a useful initiation for further discussions, challenges and refinement in future task design research.

References

Curriculum Development Council (1999). *Syllabuses for Secondary Schools: Mathematics (Secondary 1 – 5)*. Hong Kong: The Education Department.

Dudeney, H. (1907). *The Canterbury Puzzles*. London: Nelson, 1907. Reprinted Mineola, NY: Dover, 1958.

Duval, R. (1998). Geometry from a Cognitive Point of View. In C. Mammana & V. Villani (Eds.), *Perspectives on the Teaching of Geometry fro the 21st Century* (p.37–52). Dordrecht: Kluwer Academic Publishers.

Healy, L. (2000). Identifying and explaining geometrical relationship: interactions with robust and soft Cabri constructions. In T. Nakahara and M. Koyama (Eds.), *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education*, (Vol.1, pp. 103–117). Hiroshima: Hiroshima University.

Leung, A. & Lopez-Real, F. (2002). Theorem justification and acquisition in dynamic geometry: A case of proof by contradiction. *International Journal of Computers for Mathematical Learning*, 7, pp. 145-165.

Leung, A. (2011). An epistemic model of task design in dynamic geometry environment. *ZDM Mathematics Education*, Vol.43, pp. 325–336.

Author's e-mail: anthonyor@edb.gov.hk

針對學習難點的教學設計——四邊形的特性

程翠娟、柯志明
教育局教育基建分部

引言

香港的小學數學課程中，四邊形這個課題出現於二年級至四年級，橫跨兩個學習階段。由二年級以直觀認識長方形、正方形、梯形及菱形及它們的不同構作方法，接著於三年級認識平行四邊形的簡單特性，到四年級的認識及比較各種四邊形的特性及用不同的方法製作四邊形等，當中涉及視覺（visual）、描述（descriptive）以致理論（theoretical）等幾何思維的不同層次（van Hiele, 1986），是小學課程的一個重要課題。

當我們查閱全港性系統評估（Territory-wide System Assessment, TSA）的一些相關題目的學生表現時，發現部份題目的全港答對率偏低。例如 2008 年的 6MC2 第 42 題及 6MC4 第 36 題（圖一）：

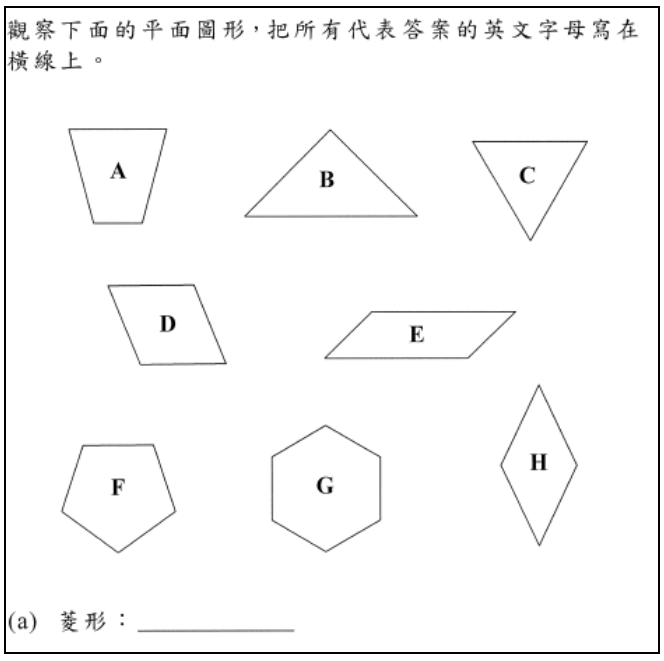


圖　一

這一題（a）部的全港答對率僅爲 19.7%，即全港少於五分之一的學生

能答對，這絕不是一個理想的答對率。答對率如此低的原因是大部分學生只能判斷 H 爲菱形，而漏填 D，相信他們都以爲 D 是平行四邊形，因爲 D 爲常見的平行四邊形的姿態。換句話說，即學生只以圖形的姿態來判斷，缺乏一套實在的判斷方法。

學生在從四邊形的特性判斷其可能形狀方面的表現一般，例如在 2007 年 6MC2 第 34 題（圖二）：

有兩種四邊形具備以下特性：兩組對邊平行及四條邊長度相等。

(a) 寫出這兩種四邊形的名稱。

(i) _____形　　(ii) _____形

圖　二

（a）（i）的全港答對率有 77.1%，但（a）（ii）的答對率則跌至 57.7%，少了差不多 20%，而當中不少學生漏答。我們估計不少學生可能只想像得到一種能滿足所示的條件的四邊形。

在方格紙上構作四邊形的表現有時也不能令人滿意。例如 2008 年 3MC1 第 28 題及 3MC4 第 27 題（圖三）：

在下圖中加上兩條直線，與原有的兩條直線組成一個平行四邊形。

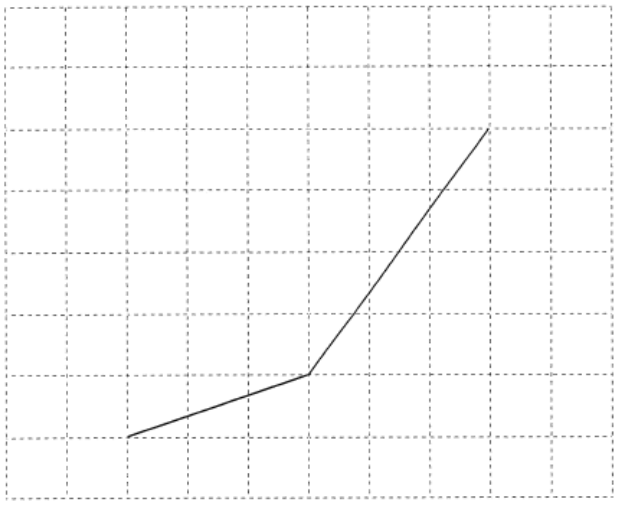


圖　三

這一題的全港答對率只有 37.3%。從考評局隨機抽取的答卷中，我們發現很多學生並不能準確地畫出平行邊（圖四）：

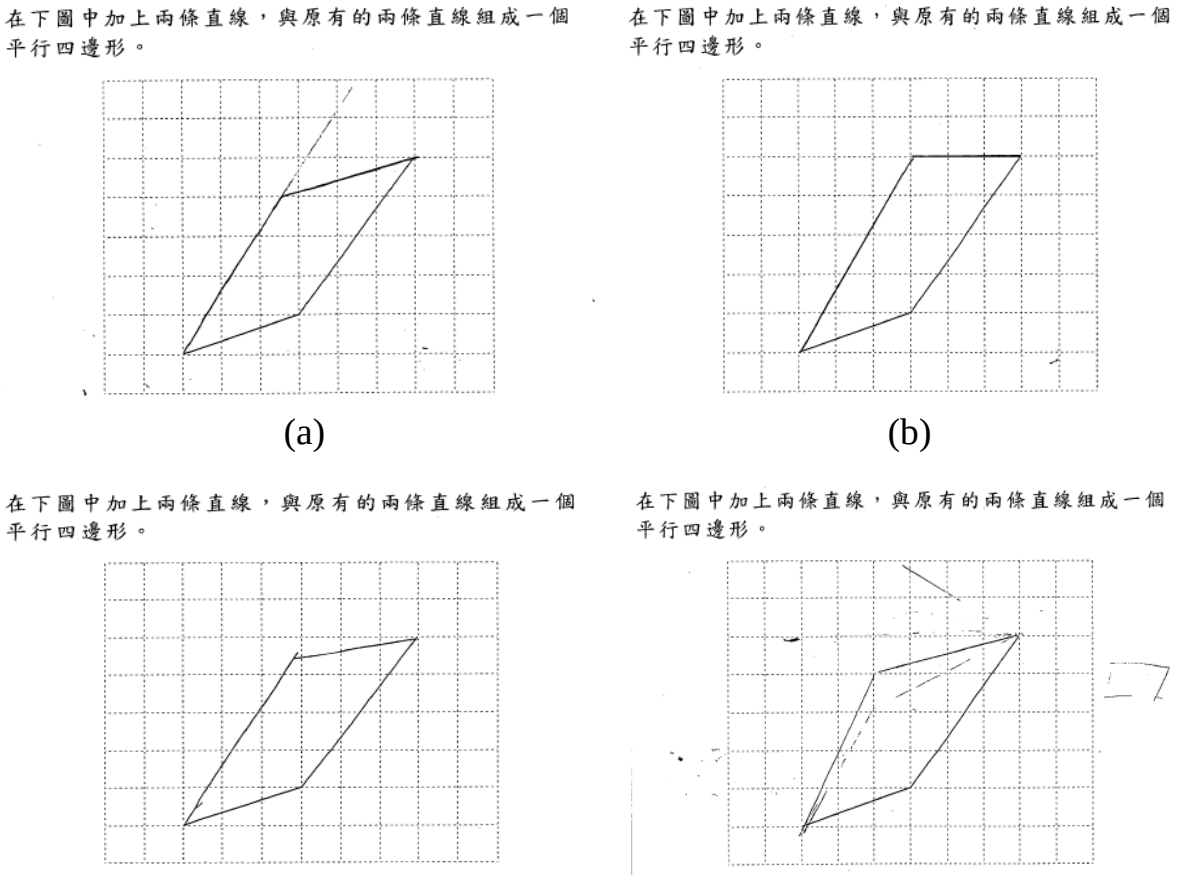


圖　四

有見及此，我們設計了一份診斷課業，透過學生訪談，了解學生在這課題的學習難點，然後針對這些學習難圖設計了一套課業。爲使課堂的討論更加互動，我們利用免費的動態幾何軟件 GeoGebra 製作動態課業供老師於課堂使用。這套的工作紙及動態圖形檔可在教育局的「網上學與教支援」網頁 http://wlts.edb.hkcity.net 裏的 KS2-S2-1 單元中找到。而本文旨在剖釋當中的設計理念如何能幫助學生學好此課題。

學習難點

學生訪談是了解他們學習難點的有效方法。我們設計了一份診斷課業，並挑選了九位學生進行訪談。綜合訪談的結果，我們發現學生有以下的學習難點：

（一） 學生大致能夠辨認已知圖形的特性，卻未能以圖形的特性來辨認圖形。他們往往只「直觀地」以圖形的姿態（orientation）來判斷；

（二） 學生能利用直尺判斷兩條邊是否相等，卻對判斷兩隻角是否相等感到困難；

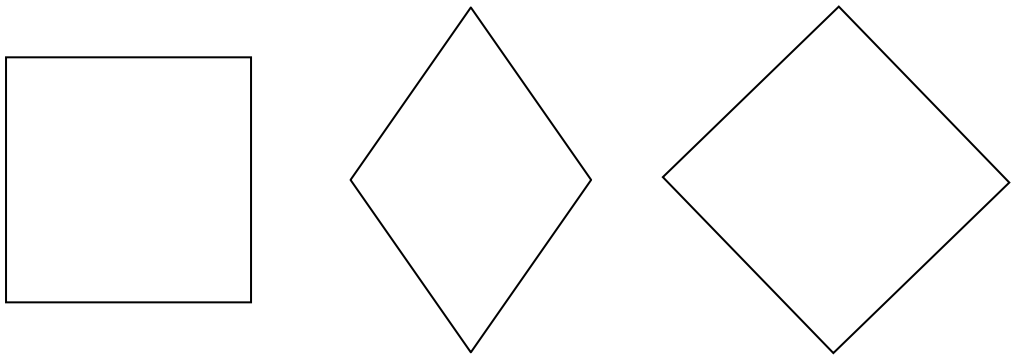
（三） 有學生將「平行」和「垂直」分別理解爲「水平」和「鉛垂」，如以爲平行線是（＝），垂直線是（||），或以爲垂直線是單一直線（|）；

（四） 絕大部份學生都缺乏在方格紙上繪畫平行線及垂直線的技巧及方法。

針對這些學習難點，我們設計了一套課業，讓學生先歸納四邊形的特性，再學習如何利用四邊形的特性把四邊形分類，並在方格紙或釘板上構作不同的四邊形。

教學設計一：由四邊形辨認其特性

二年級學生學習四邊形這個課題時，我們只要求他們能直觀辨認長方形、正方形、梯形及菱形。換句話說，即期望學生能夠一看該圖形就能說出其名稱。故此，學生所見的四邊形一般都是以常規的姿態出現，例如正方形和菱形必以圖五（a）和圖五（b）的姿態出現。所以當學生見到以非常規姿態出現的圖形如圖五（c）的正方形，學生便就會以它的姿態而判斷它爲菱形。這情況也有在訪談的過程中出現。



(a) (b) (c)

圖 五

故此在四年級開始學習這個課題時，我們把一大堆的四邊形讓學生以直觀的方式辨認出來（圖六）。對於一些常見的四邊形 (如 R)，大家都很容易得到它是正方形的共識。可是對於一些非慣常形態的四邊形如 M 和 N，有學生會以它們的姿態分別誤判爲菱形及平行四邊形。教師可利用這些具爭議性的四邊形，引導學生認識「直觀辨認」的不足之處，並思考辨認四邊形及把它們分類的準則。

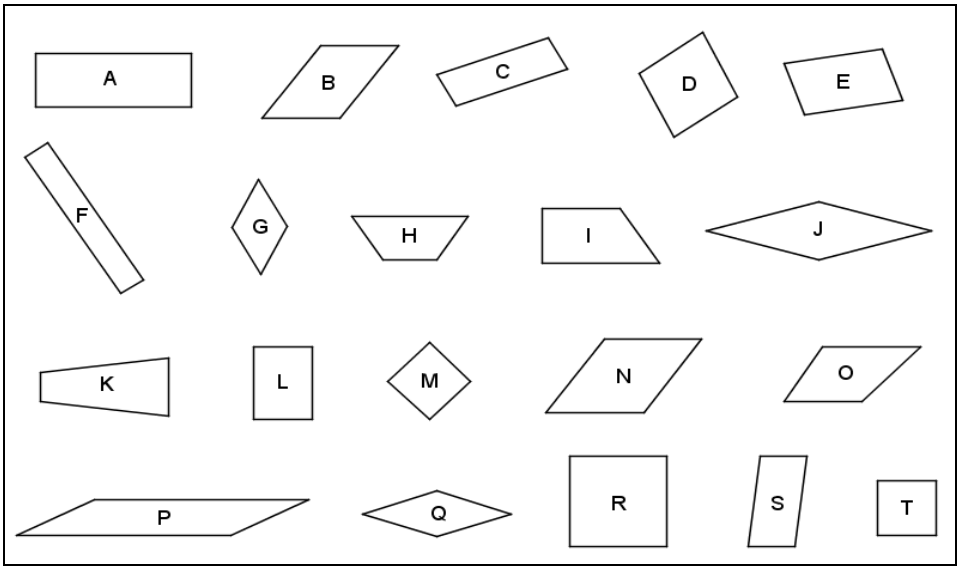


圖　六

要掌握判別四邊形的方法，先要認識各種四邊形的特性。教師可利用我們提供的簡報，先介紹或重溫如對邊、直角、對角和平行等詞彙，然後與學生一同探討工作紙的常規四邊形（圖七）的幾何的特性。教師應先和學生討論如何利用直尺探討對邊是否相等、是否有直角（以直尺的直角位來判斷）及平行邊（利用直尺量度兩線段的垂直距離）等特性，亦可運用提供的 GeoGebra 檔協助講解（圖八）。

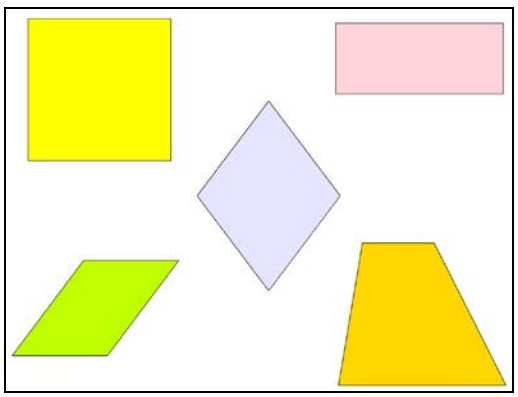


圖　七

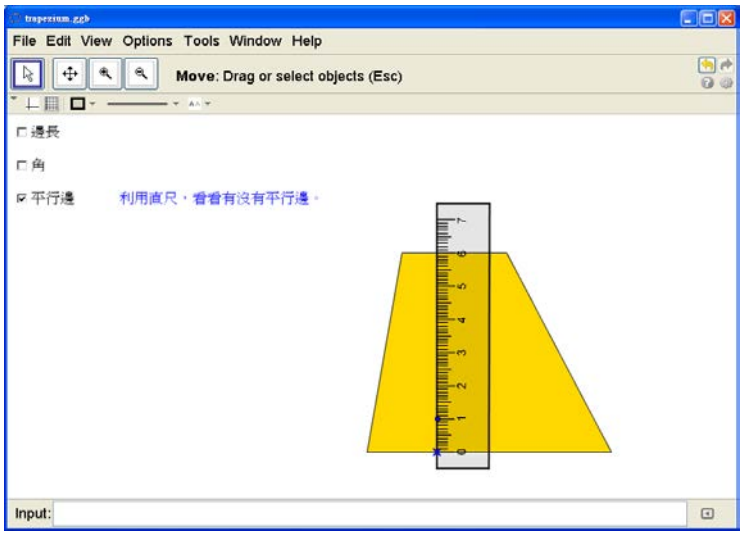
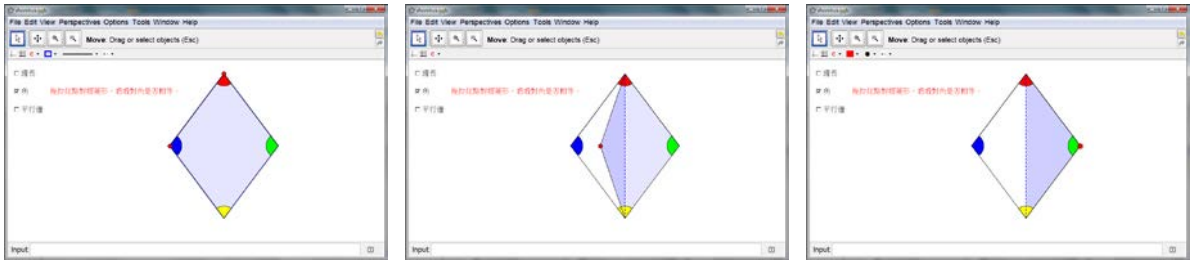


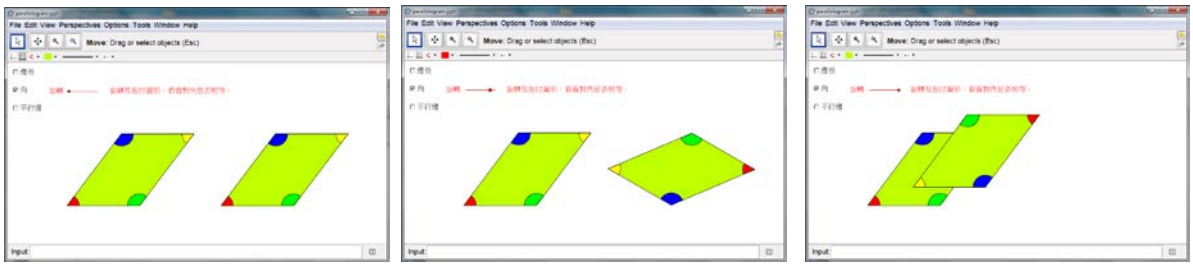
圖 八

要判斷四邊形的對角是否相等，由於沒有量角器，我們未能以量度的方法處理。我們建議把工作紙的常規四邊形剪出來，然後透過對摺（圖九）或旋轉後重合（圖十）等方法，得出菱形及平行四邊形對角相等的結果。這個做法的好處是利用了對稱性質（軸對稱及旋轉對稱）解釋了爲何菱形及平行四邊形的對角必定相等，既不需要使用超出課程要求的量角器，更把學生們的認知由「驗證」提升到「非正式論證」（informal deduction）的層次。這方法亦解釋了之前以直尺量度得出的「對邊相等」及「四邊相等」等結果，更爲學生將來學習軸對稱及旋轉對稱（增潤課程）鋪路。



(a) (b) (c)

圖 九



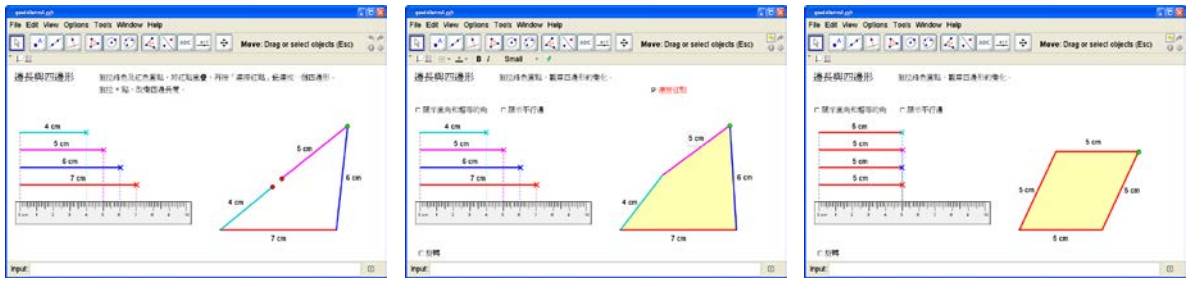
(a) (b) (c)

圖 十

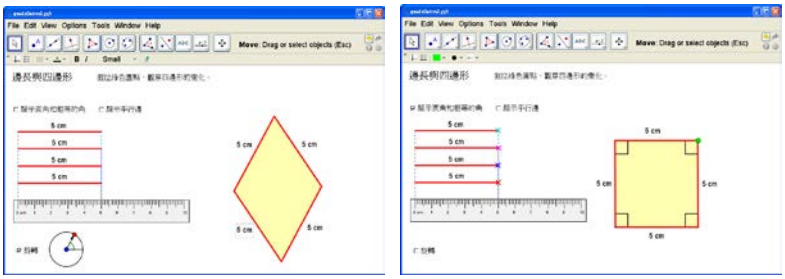
教學設計二：以特性把四邊形分類

當學生認識了各種四邊形的特性後，接著的問題就是如何利用這些特性辨認四邊形及把它們分類。其實四邊形的邊長、角和平行等特性有着因果關係，只要我們知道一個四邊形的某些特性，就足夠我們判斷它是甚麼四邊形，從而知道它的所有特性。由於直尺是學生惟一可以使用的量度工具，而且經過前述的課業活動後，他們已十分熟悉如何利用直尺去判斷相等邊、直角和平行邊等特性，所以我們設計了一個利用邊長和直角等特性去辨認四邊形的課業「邊長與四邊形」(又稱爲「電子幾何條」)，透過軟件 GeoGebra 改變四邊形的邊長，讓教師動態地呈現在「四邊相等」、「兩組對邊相等」等條件下四邊形可能出現的形狀，幫助學生掌握利用邊長及直角等特性來分辨四邊形的方法。

舉例來說，教師可以先提問學生：「若一個四邊形有四條相等邊，這是一個什麼四邊形？」部份學生可能只會想像到正方形。教師接著可開啓 GeoGebra 檔（圖十一(a)），把四條邊連接起來後（圖十一(b)），再將四條邊改變成邊長相等（圖十一(c)）。然後老師再拖拉綠點，讓學生知道即使我們固定了四邊形的四條邊長，它的形狀仍可改變。



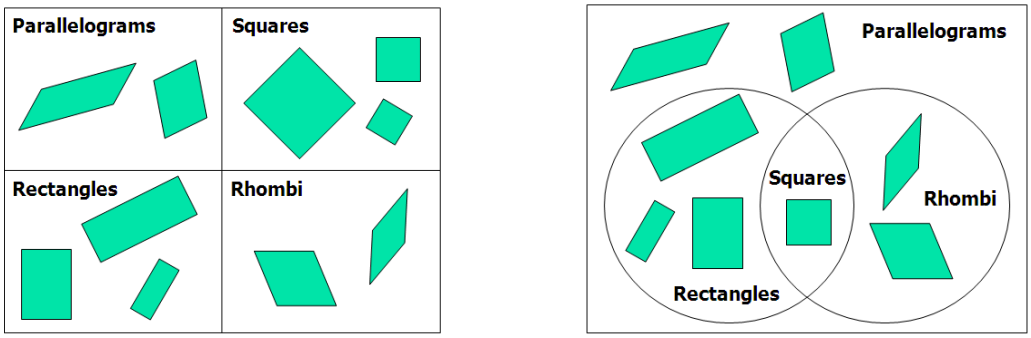
(a)　(b)　(c)



(d)　(e)

圖　十　一

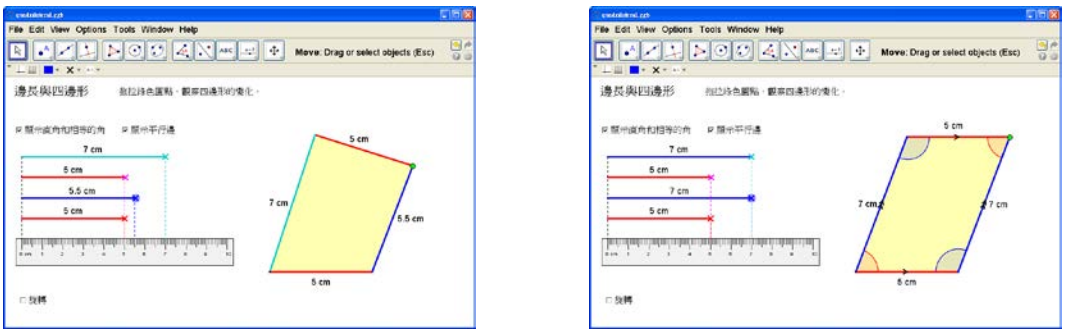
部份學生可能會以爲圖十一(c)是一平行四邊形，教師可把四邊形轉動成學生常見的姿態（圖十一(d)），讓他們知道這其實是一個菱形，從而教導學生應以四邊形的特性而非姿態來作判斷。接着老師提問學生：「四邊相等的四邊形除了菱形外，還可能是什麼圖形？」教師拉動綠點，使四隻角成爲直角（圖十一(e)）。由此學生除了可以明白到有正方形和菱形兩種四邊形能滿足「四邊相等」這個條件外，更能夠透過觀察由菱形到正方形的連續變化，認識到正方形只是菱形的一個特別情況（有直角的菱形），令他們較爲容易將四邊形的分類認知由割裂式分類（partition classification，圖十二(a)）過渡到層遞式分類（hierarchical classification，圖十二(b)。見 de Villiers, 1994）。



(a) (b)

圖 十 二

使用動態幾何軟件的另一個優點，就是能夠透過動態地改變圖形的一些已知條件（hypothesis），讓學生看到某些結果（conclusion）隨之出現，從而將因果關係（causal effects）視像化（Laborde, 2005）。例如當學生將四邊形的邊長由只有一組對邊相等（圖十三(a)）改變爲有兩組對邊相等時，他們就會發現圖形立即會出現兩組對角相等和兩組對邊平行的特性（圖十三(b)），從而認識到四邊形只要有兩組對邊相等，它就必然是平行四邊形了。



(a) (b)

圖 十 三

完成「邊長與四邊形」的活動後，我們把討論結果整合成一樹形圖（圖十四）。有別於常見的由一般到特殊的樹形圖，我們的樹形圖採用了由特殊到一般（generalization）的方法（de Villiers, 1994），讓學生由最特殊的「四邊相等」開始，逐步放寬到「兩組對邊相等」、「只得一組對邊平行」等較一般的條件，讓學生掌握以是否有相等邊和直角等特性去辨認四邊形，並把它們分類。

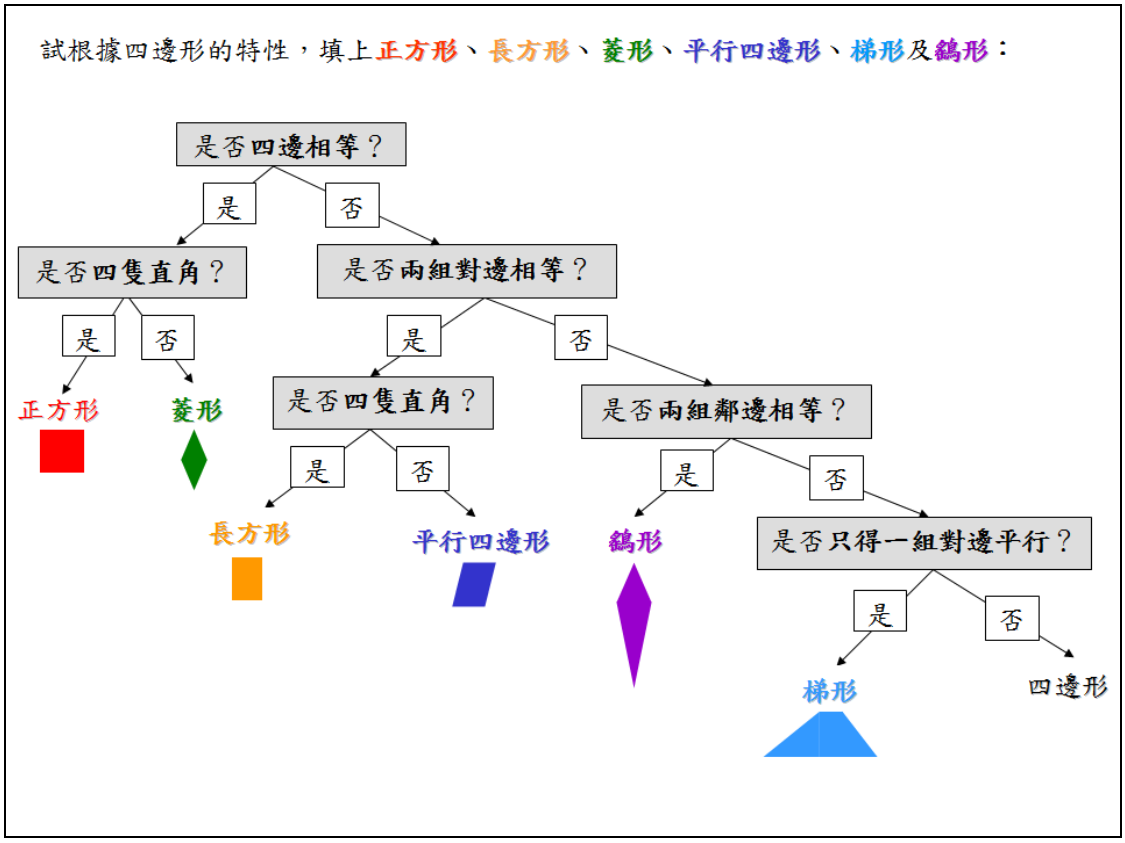


圖　十　四

學生掌握了利用邊長和直角去辨認四邊形的方法後，教師就可以回到之前將一大堆四邊形分類的課業，讓他們以所學的方法判別之前具爭議性的四邊形（如 M 和 N）。教師亦可利用所提供的 GeoGebra 檔和他們討論如何利用直尺去協助辨認四邊形，並透過 GeoGebra 檔的著色功能將他們分類（圖十五）。例如圖形 N，我們透過直尺量度知道它四邊相等（邊長 25 mm）。由於它沒有直角，所以它是一個菱形。

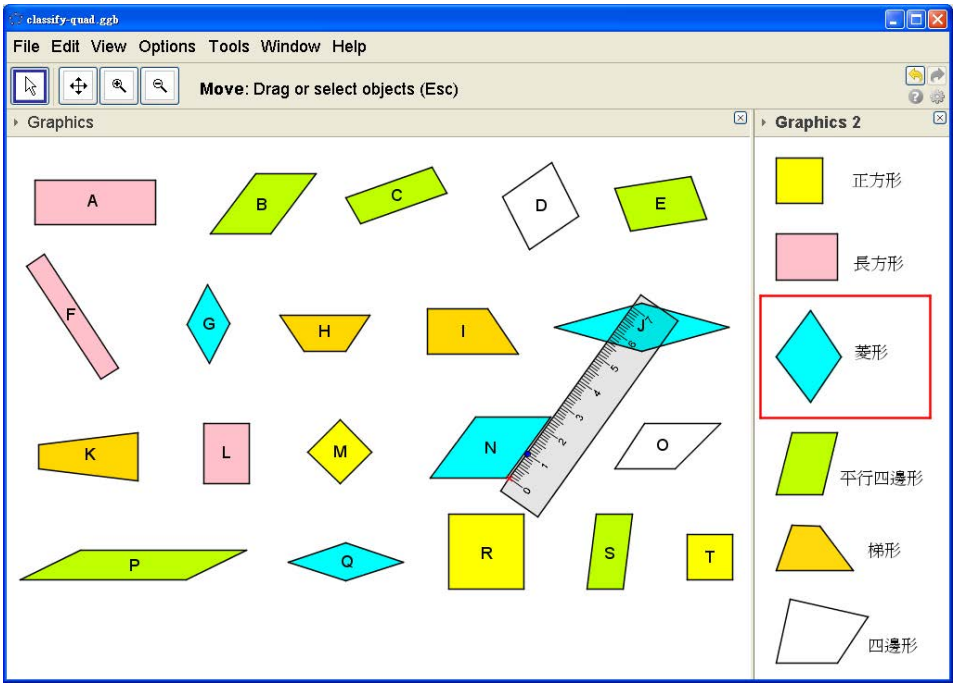


圖 十 五

教學設計三：在方格紙上構作四邊形

TSA 的數據反映學生在方格紙上畫平行四邊形的表現並不理想，最主要的原因是他們不懂得利用格線畫出與對邊平行的線。在訪談過程中，我們發現不少學生都有平行的概念，他們會憑感覺把直尺慢慢地向下移動到差不多的位置畫上平行邊，以致畫出來的平行線有欠準確。

爲幫助學生掌握利用方格紙的格線（或釘板的釘點）繪畫四邊形的方法，我們設計了一套工作紙，配合動態幾何 GeoGebra 檔使用。以繪畫平行四邊形爲例，我們著學生思考如何在一幅如圖十六般已有兩條直線的圖中加上兩條直線，組成一個平行四邊形。教師可利用 GeoGebra 檔，透過拖拉（圖十七(a)）及把直線所佔的方格著色（圖十七(b)），向學生展示如何透過數格準確地繪畫平行和相等邊。

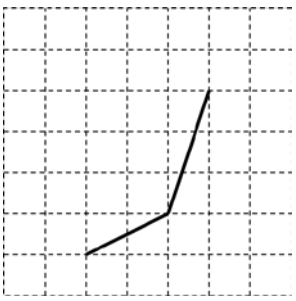
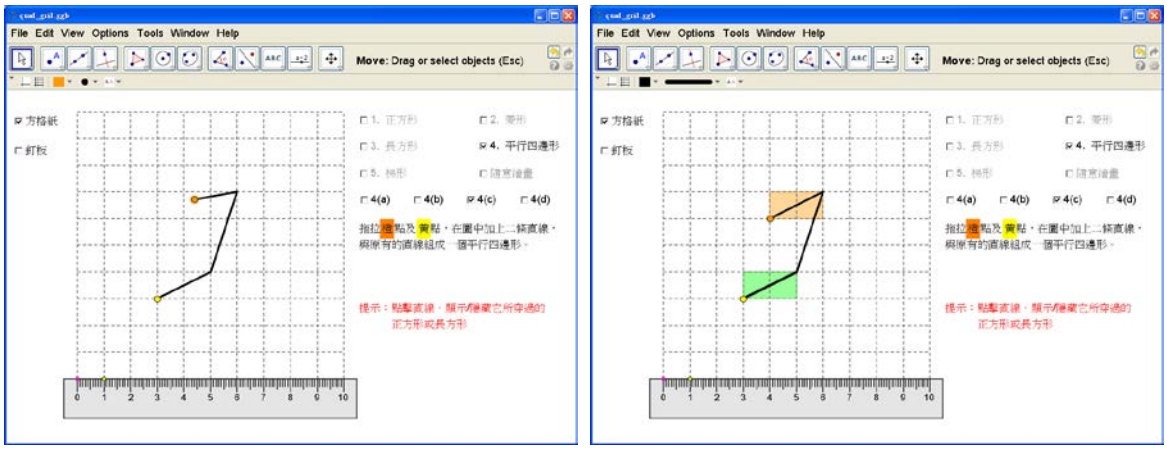


圖 十 六



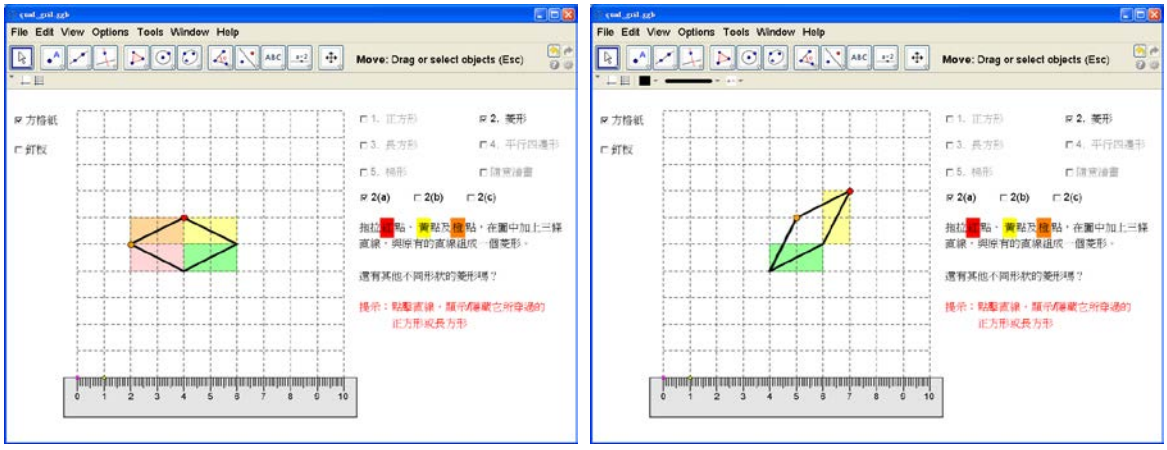
(a)　　(b)

圖　十　七

我們的繪畫菱形課業就更具挑戰性。在圖十八(a) 和 (b) 中，我們要求學生加上三條直線以組成一個菱形，並且著他們思考能組成多少個**不同形狀**的菱形。在講解圖十八(a)時，教師可以「馬行日」的象棋步法協助學生掌握繪畫三條邊的方法，如圖十九所示。

(a)　　(b)

圖　十　八

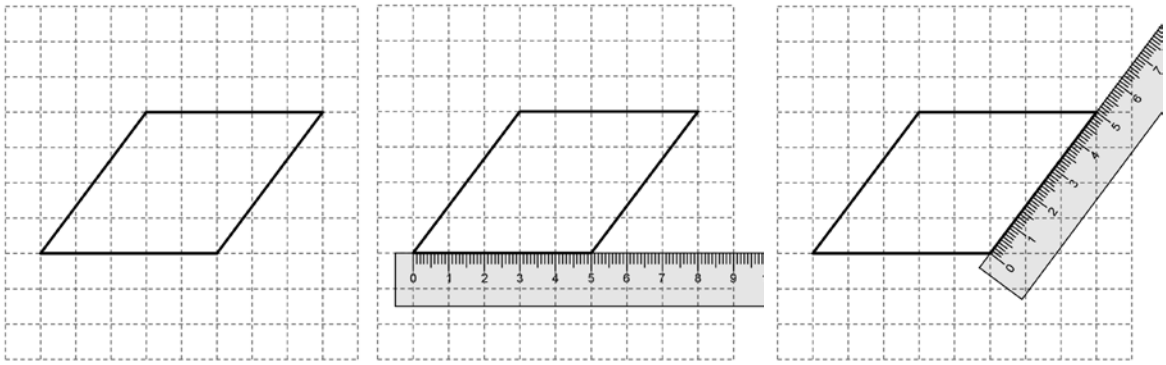


(a)　　(b)

圖　十　九

我們最初也擔心這個繪畫菱形的課業對學生來說太艱深。但在觀課中，我們發現大部份學生都能夠利用格線準確地繪畫菱形，並且十分投入學習活動，討論氣氛熱烈。在最困難的圖十八(b) 中，更有學生能夠清楚地向同學解釋繪畫第二個不同形狀菱形的方法。由此可見，我們不應該低估學生的學習能力。只要有適當的指導，他們絕對可以學得更多和更好。

當學生熟習在方格紙繪畫各種四邊形的方法後，我們可以讓他們在方格紙上隨意繪畫四邊形，並向同學指出所畫四邊形的名稱及其原因。留意這個活動可能會有出人意表的結果。例如我們觀課時見到一個學生畫出如圖二十(a) 的四邊形。大家最初都以爲這是一個平行四邊形，但當用直尺一量，發現原來四條邊的長度都是 5 cm（圖二十(b) (c)），所以它是一個菱形！雖然老師未能以勾股定理向學生解釋爲何如此「巧合」，但這也是一個很好的機會讓他們運用所學的判斷方法，並體會這方法的重要性。



(a) (b) (c)

圖 二 十

結語

四邊形這個橫跨兩個學習階段的課題涉及很多的學習重點，也帶來很多的學習機會。我們希望這套課業能夠協助老師把握這些學習機會，幫助學生釐清迷思，建立正確觀念，並且學習如何以正確的數學詞彙描述幾何特性及關係，爲他們將來學習幾何打好基礎。

鳴謝

謹此向歷年來協助我們在這個課題進行研究及試教的學校致謝：中華基督教會基全小學（2008－2009）、民生書院小學（2009－2010）、九龍灣聖若翰天主教小學（2010－2011）及基督教粉嶺神召會小學（2012－2013）。

參考文獻

De Villiers, M.D. (1994). The role and function of a hierarchical classification of quadrilaterals. *For the Learning of Mathematics*, 14(1), p.11-18.

Laborde, C. (2005). Robust and soft constructions: two sides of the use of dynamic geometry environments. In S.C. Chu, W.C. Yang and H.C. Lew (Eds.), *Proceedings of the Tenth Asian Technology Conference in Mathematics*. Cheong-Ju, South Korea: Asian Technology Conference in Mathematics.

Van Hiele, P.M. (1986). *Structure and insight: A theory of mathematics education*. Orlando, FL: Academic Press.

作者電郵：Cherie Ching　cherieching@edb.gov.hk
Anthony Or　anthonyor@edb.gov.hk

ISSN: 2305-7114