

Engaging speech pathology students in design and validation of competency based assessment in the workplace

Sue McAllister

Faculty of Health Sciences, The University of Adelaide, sue.mcallister@adelaide.edu.au

Michelle Lincoln

Communication Sciences and Disorders, The University of Sydney, milincoln@usyd.edu.au

Alison Ferguson

School of Humanities & Social Sciences, The University of Newcastle,
Alison.Ferguson@newcastle.edu.au

Lindy McAllister

School of Community Health, Charles Sturt University, lmcallister@csu.edu.au

COMPASS™: Competency based assessment in speech pathology *is an assessment tool designed and developed through a process that included student engagement. This paper will describe why student engagement was initiated, how it was facilitated during project design and validation, and the outcomes of this process. Student opinion regarding the assessment design and process will be described and compared to feedback from clinical educators. The congruence between these perspectives will be highlighted and implications for assessment design and promoting learning in the workplace discussed.*

Keywords: *assessment design, lifelong learning, professional education*

Introduction

In Australia, the discipline of speech pathology has been engaged in competency based education in a sustained manner for 15 years. Speech pathologists are university educated professionals who work in the health and education sectors, with people experiencing communication and/or swallowing difficulties. Speech pathology programs have, from their inception, combined university and workplace based education (practicum) strategies. University courses are accredited by Speech Pathology Australia (SPA), so that their graduates are deemed 'eligible' for membership of SPA. This accreditation process has been guided by the *CBOS Competency Based Occupational Standards for Speech Pathologists - Entry level* (SPAA, 2001) (known as CBOS, pronounced *cee-bos*). The CBOS were developed through a national collaborative project (Dawson, 1993a, 1993b) and subsequently revised in 2001. CBOS identifies concepts and descriptions of professional competence, and makes statements about what level of performance is required for entry into practice (Ferguson, 2006).

SPA's accreditation process focuses on 'outputs'. University programs assemble evidence through descriptions of curriculum and assessment outcomes, that their graduates meet the CBOS competencies. This is different from more 'input' based processes used by allied health professional associations that specify hours related to both university and practicum learning experiences (Coyle, 2007). This has allowed speech pathology programs to develop diverse approaches to curriculum design (Ferguson, 2006).

The common focus on 'output' or competency has also paved the way for the development and national adoption of COMPASS™. COMPASS™ was designed and validated through a PhD research program conducted over 2001 to 2005 (McAllister, 2005), to meet the need for a reliable and valid competency based assessment of speech pathology student performance on practicum. COMPASS™ embodies best practice in ensuring that the assessment effectively facilitates learning through: (a) appropriate processes e.g. authentic assessment (based in the real workplace), formative and summative components and validated rating scales (McAllister, 2005); and (b) content that is based on concepts and descriptions of professional competence that have been identified as meaningful to the profession and effectively describe the development of competence (SPAA, 2001; Dawson, 1993b).

Rationale for student engagement

The development of COMPASS™ was guided by the understanding that assessment drives learning, such that learning, teaching and assessment are inextricably intertwined. Student involvement in the design and validation of COMPASS™ was primarily initiated as a strategy to attend to aspects of the validity of the assessment tool. A secondary consideration was an interest in empowering students to manage their learning through involvement in assessment design, as discussed by Leach, Neutze & Zepke (2001).

Assessment in speech pathology practicum is negotiated within a close personal and working relationship between student and assessor, where the assessor is simultaneously the student's clinical educator. Clinical Educators (CEs) engage in multiple roles (McLeod, Romanini, Cohn & Higgs, 1997). The CE manages the caseload and holds ultimate ethical responsibility for the service provided to the client(s). At the same time, CEs understand their educative role as facilitating the student's transition from supervisee to peer (Brasseur, 1989; Anderson, 1988) and this includes contributing to the final progress decision for the student through his/her assessment of the student's performance on the practicum. Both these facets contribute to the development of a close and complex working relationship where multiple roles need to be negotiated by both the CE and the student. There is no literature in speech pathology regarding the potential impact of this complex relationship upon assessment validity. However, it is well known that the dual roles of educator and assessor create tension or stress in speech pathology (Higgs & McAllister, 2007) and other health occupations (Duke, 1996). This tension may contribute to CEs' concerns regarding objectivity and subjectivity and whether their relationship with the student impacts upon their ability to assess performance fairly (Chapman, 1998; Duke, 1996). Students have expressed the same concerns to all the authors regarding the subjectivity of the clinical education assessment process and possible impact of their relationship with their CEs. Indeed, 'personality conflict' is frequently cited as the reason for poor performance evaluations. Judgement and the legitimacy of its role in assessment also contribute to concerns about subjective influences upon this judgement (Chapman, 1998; Alexander, 1996).

Validity is further threatened when assessment tools or processes are perceived as irrelevant or unwieldy. Cross and colleagues suggest that CEs are unwilling to fully engage with assessment tools or processes in these circumstances (Cross, Hicks, & Barwell, 2001). Govaerts and colleagues (2006) contend that trust in and acceptance of an assessment system by both raters and those being rated is crucial. Lack of engagement would impact both on the learning associated with assessment (Boud, 2000) and the validity and reliability of the assessment tool. Indeed, Neary (2000) found that both CEs and nursing students would choose to not use an assessment tool as intended, if it was perceived as irrelevant to the practicum experience or used educational jargon.

Assessment tools are only valid if used in a valid manner, and while there is little evidence available regarding the reciprocal influence that students and CEs have upon each other in engaging validly with assessment content and process, it is worthy of consideration when designing practicum assessments. Other factors which contributed to the decision to engage students in the development of COMPASS™ were literature suggesting that student engagement with assessment is influenced by their perceptions of the assessment process (Maclellan, 2001), and an intention to safeguard assessment fairness (Lew et al., 2002), in combination with an interest in the rights of students to be empowered within the assessment process (Leach et al., 2001).

Methods

Direct student engagement occurred on three occasions over the design and validation of COMPASS™, and was also conducted with CEs with the questions/format modified as minimally as possible.

Early tool design

The first consultation was conducted during the early stages of assessment design, to develop an understanding of students' perspectives on assessment content and process. Consultations ended when it became apparent that no new information was being generated between the three groups involved – students, CEs, and an expert reference group (comprising the 2nd, 3rd and 4th authors). Questions were based on issues raised in the literature review regarding the nature of valid performance assessment, generic and occupational competencies, and the role of judgment in assessment (McAllister, 2005). Stewart and Shamdasani's (1990) guidelines for developing the interview questions were used to ensure minimal structure to allow for flexibility in pursuing lines of enquiring, ensuring clear wording and ordering questions from general to specific.

Semi-structured interviews were conducted with five students from two universities, with the number of students involved in this process limited by the timing of the interviews and consequent student availability and interest. The same questions were also discussed over a series of six focus groups involving a total of 31 CEs from a variety of settings, who provided placements from students from at least five different programs.

A thematic analysis of the transcripts and field notes was conducted by the first author. Each statement was summarised by a key word or phrase that was then collated into a summary of the key concepts and issues identified by participants. The source of each statement was identified to determine whether the issue was held in common or specific to students or CEs. This summary was then examined and themed categories identified that accounted for all the concepts and

issues. A similar process was undertaken independently by the second author, and any differences in categorisation were identified and resolved. The final categorisation of the transcripts was also examined by third and fourth authors. Analysis of the data showed strong convergence of opinion regarding assessment design by the CEs, students and the expert reference group.

Finalising scale design

A further, less formalised consultation was initiated later in the design phase, where specific decisions were being made regarding the most appropriate scale format for CEs to represent student performance. The literature was equivocal whether scale design was critical for reducing construct irrelevant variance (Gomez-Mejia, 1988; Fay & Latham, 1982; Kingstrom & Bass, 1981). Opinion within the expert reference group driving the project was divided with two members strongly suggesting adopting a scale format (visual analogue scale) that had never previously been used to rate speech pathology student performance (McAllister, 2005). The consultation described above had identified a number of themes that would be compatible with use of a visual analogue scale, but neither group had been specifically asked about preferred scale formats. A focused, detailed consultation interview was conducted with two groups associated with the University of Sydney by the second author. A group of 10 second and third year students and a second group comprising 12 university CEs were shown several scale formats (see Figure 1). Each group was facilitated to discuss and reach consensus regarding the perceived reliability and validity of each format and preferred options for recording their judgments of performance levels.

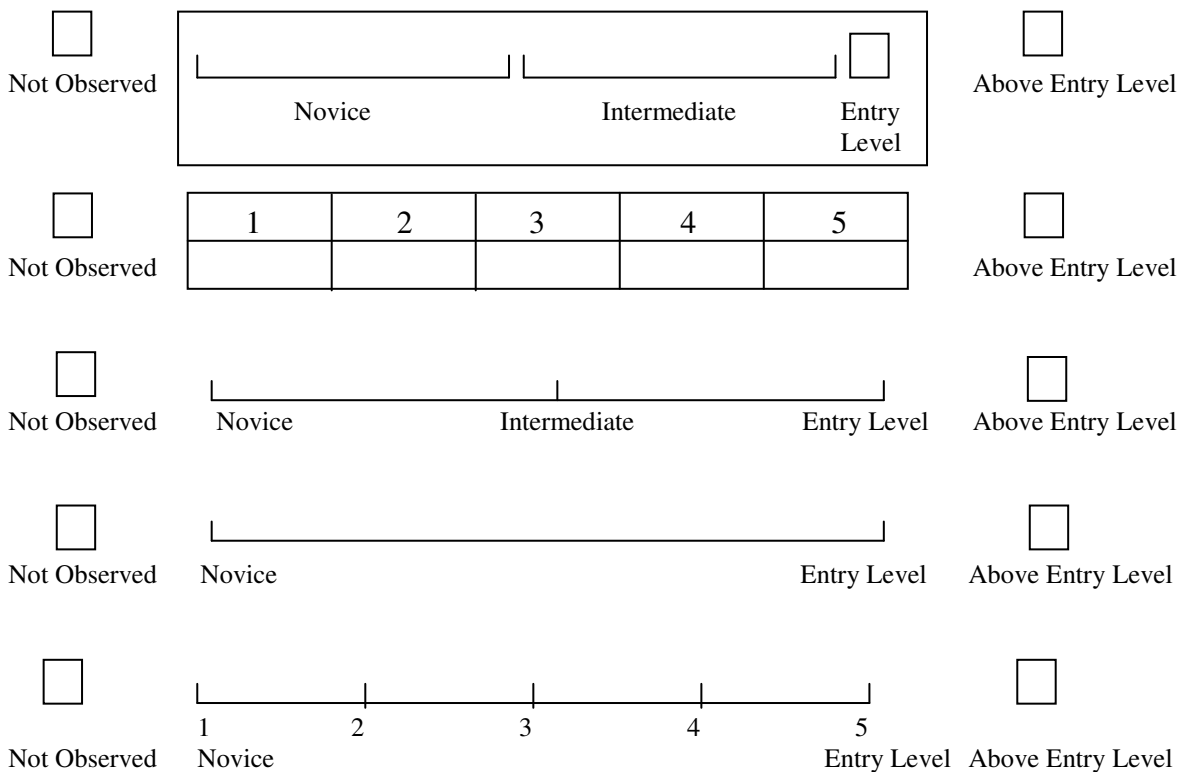


Figure 1. Scale designs presented to students and clinical educators for discussion

Validating field trial of assessment tool

The third consultation involved all students and CEs who participated in a national field trial of the prototype COMPASS™ format, being invited to contribute feedback via a questionnaire on the content and process of the tool. Full details regarding the questionnaire design, analysis and outcome have already been published (McAllister, Lincoln, Ferguson, & McAllister, 2004). In summary, students were invited to provide the same level of feedback as CEs regarding their experience of the assessment tool, perception of the validity of the assessment items, rating procedure, formative and summative assessment processes. Response modes included open and closed questions and a seven category Likert rating scales. The rating scales were analysed using Rasch (rating scale model), to identify if seven discriminable categories were represented by the responses provided (Bond & Fox, 2001; Zhu, 1996; Wright & Masters, 1982). Mann-Whitney U tests were calculated to compare rating patterns between CEs and students.

Results

Early tool design

Students and CEs all made the following points:

1. Occupational (CBOS) and generic competencies were both relevant to assessment of competency
2. Clear, detailed criteria and examples to guide ratings were required. Training and peer review was important to ensure consistency of ratings across placements
3. Clear definitions of what is being rated are essential.
‘Clearer wording need to discern between each level, what the actual meaning of the statements are.’ (CE)

‘I think you need more...probably a series of meetings and criteria to be able to move onto the next level. But I don’t know how do you know what that criteria is?’ (Student)
4. Must be applicable to different caseload and placement complexity and student experience
5. Student performance should be rated on a range of features or qualities
6. Rating scale should reflect progress over time, within and across placements, than do current scales used in assessments.
‘Continuums are good – box or line, illustrates where to aim for and visually show students this.’ (CE)
‘... more of an emerging scale of where they are at.’ (CE)

‘Well, for me, having that rating scale broadened and more defined and that way you have a better understanding of exactly where you are placing within it and whether you have actually made progress or whether it has just been a tiny little shift.’ (Student)
7. Ultimately the assessment process is subjective
8. Assessment impacts on learning and this needs to be attended to.
‘...well pretty much all the assessment comes right at the end, so you get a tiny little bit of assessing along the way ... But it tends to all come in a big lump at the

end. And I think there is a lot of room for doing that, carrying out that more continuous assessment along your placement.’ (Student)

‘It’s moving away from it being achievement and being focused on the learning.’ (CE)

9. Should include a formative and summative component with students being involved in the assessment process
10. Opportunity to make or receive comments was highly valued

Finalising scale design

There were minor differences in comments made by CEs and students, with both groups suggesting scale designs different from those offered for discussion. Both groups preferred a scale format that was very similar to a visual analogue scale, so that a continuum was represented and descriptive labels were preferred over numbers. Students differed in that they wanted vertical marks placed along the line to indicate five stages of development, as they were concerned that a very small and possibly chance variation in the point at which the line was marked would have a disproportionate effect upon their assessment result. See Figures 2 & 3 for the two suggested formats. Overall, this consultation suggested that a modified version of a VAS format might be acceptable to both groups.

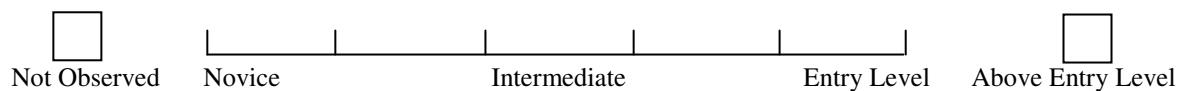


Figure 2. Students' suggested scale format

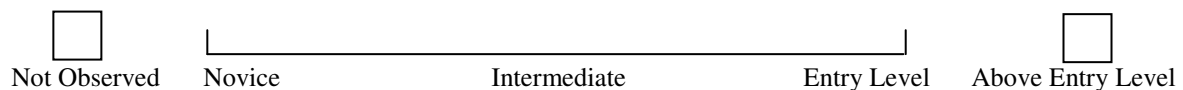


Figure 3. Clinical Educators' suggested format

Validating field trial of assessment tool

Forty-one percent (88 of 217) of students who had assessments returned responded to the questionnaire. However the overall return rate may have been higher, as some students were reported as not being directly involved in the assessment and therefore, did not return a questionnaire. CEs returned 66 of 96 (70%) feedback forms. A broad range of student and CE experience, university programs and placement types and length were represented. Rasch analysis found that ratings were meaningful only as three categories: agreement, neutral or disagreement. All ratings were positive and Mann-Whitney U tests calculated using the three categories indicated that CEs and students rated in a similar pattern for 19 of the 21 evaluative statements (see Table 1 for an example of results for an evaluative statement), including those related to the rating scale format. The remaining two items were significantly more likely to be rated as ‘neutral’ rather than ‘agree’ by students ($p < .05$).

Table 1 Example of questionnaire item and responses

Overall satisfaction			
Statement rated	Response category	CE (% responding)	Student (% responding)
How would you rate your OVERALL satisfaction with the research Assessment Tool as an assessment of your competency in your clinical placement? NB ratings were from Low (1) to High (7)	High	85.2	86
	Neutral	7.4	7
	Low	7.4	7
CE N = 66 Student N = 86	Median rating	5.5	6

Discussion

Student involvement in assessment design was very helpful, as the strong congruence of student opinions with that of CEs and experts enabled the researchers to move forward with confidence in designing the assessment tool. This early input into design is likely to have contributed to the strong agreement in student and CE evaluation of the validity of the research version of COMPASS™, including their perception of the scale type used. This eliminated potential threats to the tool’s structural and consequential validity (Messick, 1996). Structural validity, whereby the rating process is congruent with the construct domain, was supported by ensuring that rating behaviour was less likely to be negatively impacted through lack of engagement by the student and/or CE either directly or through indirect influences arising from their close working relationship. Consequential validity would have been impacted should poor engagement in assessment processes negatively affect learning.

These results also suggest that speech pathology students are sophisticated ‘consumers’ of assessment given the strong agreement between their opinion, CEs and the expert reference group. While students were concerned about fairness and how this might impact their progression through their program, they were also strongly interested in the impact of assessment upon their learning. It would appear that speech pathology students understand and critically reflect on their learning, and development of professional competence, features identified by Brodie and Irving (2007) as important for quality practicum learning experiences.

If this is the case, speech pathology students and their CEs are engaging in an active relationship, with construction of learning at its core. This process will facilitate the development of lifelong learning skills that are so important for maintaining a lifetime of professional competence (Billett & Somerville, 2004; Hager, 2004). This would equip students to identify post graduation what they need to learn and what constitutes a good performance (Boud & Falchikov, 2006), so that competence can be maintained in response to changing professional practices. This is in contrast with other findings in the literature e.g. that physicians have limited ability to self assess (Davis et al., 2006) or that students have generally unsophisticated understanding of assessment (Maclellan, 2001).

In summary, while the evidence presented in this paper is limited in scope, it highlights the effectiveness of involving students in assessment design. It also identifies that students are able to be sophisticated participants in the assessment process. Therefore, recommendations that students' lifelong learning for professional practice should be promoted through involving them as co-constructors of the learning and assessment process on placement (Boud & Falchikov, 2006) appear to be feasible and worthy of further exploration.

References

- Alexander, H. (1996). Physiotherapy student clinical education: The influence of subjective judgements on observational assessment. *Assessment & Evaluation in Higher Education*, 21(4), 357-366.
- Anderson, J. (1988). *The Supervisory Process in Speech-Language Pathology and Audiology*. Boston: College Hill.
- Billett, S., & Somerville, M. (2004). Transformations at work: identity and learning. *Studies in continuing education*, 26(2).
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*: Mahwah, NJ: Lawrence Erlbaum and Associates.
- Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167.
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment and Evaluation in Higher Education*, 31(4), 399-413.
- Brasseur, J. (1989). The supervisory process: A continuum perspective. *Language, Speech and Hearing Services in Schools*, 20, 274-295.
- Brodie, P., & Irving, K. (2007). Assessment in work-based learning: investigating a pedagogical approach to enhance student learning. *Assessment and Evaluation in Higher Education*, 32(1), 11-19.
- Chapman, J. (1998). Agonising about assessment. In D. Fish & C. Coles (Eds.), *Developing Professional Judgement in Health Care: Learning through the critical appreciation of practice* (pp. 157-181). Oxford: Butterworth-Heinemann.
- Coyle, G. (2007). Mandated hours in NZ allied health degrees. *Health Professional Education: A Multi-Disciplinary Journal*, 9(1), 67-77.
- Cross, V., Hicks, C., & Barwell, F. (2001). Exploring the gap between evidence and judgement: Using video vignettes for practice-based assessment of physiotherapy undergraduates. *Assessment & Evaluation in Higher Education*, 26(3), 189-212.
- Davis, D. A., Mazmanian, P. E., Fordis, M., Harrison, R. V., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence A systematic review *Journal of the American Medical Association*, 296(9), 1094-1102.
- Dawson, V. (1993a). Assessing the standards - Stage II of the Competency Based Occupational Standards project. *Australian Communication Quarterly*(Summer), 8-9.
- Dawson, V. (1993b). Competency based standards for speech pathologists. *Australian Communication Quarterly*(Autumn), 9-10.
- Duke, M. (1996). Clinical evaluation - difficulties experienced by sessional clinical teachers of nursing: A qualitative study. *Journal of Advanced Nursing*, 23(2), 408-414.
- Fay, C. H., & Latham, G. P. (1982). Effects of training and rating scales on rating errors. *Personnel Psychology*, 35, 105-117.
- Ferguson, A. (2006). Competency based occupational standards: Influences on Australian Speech Pathology Education. *Folia Phoniatica et Logopaedica*, 58(23-32).
- Gomez-Mejia, G. (1988). Evaluating employee performance: Does the appraisal instrument make a difference? *Journal of Organizational Behavior Management*, 9(2), 155-172.
- Govaerts, M. J. B., van der Vleuten, C. P. M., Schuwirth, L. W. T., & Muijtjens, A. M. M. (2006). Broadening perspectives on clinical performance assessment: Rethinking the nature of in-training assessment. *Advances in Health Sciences Education*.
- Hager, P. (2004). Lifelong learning in the workplace? Challenges and issues. *Journal of Workplace Learning*, 16(1/2), 22-32.
- Higgs, J. & McAllister, L. (2007) Being a clinical educator. *Advances in Health Sciences Education*, 12, 2, 187-199.

- Kingstrom, P. O., & Bass, A. R. (1981). A critical analysis of studies comparing Behaviorally Anchored Rating Scales (BARS) and other rating formats. *Personnel Psychology*, 34(2), 263-289.
- Leach, L., Neutze, G., & Zepke, N. (2001). Assessment and empowerment: Some critical questions. *Assessment & Evaluation in Higher Education*, 26(4), 293-305.
- Lew, S. R., Page, G. G., Schuwirth, L. W. T., Baron-Maldonado, M., Lescop, J., Paget, N., et al. (2002). Procedures for establishing defensible programmes for assessing practice performance. *Medical Education*, 36, 936-941.
- MacLellan, E. (2001). Assessment for learning: the differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education*, 26(4), 307-318.
- McAllister, S. (2006). *Competency based assessment of speech pathology students' performance in the workplace*. PhD thesis. The University of Sydney, Faculty of Health Sciences, Sydney (Accessible from the Online Digital Thesis Repository: <http://hdl.handle.net/2123/1130>).
- McAllister, S. M., Lincoln, M., Ferguson, A., & McAllister, L. (2004). User Evaluation of the Content and Processes of an Assessment of Student Performance in Workplace Settings. In *IALP World Congress, 2004, Brisbane, Australia*.
- McLeod, S., Romanini, J., Cohn, E.S., & Higgs, J.. (1997). Models and roles in clinical education. In L.McAllister, M. Lincoln, S. McLeod & D. Maloney (Eds.). *Interdisciplinary Approaches to Clinical Education* (pp. 27-64). London: Stanley Thornes.
- Messick, S. (1996). Validity of performance assessments. In G. W. Phillips (Ed.), *Technical Issues in Large-Scale Performance Assessment* (pp. 1-18). Washington: National Centre for Education Statistics.
- Neary, M. (2000). Supporting students' learning and professional development through the process of continuous assessment and mentorship. *Nurse Education Today*, 20, 463-474.
- SPAA. (2001). *Competency-Based Occupational Standards for Speech Pathologists (Entry Level)*. Melbourne: Speech Pathology Association of Australia Ltd.
- Stewart, D. W., & Shamdasani, P. (1990). *Focus Groups: Theory and practice* (Vol. 20). Newbury Park: SAGE Publications, Inc.
- Wright, B. D., & Masters, G. N. (1982). *Rating Scale Analysis*. Chicago: MESA Press.
- Zhu, W. (1996). Should total scores from a rating scale be used directly? *Research Quarterly for Exercise and Sport*, 67(3), 363-372.