JOURNAL OF PROFESSIONAL LEARNING

YOUR CLASSROOM • YOUR STAFFROOM • YOUR FUTURE· YOUR RESEARCH

THE ASSESSMENT JOURNEY CONTINUES: TEACHER CENTRIC ASSESSMENT AND THE ROLE OF THE IMAGE

PROFESSOR JIM TOGNOLINI

Following on from the JPL article on assessment that he wrote in 2020, Professor Jim Tognolini gives a comprehensive insight into why teacher professional judgement is at the heart of assessment…

INTRODUCTION

Assessment is an integral part of the teaching and learning process. Tognolini and Stanley (2007) suggested that assessment involves professional judgment about student progress along a developmental continuum.

Central to such judgment are the *images* formed by the observed performance of students and knowledge of the standards that differentiate performance within the curriculum. Teachers are closest to their students and have many opportunities to observe and test their performance. They are also the primary agents in assessment and assess informally every day. However, the interest in assessment goes beyond the classroom. There are many other players who have a stake in assessment and in understanding what it means.

Many of the misapplications of assessment come from divorcing it from its natural role in the teaching and learning process and from misunderstandings about its nature and function in that process. This article shows how conceptualising learning as progress along a developmental continuum brings together curriculum, teaching and learning, and assessment as parts of one continuous process centralised on the teacher.

ASSESSMENT IS USED TO TRACK GROWTH

Assessment is about evidence of progress in the growth of knowledge, understanding and skills. This developmental emphasis shifts the focus of attention in assessment towards monitoring student progress in learning.

The key idea is that the students' progress or growth, in what is required to be learned, is monitored along a developmental continuum.

Development is a fundamental concept in education. Teachers' interactions with students facilitates their progressive development of knowledge, skills and understanding. Classroom activities are designed in a context of curriculum and syllabus specification about the content, level of knowledge and skills to be developed.

THE DEVELOPMENTAL CONTINUUM

The monitoring of student growth along a continuum requires the continuum to be defined and levels of performance to be articulated using pre-determined standards of performance. Effective curriculum frameworks and syllabus documents set out a developmental sequence, commonly in the form of statements of learning or outcomes. Outcomes describe what students are expected to know and be able to do at different stages along the continuum. They provide the basis for the development of the teaching and learning (including assessment) sequence and activity within a subject or course.

Curriculum requirements differ across systems in the degree of explicitness about content to be taught and mastered. This can be seen in a developmental sequence of outcomes from a primary syllabus which is shown in Figure 1. It shows a sequence of outcomes for understanding whole numbers.

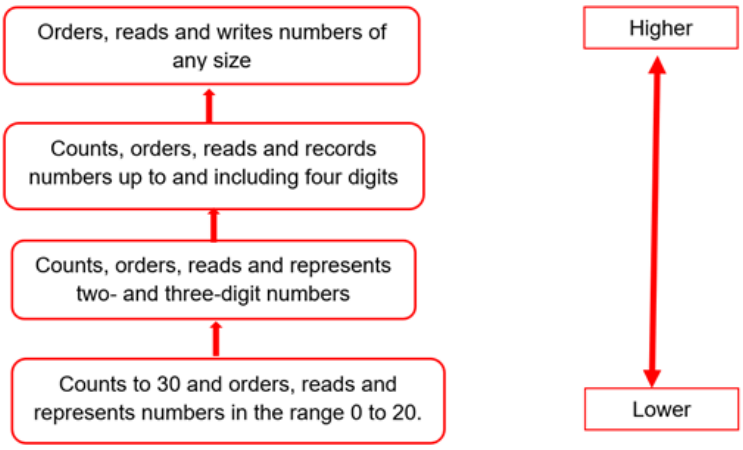

Figure 1: *Developmental sequence for understanding whole numbers*

Classroom activities are designed to enable students to progress through the developmental sequence and evidence of progress needs to be obtained for each student through appropriate assessment opportunities provided by the teacher. The developmental process is often represented in terms of a series of stages.

Progression from one stage to the next commonly involves a transition process. During transition performance may go backwards before it improves. This may be due to the next stage of learning requiring an ability to re-organise previous understanding into a new perspective. Consequently, there may be some uncertainty and inconsistency in performance until the new perspective is dominant.

Development implies improvement in performance. If there is no evidence of students improving, then there is no evidence of learning occurring. Whether formal or informal, assessment provides the evidence as to where a student is located on the developmental continuum which underpins the curriculum. To this extent curriculum, assessment, teaching and learning need to be closely integrated. A good question to ask when preparing lessons and associated assessments is: How are we helping our students move along each developmental sequence of knowledge?

TEACHING AND ASSESSING FOR DEVELOPMENTAL PROGRESSION

When designing an assessment program, the purpose is to provide information, which helps teachers understand student progress along the developmental continuum, which underpins the curriculum. To progress along the continuum, students must become more proficient in the subject. Outcomes that are further along the continuum are intended to be more cognitively demanding for the students. They require more of the 'attribute', 'trait' or 'construct' that enables the students to demonstrate proficiency. Progressing along the continuum means that students are becoming more proficient in the subject.

Progress is generally represented by a series of stages that are cumulative in nature. Skills, understanding and knowledge that students demonstrate, at different stages along the developmental continuum for a subject or learning, are typically captured by generic descriptors with broad descriptions of standards. Teachers in schools can locate students along these developmental continua by comparing their 'images' of students, informed by the assessments, to these broad standards and using their professional judgement say, on balance, that the student is located at a 'stage 3' (or 'level 3' or 'band 3') at this point in their learning because the description of the standard best aligns with the image of the student.

THE IMAGE

The concept of *image* is a central aspect of teachers' professional judgement about student growth.

Teachers build images of what students know and can do based upon all the information that is collected from various assessment techniques, not just formal or standardised assessments. The image of a given student

is built up from such information and, if new evidence comes in, then the image of that student will need to take account of the latest evidence. The image is critical to the teaching and learning process. It is not based on subjective opinion because it needs to be consistent with evidence.

Generally, the information that emerges from students completing classroom tasks, answering questions, from students talking to each other, and taking classroom tests, standardised tests or examinations is expected to be consistent with the images.

Sometimes it is not, and the teacher then asks the question, "Why not?" There are many students who perform well in classroom activities and yet perform poorly on a standardised test or examinations; this atypical performance is of interest to teachers. It could be that a student has some difficulties which have been identified by the performance on the test and there is a need to collect further evidence to see if there is a need to adjust the image of that student. Alternatively, it could be that the result may have been caused through other reasons that would not warrant a substantive change in the image e.g., the student was sick on the day of the assessment or did not try to do well on the assessment.

In summary, therefore, teachers use assessments to form an image of what students know and can do. As more information becomes available from a variety of assessment sources, it is incorporated into the image. The various sources of assessment are targeting the same material from different, but interrelated perspectives. Consequently the "fairest image" emerges when teachers use a range of assessment techniques and assimilate the information from the multiple sources using their professional judgment. Teachers are constantly assessing students or their actions, taking the latest information back to the image and making informed decisions as to what to do next. In this way assessment is fully integrated into teaching.

Figure 2 shows the usefulness to teachers of various assessment **activities** ordered from more-useful to less-useful in producing the image.

The key point is that assessment is teacher centric. All data, whether it has been collected from classroom interactions or formal tests should be interpreted by the teacher using professional judgement. One of the questions that is often asked by teachers around the world is "How do we bring together formative and summative assessments?". The response is that it is done through the process of professional judgement described above. Summative tests and formal assessments provide just one more piece of evidence that is used to inform the image which is used to monitor student growth.

***Figure 2**: Usefulness to teachers of various assessment methods in developing the image*

	Example of student assessment *methods*	Examples of other *methods*
More useful ↑	Classroom observations, questioning, etc.	Ad hoc observations with minimal data collected for post hoc analysis
	Classroom assessments administered and marked by an individual teacher	Questionnaires with minimal effort to validate them before administration and little attempt to carry out post hoc validation
	Classroom assessments (could include tests, checklists, project work, etc.) where there has been some degree of collaboration with other participating teachers in constructing the assessment and achieving comparability across different markers across the school	Surveys with self-report data where there has been some effort to validate the questionnaire, interview or observation schedule before administration
	Published achievement tests administered in the classroom (NAP self-administered tests on civics and citizenship, science literacy and ICT)	Well-constructed self-report and validated surveys and teacher quality data where the judgement of the evidence is made independently against standards and the evidence is available for further judgements if required
↓ Less useful	Standardised achievement tests (e.g., NAPLAN and HSC)	Controlled census type data collections (e.g. attendance data, student welfare data)

USING THE IMAGE TO MONITOR STUDENT PERFORMANCE AGAINST STANDARDS USING TEACHER PROFESSIONAL JUDGEMENT

There are numerous advantages for students and teachers in using a system whereby student images are referenced, using teacher professional judgement, to pre-specified standards of performance. One advantage is that reporting of student performance is focused on individual progress on the developmental continuum rather than on performance relative to other students or on so called "mastery" of content. That is, there is a desire to see growth in the individual student and that is outcome is provided by the developmental continuum. A second is that continua, with descriptions of performance, provide a picture of what it means to improve in learning in different areas. A third is that teachers can help students (and others) know what is required and what it is that they must do to progress along the developmental continuum.

For students to demonstrate where they are along a developmental continuum, they must be given the opportunity to show what they know and can do in relation to the outcomes of the subject. Tasks, activities and test items provide them with this opportunity. This is important in differentiating learning. If very able students are not given the opportunity to show that they have developed in their learning, by giving them opportunities to demonstrate greater levels of cognitive depth, then it is not possible to locate them on the developmental continuum with a degree of consistency or accuracy. This is not fair for the students.

Vygotsky (1978) used the concept of *the zone of proximal development* as the region on the developmental continuum to describe where students can learn best. Located between that which is too easy and that which is too hard, it is where the guidance of a person more competent in a task (generally the teacher, but could be a student's peers, parents, etc.) can help a student to reach his or her potential. The most effective way in meeting the learning needs of individual students is to locate the student on the developmental continuum and then work within the region where they are located.

Differences in the pace of student learning can be due to some having a slower path of development, reaching a plateau at a lower level of performance to others or needing to develop other capacities first. While such differences are quite common, especially in non-streamed school classes, many believe that growth paths should 'close the gap' between the lowest and the highest performers. However, in practice this may lead to holding back students who reach the need for the next step earlier. The important task is to help all students to progress along the developmental continuum as quickly as they can. As Masters (2013) has argued current school organisation and grading practices do not deal adequately with individual differences in growth.

Teacher judgement of student progress affects how they structure teaching and learning activities to enable students to progress through the developmental sequence in achievable steps. Evidence of student growth can take many forms but should be considered in terms of how well it satisfies the needs for practicality, fairness, validity and whether it provides feedback to assist the next step in the developmental pathway for an individual. Timely feedback is essential to assist learning (Hattie & Timperley, 2007).

Different sources of evidence about student growth should converge. For example, if in a particular case there are different signals coming from external tests than classroom observation, rather than discarding one source there will be value in adopting a forensic approach to understanding why such discrepancy has occurred. The product of such analysis should lead to a more effective understanding of and eventually improvement in student learning.

CHALLENGES IN IMPLEMENTING AND MONITORING A CONSISTENT APPROACH TO GROWTH

The consequences of such a model of assessment requires re-negotiating the processes of curriculum, teaching and assessment towards a holistic emphasis on how growth occurs and on what evidence should be gathered to show that it is occurring. If curriculum requirements are not organised with respect to developmental outcomes, which clarify expected learning pathways and progress maps, then teaching programs are unlikely to yield evidence of depth of learning.

It takes time and resources to develop research-based learning developmental continua and, so far, most attention to such development has been in areas such as literacy, numeracy and science (e.g. see Black et al.,

2011; McNamara & Hill, 2011). These areas have been given special attention because of their core nature and apparent tractability to a developmental pathway.

Digital technologies have much potential to assist in the process of learning. They can present varied assessment tasks with useful feedback customised to individual developmental levels. One senses many opportunities for improved assessment from educational use of such technologies.

THE IMAGE OF A STUDENT FORMED BY PROFESSIONAL JUDGEMENT

From the discussion to this point, it should be clear that the image of the student formed by professional judgment is central to modern assessment in education.

The image of a student is defined in terms of the observation and experienced-based impression of their current level of performance. When this point of view is expressed to teachers and students, one of the first responses is that the image appears to be a very subjective concept.

This leads to some potentially awkward questions: Is not good assessment supposed to be objective and unbiased? Why is such a subjective term like *image* considered central?

Clearly, there is a need for assessment to be fair and unbiased, and it is important to examine how this can be achieved in practice. Recognising the centrality of professional judgment in assessment does not mean that assessment is primarily a subjective activity, where 'subjective' implies arbitrariness or inconsistency.

Observation in science involves professional judgment using agreed protocols for collecting evidence. This evidence is then tested against other evidence. The outcome of such observation is accepted as part of the scientific endeavour and is not considered subjective. Similarly in assessment it is possible to have confidence in the outcomes, provided careful attention is paid to the processes of observation and how the conclusions about student performance are determined. It does require a level of assessment literacy of teachers that may or may not be evident at this point. However, building the capacity of teachers in assessing and making consistent judgments of student performance against standards would seem to be a worthy goal given the importance of assessment and data literacy to teaching and student learning.

There is a need to consider how acceptable information can be generated to test and refine the image developed of a student. It is important to look carefully at the different sources of information and their respective contribution to the overall image.

All evidence collected needs to be considered carefully. This includes so-called 'objective' test data. Just because a multiple-choice test can be marked objectively does not mean that it is free from professional judgment in its construction, or that it always gives more valid information. The person writing objective test items has an image in mind of what knowledge and skills can be demonstrated by students responding to the test. This image is used to make decisions regarding choice of test format and item content.

For some purposes, a multiple-choice item may be the most efficient way of testing particular knowledge. In other cases, by providing a frame for student responses, the construction of a multiple-choice item may be seen to limit the opportunities for students to show creative use of the knowledge and skill they possess.

Depending on the purpose of the assessment, a better solution may be achieved by substituting an open ended short-answer question for the multiple-choice item. Every time a formal test is devised there is a series of judgments that need to be made to ensure that the information gained helps our understanding of student achievement.

The key to good assessment is to understand both the centrality of professional judgment in the collection of information that leads to the formation of the image of the student being assessed and ways of ensuring that the professional judgement is well grounded in evidence.

The initial image may well be formed by partial information and hearsay. It is important to move beyond this to classroom observation, more formal and informal data informing the image that is used to drive teaching and learning.

Why is this so important? The literature on teacher expectation suggests that untested impressions are likely to be unfair and lead to unsound and unproductive further teacher-student interactions.

Most teachers have heard of the Pygmalion Effect studied by Rosenthal & Jacobson (1968) in which it was claimed that impressions of students' ability formed by teachers influenced their actual student achievement. Ever since then concerns about teacher expectation effects and self-fulfilling prophecies have led to worry about judgments by teachers leading to unfair and biased outcomes for students.

Being worried about it is a good sign. Knowing the potential effects of unfounded and untested assumptions about the students is essential if teachers are to avoid making mistakes about them.

Subsequent debate about the relationship between expectancy and performance suggests that it could be just as easily claimed that teacher expectation effects were due to student effects on teachers rather than the other way around (Brophy & Good, 1970). Like most controversy, there is some evidence in favour of both directions of expectancy effects. Interactions with students provide a strong basis for our understanding of what they can do.

While there may be contexts in which the expectations of performance are not well formed by evidence, this is not ground for asserting that all images of the performance capability of students are necessarily subjective and untrustworthy.

The image a teacher may have of a student is initially formed by expectation and professional judgement and needs to be continually challenged and revised by evidence collected during everyday classroom experience as well as test data. As mentioned previously, assimilating information about performance of students from several sources and over several occasions leads to more reliable and valid images.

Teachers must always believe in the possibility that their students will continue to develop. The image that each student presents in terms of performance and achievements should help guide the teacher in the next step to develop the student. However, for the next step to be achievable, there is a need to have a well-grounded view of the student's current level of knowledge and skill. To achieve this does not mean that that there is a need to collect a large amount of evidence. Sometimes uneasiness about how much evidence is needed to have an appropriate image of their students leads teachers to become overzealous in collecting a large portfolio of student work.

To have a well-grounded basis for the image of students, teachers must have confidence in the observations they make about student performance. The quality of the evidence is more important than the amount of evidence. Classroom engagement with students through discussion and observation adds to any assignment or test data in forming the image.

RELIABILITY OF CLASSROOM ASSESSMENT

Observing and making professional judgements about students every day, as they engage in classroom activities and conversations, is an integrated part of the work of teachers and of good teaching. As the interactions are many, and occur over several occasions, assessment based on these interactions is more reliable than assessments made based on a one-off test. In principle the reliability of assessment increases with the number of observations made.

Nevertheless, there are concerns about how to ensure the reliability and validity of teacher assessment, especially where there is performance management based on student outcomes. External standardised tests often claim to be more reliable and independent even if they can be perceived to be limiting the scope of what is taken as evidence of student achievement.

Much of the educational research literature on the reliability or validity of teacher assessment is embedded in contexts, that may not fit well into modern system-wide reporting and accountability frameworks.

In considering classroom assessment practice it is essential to distinguish between judgments based on formal written work, such as essays and assignments of varying structure and content, and those based on dynamic interactions in the classroom.

Different classroom teaching and learning situations vary in opportunities to observe and record information to inform judgments about student achievement. Teacher assessment practices differ in the extent of data collection and recording (ranging from detailed protocols to 'on-balance' judgments of achievement of assessment criteria). As with external tests and examinations, it would be expected that different requirements would show different degrees of reliability.

Reliability of a measure may be improved in two ways - by making the assessment(s) underpinning the measure longer and by improving the properties of the assessment tasks.

Tasks may be critiqued to remove ambiguities, or the difficulty of the tasks may be adjusted to make them more consistent with the average ability of the student group being tested. Some parts of the task may be substituted with items that are inherently more reliable (e.g. short answer or multiple-choice) or the marking scale may be refined to obtain greater clarity of the relationship between the quality of an answer and the marks/grade awarded.

Importantly the key point in this article is that as teachers base their images on data collected every day and in multiple ways throughout the school year, the image is based on many, many more observations than a standardised assessment and, because of this, outcomes from the assessments are likely to be more reliable.

WORKLOAD ISSUES

In classroom assessment there are inevitable tensions, that arise from the interaction of the following aspects:

- The range and quantity of work on which teachers' judgments are made
- The manageability of making such judgments during teaching
- The recording and storage of evidence

To ensure the validity and authenticity of assessment, it is desirable that teachers' judgments are based on observations of a student performance on a wide range of activities. This is to ensure that a student is given every opportunity to show their level of functioning in relation to the curriculum standards. However, tension arises as to the manageability of recording such observations for all students in the context of a busy classroom.

Concerns about the reliability and validity of school and classroom-based assessment sometimes creates a tension between quality of measurement and good teaching practices. The former places an emphasis on standardisation so that students are being compared fairly on the same or similar tasks. On the other hand, the latter often requires differentiation, where teachers may give more structure and more help to lower ability students and greater autonomy to high ability students.

Some classroom assessment systems like the English Assessing Pupil's Progress (APP) suggested that the teacher take notes on every observation that might contribute to an assessment. While this has the virtue of giving a complete picture of the student over the full range of educational activities, teachers tend to become overwhelmed by the sheer amount of data collected. Moreover, students may feel that they are always under observation. Such effects may interfere with the normal teaching process (Stanley et al., 2009).

Another approach lets the cumulative effect of informal observations create a judgment of what a student knows and can do. This more informal approach is not dissimilar to how teachers usually form an image of their students' capabilities. Memorable observations that indicate atypical performance are recorded to check the confidence of the teacher that the student has reached the presumed level of performance.

SUMMARY

The central concept in the teaching and learning process is the idea of developmental continua underlying the domains of knowledge and skills being taught. Assessment enables the progress of students to be monitored along these continua and provides essential feedback to assist in designing the next step in student learning. There needs to be close alignment between the curriculum, teaching and learning and assessment.

The central concept of this article is that teachers are assessing their students continuously and building an image of what they know, can do and understand. Using this image, and a defined underlying developmental continuum based on agreed upon standards, student progress can be defined, observed and communicated in tangible ways and the teaching and learning process can be modified to take individual student needs into account without overwhelming the teacher with formal assessment processes or data. Furthermore, it is likely the corpus of information, collected in such a manner, will be as, if not more valid and reliable than one-off assessments conducted at a single point in time, typically encountered in standardised test. While such assessments provide good quality data, they are just one more piece of evidence the teacher should use to adjust their

image of their students relative to the developmental continuum.

The view of assessment advanced in this article puts the teacher at the centre of assessment relative to the teaching and learning process. Finally, this process will only work if there is close alignment between the curriculum, what is taught and what is measured by the assessments.

REFERENCES

Black, P., Wilson, M. & Yao, S. (2011). Road maps for learning: A guide to the navigation of learning progressions. Measurement: Interdisciplinary Research and Perspectives, 9,2-3, 71-123.

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

Krajcik, J. (2011), Learning progressions provide road maps for the development of assessments and curriculum materials. Measurement: Interdisciplinary Research and Perspectives, 9, 2-3, 155-158.

Masters, G.N. (2013). Reforming educational assessment: Imperatives, principles and challenges. Australian Education Review Number 57.

McNamara, T. & Hill, K. (2012). A response from languages. Measurement: Interdisciplinary Research and Perspectives, 10(3), 176-183.

Rosenthal, R. & Jacobson, L. (1968). Pygmalion in the classroom: Teacher expectation and pupil intellectual development. Holt, Rinehart & Winston.

Stanley, G., MacCann, R., Gardner, J., Reynolds, L. & Wild, I. (2009). Review of teacher assessment: Evidence of what works best and issues for development. Report on QCA Contract 2686. http://www.qcda.gov.uk/27194.aspx.

Tognolini, J. & Stanley, G. (2007). Standards-based assessment: A tool and means to the development of human capital and capacity building in education. *Australian Journal of Education*, 51(2), 129-145.

Vygotsky, L.S. (1978). Mind in society: Development of higher psychological processes. Harvard University Press.

ABOUT THE AUTHOR

Professor Jim Tognolini is Director of The Centre for Educational Measurement and Assessment (CEMA) which is situated within the University of Sydney School of Education and Social Work. The work of the Centre is focused on the broad areas of teaching, research, consulting and professional learning for teachers.

The Centre is currently providing consultancy support to a number of schools. These projects include developing a methodology for measuring creativity; measuring 21st Century Skills; developing school-wide practice in formative assessment. We have a number of experts in the field: most notably, Professor Jim Tognolini, who in addition to conducting research offers practical and school-focused support.