

Teacher ‘rater’ training in Hong Kong and Australia: Different contexts, same problems?

Chris Davison, University of New South Wales,
Australia

Overview

1. Background: The rise of the teacher 'rater'
2. A sociocultural approach and its application in an Australian educational system
3. The transferability of models of teacher assessment decision-making: The case of Victoria, Australia vs. Hong Kong
4. Conclusions

1. Background: The rise of the teacher ‘rater’

- The international movement towards the more effective collection and use of assessment information to improve learning in schools (Darling-Hammond & McCloskey, 2008) has led to a number of recent developments in teacher-based English language assessment, including a focus on school-based assessment and the adoption of assessment *for* learning (AfL) to help drive improvements in student learning (Davison & Leung, 2009), especially in areas of English language development that have traditionally been difficult to assess with large-scale assessments, such as speaking skills.
- However, at the same time there has been a push to increase teachers' assessment literacy as well as that of other stakeholders (Taylor, 2009), in part due to concerns that teachers may lack sufficient training in what educational assessment entails (Malone, 2011; Stiggins, 1991, 2008, 2009; Taylor, 2009) and/or “may lack the confidence or skills required to perform their assessment duties in a competent manner” (Stiggins, 2008, p. 8).

- Particular concerns have been raised about the ‘ordinary’ teachers’ ability to be able to assess their own students consistently, accurately and fairly, especially for high-stakes purposes (Lynch & Shaw, 2005). This has foregrounded the issue of teacher ‘rater training’, the topic of this presentation.
- In traditional rater training for large-scale external assessments, the focus is usually on ensuring inter and intra-rater reliability (which for students is often identified as a ‘fairness’ issue), that is, is there consistency across contexts/assessors? (Lynch & Shaw, 2005).
- Variability in assessment is traditionally viewed as assessment bias, that is:
 - Unfair: Disadvantageous to one subgroup but beneficial to another, and/or
 - Offensive: Driven by “negative stereotypes of certain subgroups” (Popham, 2014).

Many research studies have found variability to result from different factors in the rating process, including:

Experience (Leckie & Baird, 2011; Barkaoui, 2011);
Weigle, 1998)

Teacher
(rater)-
related
factors

Gender (Eckes, 2005; Lumley & O'Sullivan, 2005;
O'Loughlin, 2002)

Language backgrounds (Brown, 1995; Caban, 2003;
Johnson & Lim, 2009; Lee, 2009; Shi, 2001; Xi &
Mollaun, 2009; Yan, 2014)

Education, qualification and training (Brown, 1995;
Fayer and Krasinski, 1987)

| | |
|---|--|
| Student- (ratee) related factors | First Language (Ehrenberg, Goldhaber & Brewer, 1995; Dee, 2005; Ouazad, 2008) |
| | Gender (Eckes, 2005; Lumley & O'Sullivan, 2005; O'Loughlin, 2002) |
| Task- related factors | General task effects (Fayer & Krasinski, 1987; Galloway, 1980; Hadden, 1991; Kim; 2009) |

However, the rise of school-based assessment and the increasing use of classroom teachers as 'raters', even for high-stakes assessment, has resulted in a paradigm shift

The assessment paradigm shift

Psychometric vs 'classometric' (Brookhart, 2003)

| Focus | Large scale assessment concepts | Classroom assessment concepts |
|--------------------|--|---|
| Validity | Measure is external to inferences made and actions taken | Inferences made and actions taken are internal to the measurement process |
| Validity | Measurement context is construct-irrelevant | The measurement context is construct-relevant |
| Reliability | Reliability is consistency over irrelevant factors | Reliability is sufficiency of information |

Classometric theory of assessment (Brookhart, 2003)

1. Classroom assessment is an internal process

- Traditionally, a test is a ‘dipstick’ into an oil tank – *external*, but how instruction is conducted and learning occurs is not considered.
- Classroom assessment is *internal*. Inferences and actions are internal to the assessment process.
- Classroom assessment is part of the learning process.
- Teachers are internal to the assessment process, because their assessment reflection and feedback-feedforward are part of the assessment information (Black & Wiliam, 1998)

2. The assessment context is construct-relevant

- Classroom assessment is the integration of instruction and assessment.
- The same task is perceived differently in different contexts.
- How instruction is conducted must be taken into consideration.
- Primary purpose of all assessment must be for student improvement, so classroom assessment functions as both a tool for evaluating student progress/achievement, and a critical mechanism for improving learning and enhancing instruction, i.e. both formative and summative assessment.
- Students are no longer objects of assessment, but proactively involved (Anderson, 1998) as primary users of assessment results.

So what about reliability?

The concept of trustworthiness

- In classometric assessment, reliability is determined not so much by the consistency but by the **quality** and **trustworthiness** of teacher judgements
- With human assessors, there is always a problem of interpretation, hence professional dialogue and interaction need to be central to decision-making process, foregrounding the importance of the establishment of ‘interpretative assessment communities’
- In order to establish and maintain trust such “communities will have to show that their procedures for making judgements are fair, appropriate and defensible (i.e. that they are valid), even if they cannot be made totally transparent” (Wiliam, 2001, p.173-4).
- We need to improve the ability of all key stakeholders to draw inferences or derive judgments from the data, especially those who are meant to be doing the improving (i.e., students, parents, teachers and policy-makers), as in naturalistic inquiry ...

The concept of trustworthiness

Guba, 1981, *Criteria for assessing the trustworthiness of naturalistic inquiries*

| Problem | Rationalistic treatment | Naturalistic treatment |
|---|--|---|
| Masking or competing factors (confounding) | Control, randomise, <i>Internal validity</i> | engagement, triangulation, member checks <i>Credibility</i> |
| Situational variations (atypicality) | Probability sampling, <i>External validity</i> | thick description, purposive sampling <i>Transferability</i> |
| Instrumental drift or decay (instability) | Replicate, <i>Reliability</i> | stepwise replication, audit trail, <i>Dependability</i> |
| Investigator predilections (bias) | Insulate the investigator, <i>Objectivity</i> | triangulation, practice reflexivity <i>Confirmability</i> |

Improving teacher assessment decision-making

- Teacher-based assessment derives a major part of its validity from its location in the actual classroom where assessment activities are embedded in the regular curriculum, and assessed by a teacher familiar with the student's work and with a stake in their improvement.
- However, if assessment is to be authentic (and not distort classroom activities and student and teacher behaviour) schools and teachers need to be granted a high degree of trust and autonomy in the design, implementation and timing of assessment tasks.
- Students also need to be given sufficient time and support to demonstrate their best - to show what they **can do** - and for the assessor to be able to confidently assess their output, but even more importantly, 'test' their informal judgments of students' language levels and achievements.

Improving teacher assessment decision-making

- More appropriate tools, processes and forums are also needed which encourage the teacher to stand back and reflect on their implicit or explicit assumptions about individual students' capacities, compare those assumptions with careful analysis of examples of their own other students' actual performance, and then subject their judgments to explicit scrutiny and challenge, or confirmation by others, hence the critical importance of establishing forums and processes for disagreement (a pre-condition for establishing sufficiency of information)
- Research shows that the teacher assessment is never 'objective'; the teacher always has preconceived ideas or assumptions about a student's level, but they need to 'test' their own informal judgments through participation in common assessment tasks, benchmarking work samples, and social moderation in which their judgments are made explicit and open to discussion with fellow teachers.

Improving teacher assessment decision-making

- To do this, it is necessary to establish common understandings of key constructs and tasks, publicly agreed and explicit common assessment tasks, criteria and levels of performance, and strong moderation among teacher-assessors and their communities, including students.
- However, teacher-based assessment systems are still too often evaluated against traditional psychometric criteria developed for formal testing programs
- Hence, as well as developing appropriate tools and resources to ensure quality and trustworthiness in teacher assessment decision-making, there is also a need for better theorization of teacher-based assessment practice.

2. A sociocultural approach and its application in an Australian educational system



Department of
Education & Training



UNSW
AUSTRALIA



catholic education commission of victoria ltd
ACN 119 459 853



Australian Government



Tools for Enhancing Assessment Literacy for Teachers of English as an Additional Language

[Home](#) [Professional Learning](#) [Assessment Tools](#) [Annotated Units of Work](#) [Discussion Forum](#) [Contact Us](#)

Welcome to Tools to Enhance Assessment Literacy



The TEAL web project, launched in June, is an online resource for teachers of primary and secondary level children who are learning English as a second (ESL) or additional (EAL) language in Australia. It brings together a range of tools and advice for the assessment and reporting of the English language proficiency and progress of students.

For more information on the TEAL web project, its contents and its intended use, please follow the links below to read more. If you have any further questions or comments, don't hesitate to visit our

Sociocultural (Vygotskian) theory

..is a theory of mediated mental development.
(Lantolf, 2006, p.4)

..is heavily focused on the impact of culturally organised and socially enacted meanings on the formation and functioning of mental activity. (Lantolf, 2006, p.2)

Mediation is the process through which humans deploy culturally constructed artifacts, concepts, and activities to regulate (i.e. gain voluntary control over and transform) the material world or their own and each other's social and mental activity.(Lantolf, 2006, p79)

Teacher assessment practice as semiotic mediation and co-regulation

Co-regulation is the combined semiotic mediation of:

Object-regulation - material and ideational 'things'
signified objects

Other-regulation - physically or ideationally present people
significant others

Self-regulation - mediating object and others as
psychological tools

(Frawley, 1987)

Trustworthiness as co-regulated assessment activity

Other regulation

community of assessment practice
expertise in use
modelled, guided practice

*Trustworthy
assessment*

Self-regulation

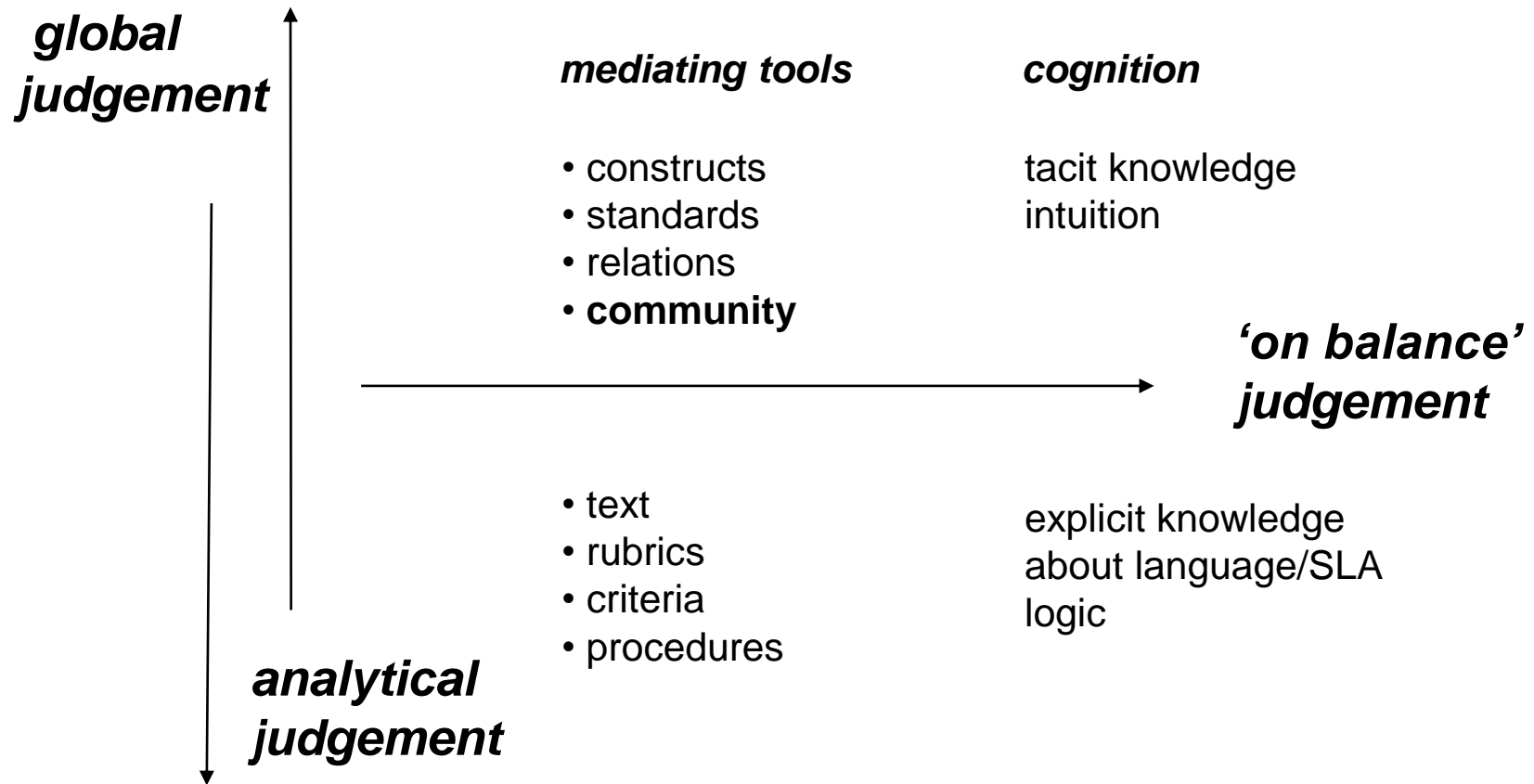
other/object mediation,
appropriation, internalization,
reflexive assessment
judgement, meta-cognition

Object regulation

assessment tools, instruments
frameworks, rubrics, criteria, methods

A dialectic model of teacher assessment judgement

(Michell & Davison, in preparation)



“To function dialectically means to be able to hold in one cognitive space notions that on the surface appear to be contrary... and to come to understand how these seeming contraries fit together as necessary components of the object of study”.

(Exploring the Dialectic: An Interview with James Lantolf, 2007)

TEAL: Tools to Enhance Assessment Literacy for Teachers of EAL

- Drawing on Assessment for Learning (AfL) principles and Vygotskian theory, TEAL is an online assessment advice and 'toolkit', <http://teal.global2.vic.edu.au/>, for use by all Victorian school teachers to help assess the stage of development for EAL students in speaking and listening, reading and writing, and to improve learning and teaching
- All tools are aligned against the *Victorian EAL Standards (VELS)* and the *EAL Developmental Continuum*, now being revised to align against the new EAL curriculum, with potential for alignment to other standards by other jurisdictions.

TEAL as an assessment activity system

Means/tools/ instruments

Knowledge tools

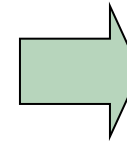
- EAL developmental continuum
- text framework
- standard assessment task
- scoring procedures
- assessment criteria
- benchmarked student work samples
- performance levels, rating scales
- reporting framework

Tacit knowledge

- pedagogical language assessment content Knowledge
- language constructs
- learner needs

ESL teacher
teacher as assessor

Object
trustworthy
assessment
judgements



Outcome

- fair, valid, reliable classroom EAL assessment
- **trustworthy teacher-based EAL assessment**

Rules

- professional norms, accountability
- program policy and procedures,
- curriculum standards
- fairness, validity, reliability
- assessment criteria
- **trustworthiness**

Community

- community of EAL teachers
- assessment stakeholders/consumers
- assessment reform stakeholders
- education authority
- assessment culture

Division of labour

- Teacher /student assessors
- assessment developers
- experts, novices,
- project participants

School of Education



UNSW
THE UNIVERSITY OF NEW SOUTH WALES

An online teacher-mediated EAL assessment resource centre continually being expanded, comprising:

Teacher professional learning resources

- Self-directed video and text-based resources and tasks on assessment for learning principles and processes to develop assessment literacy.

An assessment tools bank

- A range of assessment tools and tasks, including computer-adaptive tests, organized around macro-skills cross-referenced by assessment type, EAL stages and year levels, with extensive annotated exemplars and advice for teachers and students.

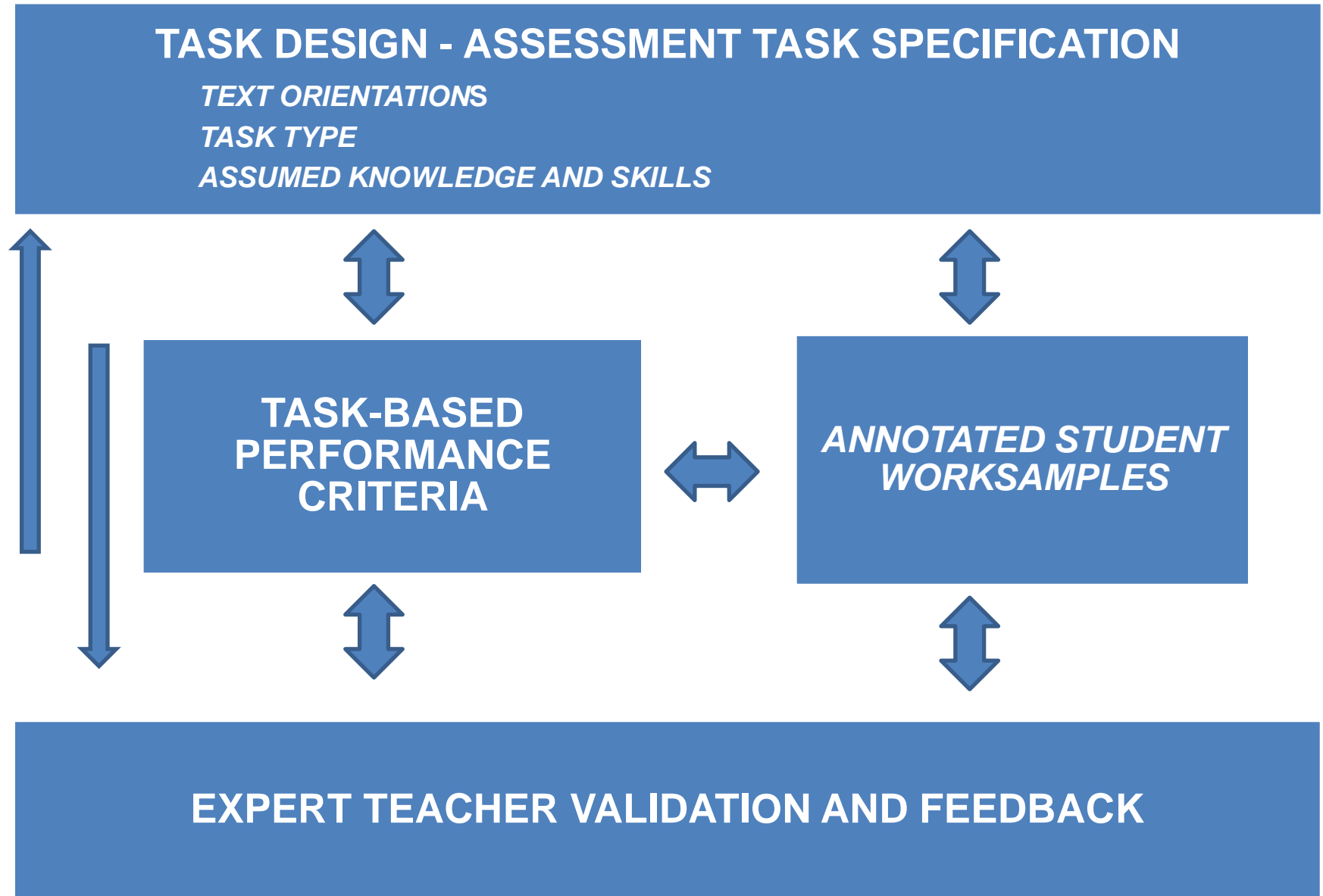
Assessment for teaching and learning exemplars

- A selection of annotated units of work across a range of subject areas and year levels, including in the mainstream, showing assessment tasks with self-peer assessment and formative feedback embedded within a teaching/learning cycle.

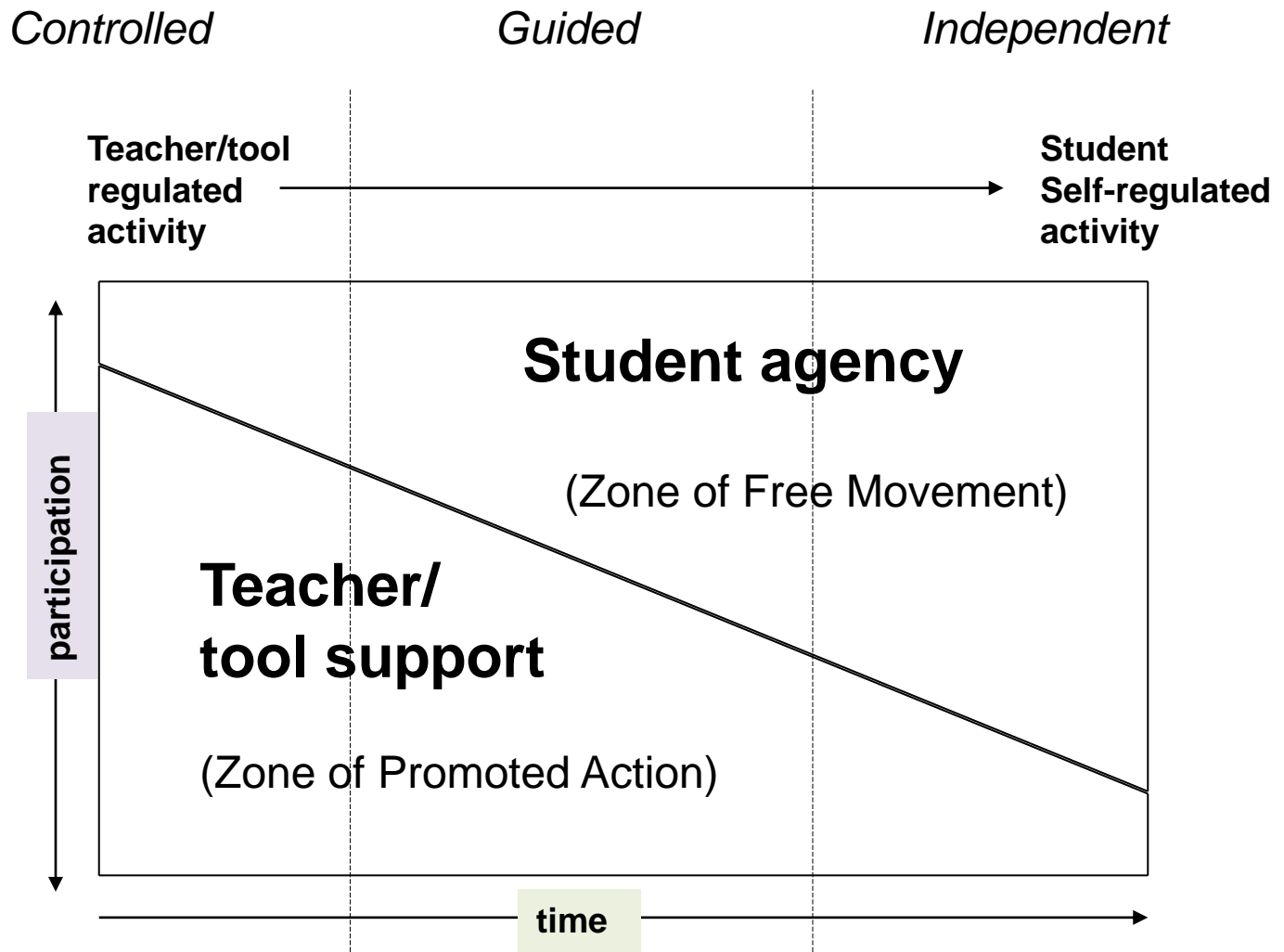
An online teacher discussion and benchmarking forum

- A password-protected area for teachers to share problems, strategies and work samples and engage in moderation/benchmarking.

Assessment task design as co-regulation



Scaffolded assessment tasks as co-regulation



3. The transferability of models of teacher assessment decision-making

- To what extent can such approaches be applied across different assessment contexts? eg. Australia vs. Hong Kong
- To some extent the wrong question as theoretical underpinning of sociocultural approaches is mediation... so no two assessment systems can be the same, but some commonalities and differences.
- Our experience shows that whilst many of the same problems arose in the development of teacher training programs for school-based assessment, there were some significant differences due to important contextual differences.

- In terms of its assessment demands and practices, Australia is very different from other English-speaking countries and from Asia, e.g., http://www.cese.nsw.gov.au/images/stories/PDF/Re-assessing_Assessment_v6.pdf

“Australia tends to sit in the middle when comparisons are made between it and other countries, both in terms of the performance on international assessments; and in terms of an assessment system which is balanced between classroom level assessment and standardised assessment, with a reasonably strong focus on classroom assessment ... Teachers are encouraged to use a variety of assessment techniques that are valid, reliable and appropriate to the age and stage of learning. The development of these skills has been a major focus of teacher professional learning opportunities in NSW schools over the past decade. Teacher survey data from OECD’s Teaching and Learning International Survey (TALIS) indicates that 80 per cent of NSW teachers who participated in professional development relating to student evaluation and assessment practices reported that it had a moderate to large impact on their teaching (CESE 2015)”

- However, there is evidence that Australia has some way to go to ensure that teachers understand how to interpret and understand assessment data and effectively embed assessment within a framework of teaching and learning. For example, an OECD review of Australian assessment practices undertaken in 2011 found that when teachers graded against national A-E standards, the consistency of their judgements within a school was weak (Goss et al 2015).
- In EAL in Australia, as we co-developed TEAL, we were conscious of the need to ensure teachers had a deep understanding of the construct being assessed (e.g., speaking), access to well-designed common assessment tasks, and the ability to make clear and consistent assessment decisions through understanding key terminology and assessment criteria, benchmarking multiple work examples and participating in extensive opportunities for social moderation to test their judgements against their own expectations and those of their peers (and their students)

Log In

[Home](#) > [Assessment Tools](#) > Common Oral Assessment Tool

Common Oral Assessment Tool

Guiding Principles and Advice

1. [Preparing to Assess Students' Learning](#)
2. [Getting Students Ready for an Assessment](#)
3. [Choosing an Oral Assessment Task](#)
4. [Using the Assessment Criteria](#)
5. [Making Trustworthy Assessment Decisions](#)
6. [Aligning your Assessment with the EAL Continuum](#)
7. [Using the Assessment Data to Improve Learning](#)
8. [Glossary](#)

Assessment Tasks and Criteria

[Oral Assessment Tasks and Criteria](#)



However, we also identified many areas needing more research, e.g.,

Phung (2018) examined to what extent the speaking assessments of EAL teachers new to TEAL were consistent, the factors influencing their assessments, and the common characteristics of teacher decision-making. Employing a mixed-method research approach, this study involved twelve experienced EAL teachers who participated in a survey, an assessment activity and a think-aloud protocol followed by individual interviews. The findings revealed that teachers were different from each other in their decision-making processes and in their perception of student performances and that the differences were influenced by factors related to teachers' and students' backgrounds and the characteristics of assessment tasks, with three different pathways of teacher decision-making identified, namely: (1) integrated assessment, (2) conflicted assessment, and (3) automated assessment. These decision-making pathways provide a new lens for explaining variability in teachers' judgement of student outputs.

| | Communication | Cultural conventions of language use | Linguistic structures and features | | | | Strategies |
|---|---|---|--|--|--|---|--|
| | | | Text structure | Grammatical features | Vocabulary | Phonology | |
| 4 | <ul style="list-style-type: none"> Describes the plot in detail, and relates to themes and issues Describes characters and how they illustrate or relate to themes or issues Describes key events and how they relate to themes or issues Relates personal evaluation of the work to elements of the work Relates work to self Comments on elements related to filmography, literary techniques or devices and their impact | <ul style="list-style-type: none"> Very fluent interaction, responding to questions | <ul style="list-style-type: none"> Long turns Extended statements Details of text related to themes and issues, and deeper personal responses | <ul style="list-style-type: none"> Use of a wide range of tenses used appropriate and consistently in expressing different types of meanings Use of additional verb tenses, modals verbs to discuss hypothetical and conditional meanings use of a range of adverbs to express modality and qualify or emphasise, <i>probably possibly, actually, etc</i> | <ul style="list-style-type: none"> Uses and explains a range of specialised terminology from the work Uses terminology related to the themes and issues e.g. <i>gender roles, tradition and change</i> | <ul style="list-style-type: none"> Clearly intelligible with no problems for audience | <ul style="list-style-type: none"> Self-sustained presentation with little or no reference to notes or prompts |
| 3 | <ul style="list-style-type: none"> Describes the plot and events in details Describes events and their significance, and explains the significance related to themes and issues Gives a personal response, relates elements of the text to self Makes evaluative comment on elements of the work Makes evaluative comment about the work as a whole | <ul style="list-style-type: none"> Fluent interaction, answering questions | <ul style="list-style-type: none"> Describing text describing events and relating them to themes Evaluative comments on the aspects of the work | <ul style="list-style-type: none"> Consistent tense use, either past or present, to describe plot, characters, and make evaluative comments Some use of conditional <i>If I were</i> in relating story to self use of adverbs like <i>done quite well, I'm pretty sure, actually</i> to qualify or emphasize | <ul style="list-style-type: none"> Uses and explains some specialised terms from the work e.g. <i>taiaha</i> Uses some terminology relevant to issues and themes in the work – <i>gender equality</i> | <ul style="list-style-type: none"> Clear intelligible pronunciation, though non – standard pronunciation of some words e.g. <i>Maori</i> pronounced as <i>my-ori</i> | <ul style="list-style-type: none"> Explicit request for assistance, <i>I'm not sure ...</i> Some use of notes |
| 2 | <ul style="list-style-type: none"> Describes main characters and significant events of the plot Relates elements of the work to the themes or issues Makes some evaluative comments about elements of the text, such as the believability of a scene using props Fluent, but some hesitation at times | <ul style="list-style-type: none"> Interacts answering questions and providing reasons and explanations Appropriate turn taking, and sharing of ideas | <ul style="list-style-type: none"> Moderate turns, long turn in describing the plot of the text Discussion of plot, characters themes Simple personal reactions To issues and parts of the text e.g. <i>I liked the story ...</i> | <ul style="list-style-type: none"> Mostly consistent use of present or past tense to re tell narrative Mixture of present and past tense used in discussing characters Mixture of present and past tenses used to discuss issues and give responses to elements of the text | <ul style="list-style-type: none"> Uses terminology names, places, ideas etc Limited range of vocabulary for evaluative comments | <ul style="list-style-type: none"> Intelligible, sounds clearly articulated | <ul style="list-style-type: none"> Responds to questions and adds more information Uses gestures to add meaning, including actions depicted in the film May rely on notes |
| 1 | <ul style="list-style-type: none"> Identifies main characters and events of the plot Identifies theme or issue in the movie or film | <ul style="list-style-type: none"> Responds to questions asked by conversational | <ul style="list-style-type: none"> Questions and answers about plot and main characters | <ul style="list-style-type: none"> Inconsistent use of past tense to retell narrative elements of the story Present tense used to | <ul style="list-style-type: none"> Uses minimal terminology relevant to the work Limited vocabulary | <ul style="list-style-type: none"> Intelligible, but some perceptive errors of production, such as | <ul style="list-style-type: none"> Requests for clarification – <i>What do you mean?</i> May avoid answering difficult questions |



School-Based Assessment in Hong Kong



- In Hong Kong similar constructs in terms of speaking skills and the learning – teaching processes, but different from Australia in who is being rated, who is rating, and the consequences of the rating process, so need to accommodate those differences, eg., see https://www.hkeaa.edu.hk/DocLibrary/SBA/HKDSE/Eng_DVD/sba_standard.html
- For these reasons in Hong Kong we needed to add
 1. Extensive engagement with students and parents from the beginning of the development of the school based-assessment of oral skills, in particular, with the processes to ensure fairness and objectivity in L1 (Cantonese).
 2. Support and training for how teachers needed to disagree, as disagreement was shown to increase reliability (and the sufficiency of information)
 3. The development of nested moderation systems (intra, inter and across schools), auditing processes and statistical moderation, resulting in a successful and sustainable system of SBA.

2012 HKDSE English Language Examination Introduction to the School-based Assessment Component



Introduction to School-based Assessment

SBA in the 2012 HKDSE English Examination

Assessment, Teaching and Learning:
from Principles to Practice

Extensive Reading Programme

Case Studies: Sample Teaching, Learning and
Assessment Sequences

Appendices

Acknowledgements

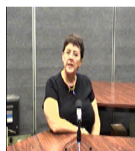


Faculty of Education
The University of Hong Kong



香港考試及評核局
Hong Kong
Examinations and
Assessment Authority

Each of the domains has a set of **descriptors** at each of six levels. This set of levels and descriptors make up the **assessment criteria**. Click [here](#) to view the assessment criteria.



- To [view](#) the explanation of the domains provided by Prof Liz Hamp-Lyons.
- To [view](#) the explanation of the demands of the task types.
- To [view](#) the explanation of the levels of performance.

Note:

Remember that the ability of a student is measured by the quality and range of the oral language produced for assessment, including both fluency and accuracy, not by the speed of completion, or total number of words.

The assessment criteria provide teachers with a common language to describe students' oral language use in ways which are comprehensive, specific and yet constructive.

[Example 1](#) [Example 2](#) [Example 3](#) [Example 4](#) [Example 5](#) [Example 6](#) [Example 7](#)



- To [view](#) the individual presentation in which the student was promoting the film, *Amelie*.

Comment:

This is an individual presentation promoting a film called *Amelie*. The student rather cleverly uses a "poster" about the movie to show to her classmates and to paste her notes onto the back. Her voice projection is fine, and she pronounces common sounds / sound clusters clearly and accurately, though there are some articulation errors and a few words are mispronounced. Her pacing is a little slow and there is some hesitation, suggesting that she is not yet in control of spoken delivery. Her vocabulary is mostly appropriate and she attempts a range of language patterns. Although there are a number of noticeable language errors, none of them seriously impede communication. She shows awareness of errors and some attempts at self-correction. She gives an overview of the film as well as describing the most interesting part of it. She manages to maintain a clear structure without relying too heavily on any notes. She presents her own opinion, although her ideas could be elaborated a bit more. She tries to promote the film by showing her classmates a picture related to the film; although she seems a bit distracted by her own picture and looks at it more than she looks at her audience.

[Back To Top](#)

Before the official reporting date to the HKEAA, there will be a formal within-school standardisation meeting of all the English teachers at each level, chaired by the school SBA Coordinator, to review performance samples and standardise scores. Sample video or audio recordings need not be submitted to the HKEAA but will greatly facilitate quality assurance and within-school standardisation. They can be used for professional development and induction of new S4 teachers into the assessment process. They are very useful for providing feedback to students - for evaluating students' strengths and weaknesses and identifying areas for improvement. Again this is an opportunity for clarification and discussion. The adjusted marks for each student are then listed on a Class Record.

Note:

Within-school standardisation meetings are critical for developing:

- *agreement about what a standard means, i.e. validity;*
- *consistency in and between teacher-assessors, i.e. reliability;*
- *public accountability; and*
- *professional collaboration / support.*

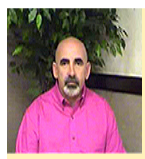


To view Ms Vicky Cheung and Ms Connie Wan describing the intra standardisation process in their school.



To view how in a SBA standardisation session teachers speak freely to each other, disagreeing when necessary and providing their reasons.

At the end of the school year, the SBA District Coordinator will organise an inter-school meeting for professional sharing among the schools within the group. The school SBA Coordinator should take the class records, and sample video / audio-recordings and Assessment Records to this inter-school meeting. The purpose of this meeting is for professional sharing of different schools' assessment processes and student samples in order to help teachers better interpret the standards. Samples of student performance from different schools will be viewed / heard and discussed with reference to the Assessment Criteria. Scores need not be changed at these inter-school meetings, but SBA Coordinators will be asked to report back to the team in their own school. If it is apparent that a particular school's scores are markedly higher or lower as a whole than those from the other schools as a whole, the school team may wish to review their scores. This opportunity for sharing and reflection on the assessment process is critical to ensure reliability and validity across all Hong Kong schools.



To view Prof Dylan Wiliam, Institute of Education, London, talking about the importance of discussion among teachers to ensure validity and consistency in the interpretation of criteria.

4. Conclusions

- We need new concepts, tools and terminology for talking about teachers as ‘raters’ and the quality and trustworthiness of their assessment decision-making
- School-based assessments need to incorporate mechanisms to ensure teachers (and students) are making high-quality and trustworthy assessment decisions, eg., through deep understanding of the assessment construct, repeated participation in common assessment tasks, consistent use of assessment criteria and rubrics, regular benchmarking with other classes/work samples to set exemplars of performance expectations/levels, recording and continually reviewing of data, active engagement in social moderation with other teachers and with their own students, and contributing to online communities of assessment practice.
- Teacher educators, policy makers and researchers need to challenge and reject the still dominant “old” concepts and paradigms, and together develop new tools and resources to support teacher decision-making.

Further reading

- Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement: Issues and Practice*, 22(4), 5-12. Retrieved from http://datause.cse.ucla.edu/docs/smb_dev_2003.pdf
- Davison, C. (2019). Using assessment to enhance learning in English language education. In X. Gao (Ed.), *Second Handbook of English Language Teaching*. Switzerland: Springer Nature.
- Davison C, (2019), Teacher assessment literacy: how do we know what we need to improve?, *English Language Learning*, Vol. 9, pp. 45 - 50, <http://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&filename=YYXI201909010&dbname=CJFDTEMN>
- Davison C, (2007). Views from the chalkface: School-based assessment in Hong Kong. *Language Assessment Quarterly: an international journal*, vol. 4, pp. 1 - 32
- Davison C, (2004), The contradictory culture of classroom-based assessment: Teacher assessment practices in senior secondary English, *Language Testing*, vol. 21, pp. 305 - 334, <http://dx.doi.org/10.1191/0265532204lt286oa>
- Davison C, (2013), Innovation in assessment: common misconceptions and problems. In Hyland K; Wong LLC (ed.), *Innovation and change in English language education*, Routledge, Oxon, pp. 263 - 275
- Davison, C & Leung, C. (2009) Current issues in English language teacher-based. *Assessment TESOL Quarterly*, 43 (3), 393-415. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/j.1545-7249.2009.tb00242.x/pdf>
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Michell M & Davison C, 2020, "Bringing the Teacher Back In": Toward L2 assessment praxis in English as an Additional Language education', *Educational Linguistics*, pp. 23 - 41, http://dx.doi.org/10.1007/978-3-030-35081-9_2
- Moss, P. A. (2003). Reconceptualizing validity for classroom assessment. *Educational Measurement: Issues and Practice*, 22(4), 13-25. Retrieved from <http://deepblue.lib.umich.edu/bitstream/handle/2027.42/72390/j.1745-3992.2003.tb00140.x.pdf?sequence=1>
- Smith, J. K. (2003). Reconsidering reliability in classroom assessment and grading. *Educational Measurement: Issues and Practice*, 22(4), 26-33. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1745-3992.2003.tb00141.x/pdf>
- Timperley, H. (2011). Using student assessment for professional learning: Focusing on students' outcomes to identify teachers' needs(pp. 1-23). DEECD, Victoria. <http://www.education.vic.gov.au/Documents/about/research/timperleyassessment.pdf>
- Timperley, H., Wilson, A., Barrar, H. & Fung, I. 2008, Teacher professional learning and development: Best evidence synthesis on professional learning and development, Report to the Ministry of Education, Wellington. <http://www.oecd.org/edu/school/48727127.pdf>
- Popham. W. J. (2009) Assessment literacy for teachers: Faddish or fundamental?, *Theory Into Practice*, 48(1), 4-11. <http://www.tandfonline.com/doi/pdf/10.1080/00405840802577536>
- Willis, J. (2009). Assessment for learning: A sociocultural approach. In: *Proceedings of : Changing climates : Education for sustainable futures*, 30 November -4 December 2008 , Australia, Queensland, <http://eprints.qut.edu.au/29323/1/29323.pdf>