



Archived at the Flinders Academic Commons:

<http://dspace.flinders.edu.au/dspace/>

This is an Accepted Manuscript of an article published by Taylor & Francis in *the International Journal of Speech-Language Pathology* on 30 August 2013, available online:

<http://>

www.tandfonline.com/10.3109/17549507.2013.809603

<http://dx.doi.org/10.3109/17549507.2013.809603>

Please cite this article as:

Hill, A., Davidson, B., McAllister, S.M., Wright, J. and Theodorus, D. (2014). Assessment of student competency in a simulated speech-language pathology clinical placement. *International Journal of Speech-Language Pathology*, 16(5) pp. 464-475.

© 2014 The Speech Pathology Association of Australia Limited. All rights reserved. **Please note** that any alterations made during the publishing process may not appear in this version.

Assessment of student competency in a simulated speech-language pathology clinical placement

ANNE E. HILL¹ BRONWYN J. DAVIDSON² SUE MCALLISTER³ JUDITH WRIGHT¹
DEBORAH G. THEODOROS¹

¹The University of Queensland, Brisbane, Australia

²The University of Melbourne, Melbourne, Australia

³Flinders University, Adelaide, Australia

Correspondence to:

Anne E. Hill, School of Health and Rehabilitation Sciences, Division of Speech Pathology, the University of Queensland, St Lucia, QLD, Australia, 4072.

Tel: +61 7 3365 1461

Fax: +61 7 3365 1877

Email: ae.hill@uq.edu.au

Abstract

Clinical education programs in speech-language pathology enable the transition of students' knowledge and skills from the classroom to the workplace. Simulated clinical learning experiences provide an opportunity to address the competency development of novice students. This study reports on the validation of an assessment tool designed to evaluate speech-language pathology students' performance in a simulated clinical placement. The Assessment of Foundation Clinical Skills (AFCS) was designed to link to concepts and content of COMPASS[®]: Competency Assessment in Speech Pathology, a validated assessment of performance in the workplace. It incorporates units and elements of competency relevant to the placement. The validity of the AFCS was statistically investigated using Rasch analysis. Participants were 18 clinical educators and 130 speech-language pathology students undertaking the placement. Preliminary results support the validity of the AFCS as an assessment of foundation clinical skills of students in this simulated clinical placement. All units of competency and the majority of elements were relevant and representative of these skills. The use of a visual analogue scale which included a pre-Novice level to rate students' performance on units of competency was supported. This research provides guidance for development of quality assessments of performance in simulated placements.

Introduction

Tertiary programs in speech-language pathology incorporate clinical placements to provide students with the opportunity to develop essential clinical competencies prior to graduation (McAllister, Lincoln, Ferguson, & McAllister, 2011). Clinical placements bridge students' learning from classroom to workplace practice by providing them with opportunities to translate their theoretical knowledge into practical applications, be socialised into the profession and achieve competencies expected by the profession (Lincoln, 2012; McAllister, 1997). For many students, however, the gap between classroom and clinic appears wide and challenging due to difficulty in adapting to alternative demands of the learning environment and inability to attend to the multiplicity of factors implicit in developing competency. The connection between theory learnt in the classroom and its application in practice can be tenuous and arbitrary, often leading to loss of what was already known (Le Maistre & Paré, 2004).

The inclusion of guided clinical learning experiences prior to graduation is expected of speech-language pathology programs around the world and is enshrined in their accreditation requirements. For example, the American Speech-Language-Hearing Association (ASHA) requires that graduates accrue 400 hours of supervised clinical practice prior to graduation (ASHA, 2009), while the Canadian Association of Speech-Language Pathologists and Audiologists (CASLPA) prescribes 350 hours of supervised clinical experience (CASLPA, 2012). The Speech Pathology Association of Australia (SPA) mandates that graduates demonstrate competency in direct clinical practice at an appropriate level with a range of clients (Competency-based Occupational Standards for Speech Pathologists: Entry Level - Revised [CBOS]; SPA, 2011). Irrespective of the country context, achievement of expected graduate skills is obtained through a range of different placement

models, with traditional models (1:1 student-to-supervising speech-language pathologist in a block or weekly placement) most prevalent (Sheepway, Lincoln, & Togher, 2011).

It is common practice for Australian universities to provide clinical learning opportunities for students prior to them undertaking workplace placements facilitated by practising speech-language pathologists (Lincoln, 2012). Early clinical placements commonly include four to six students with one clinical educator in a collaborative or group supervision model (Sheepway et al., 2011). In early stages of clinical programs, clinical educators are focussed on students' development of foundation clinical skills through providing, for example, opportunities for observation, self-evaluation and feedback (Linguist, Engardt, & Richardson, 2004; McAllister, 2005). The structured nature of foundation placements meets the needs of novice students for both the knowledge and the scaffolding they require to manage clinical placement requirements effectively (Billett, 2011; McAllister et al., 2011).

A number of authors (Kramer, Copley, & Nelson, 2004; O'Kane, 2010; Sherer, Morris, Graham, & White, 2006; Weddle & Sellheim, 2011) have reported allied health students' perceptions of the impact of experiences designed to facilitate their transition to placements in real world workplaces. While these studies are primarily descriptive in nature, they also provide evidence of the perceived value of such experiences. O'Kane (2010) reported on a program for nutrition and dietetics students designed to provide enculturation into their profession and a timely application of learned theory. Students participated in sessions in a hospital with some ward activities (including history-taking and talking with patients) and debriefing workshops prior to attending workplace placements. Students reported that the program increased their communication skills, confidence and skills for working with clients. The safe setting and real life experiences were considered to be useful in preparing students for further workplace placements (O'Kane, 2010).

Group supervision is also seen as valuable for students at early stages of their clinical learning. Studies of physiotherapy (Sherer et al., 2006; Weddle & Sellheim, 2011) and occupational therapy students (Kramer et al., 2004) engaged in early, group supervision clinical placement models have concluded that students find such experiences support their transition between classroom and workplace clinic. They also assist students in understanding their professional role, preparing them for future practice, and increasing their confidence. In addition, learning from peers within a group supervision model was highly valued (Weddle & Sellheim, 2011).

Whilst each of the reported studies differs in terms of the nature of the early clinical learning experience, commonalities exist. These include a gradual introduction to clinic through practice with clients and working within groups, the value of peer learning and focus on educator feedback, and the opportunity to develop confidence and competence in skills. Each of the studies reported on clinical placements involving real clients in either university or external contexts. The provision of traditional placements such as these is becoming more problematic due to increasing numbers of speech-language pathology programs throughout Australia. The subsequent increases in student numbers and required placements (Lincoln, 2012) warrant further consideration with respect to potential alternatives.

Simulated learning environments have been proposed to offer a valuable alternative to traditional placement models for novice students (Hill, Davidson, & Theodoros, 2010). Standardised patients are an accessible form of simulation and are commonly utilised in health professional education programs (Hill et al., 2010; Paparella-Pitzel, Edmond, & DeCaro, 2009). Standardised patients are actors who are trained to consistently portray a designated role in order to meet specified learning objectives (Barrows, 1971). They are reported to offer a number of advantages for student learning: they allow for targeted skill development; provide practice without fear of making mistakes; and offer a variety of

opportunities for receiving feedback (Bradley, 2006; Lane & Rollnick, 2007; Lysaght & Hill, 2010; Zraick, 2012).

A small number of studies from the field of speech-language pathology clinical education have reported on the use of standardised patients in educating students in the management of aphasia (Edwards, McGuinness, & Rose, 2000; Zraick, Allen, & Johnson, 2003), interaction with clients with voice disorders in a range of typical clinical tasks such as explanation of therapy techniques (Syder, 1996), and supporting students' development of foundation clinical skills such as interaction and communication (Hill, Davidson, & Theodoros, 2012). These studies have supported the use of standardised patients within speech-language pathology programs.

In addition, a recent evaluation of the use of simulation in speech-language pathology programs in Australia found that simulated learning environments (including standardised patients) have significant potential for complementing traditional placements within clinical programs (Theodoros, Davidson, Hill, & MacBean, 2010). However, it acknowledged that there is limited evidence of their suitability within speech-language pathology and that there is difficulty in generalising evidence gained from research in other professions (Theodoros et al., 2010). The effectiveness of simulated learning programs has primarily been evaluated via student and educator perceptions. Valid assessment of student performance in such environments would provide further information on how effectively this educational strategy supports the development of students' ability to apply knowledge to clinical practice.

Assessment of clinical competence is complex, not least in its consideration of the conceptualisation of competency and how assessment should occur (McAllister, Lincoln, Ferguson, & McAllister, 2010). Importantly, an assessment must offer the opportunity to provide targeted, formative feedback to students throughout the assessment process to assist their clinical learning (Hancock & Brundage, 2010; Norcini & Burch, 2007). Competency

Assessment in Speech Pathology (COMPASS[®]: McAllister, Lincoln, Ferguson, & McAllister, 2006) is embedded within Australian and New Zealand speech-language pathology curricula. It is used to undertake formative and summative assessment of students' development of competency when working directly with clients. COMPASS[®] is a psychometrically validated assessment tool that is educationally sound and is designed for use in workplace settings (McAllister et al., 2006; McAllister et al., 2010). For COMPASS[®], speech-language pathology (SLP) competency is considered to be observable behaviour that arises from combinations of occupational and generic aspects of practice that in turn arise from combinations of various types of knowledge, skills and personal qualities (McAllister, 2006). In addition, quality assessment practices such as provision of formative feedback and ensuring multiple observations of student performance are embedded in its design. COMPASS[®] provides a framework that supports students' and clinical educators' understanding and assessment of clinical competency and is likely to be relevant to assessment of students' performance in a simulated environment. Furthermore, using the same framework in both types of clinical learning environments may support students' transition in applying knowledge gained in a simulated learning experience to working directly with real clients.

Validation of assessment tools has traditionally encompassed the concepts of content-related, criterion-related and construct-related validity. Messick (1995) extended these traditional components into a unified construct framework which integrates content, criterion, and consequences of test use together with investigation of score meaning and interpretation (Messick, 1995). Six interrelated validity categories are relevant when evaluating the validity of measures yielded by an assessment tool and the validity of using these measures for assessment decisions. These validity categories are operationalised as follows:

Content: the content should be related to the construct of speech-language pathology competency.

Substantive: the competencies rated should provide a good sample of the content and processes of 'speech-language pathology competency' and the assessment should provide examinees with the opportunity to demonstrate competency in speech-language pathology practice.

Structure: the scoring should relate to what is already known about the structure of the construct, in this case, that speech-language pathology competency is developmental (McAllister et al., 2011).

Generalisability: the degree of speech-language pathology competency represented by the assessment results is likely to be represented in other tasks that sample speech-language pathology competency.

External: the assessment scores relate logically to other measures of speech-language pathology competency or behaviours that are representative of speech-language pathology competency.

Consequential: the consequences of the way in which assessment results are or might be used are considered.

This study aimed to investigate the validity of an assessment tool developed to specifically assess students' foundation speech-language pathology clinical competencies in a simulated clinical placement based on Messick's framework (1995).

Method

Clinical context

The simulated clinical placement reported in this study was designed to support foundation clinical learning of speech-language pathology students in the first stage of their clinical program. A hybrid model of simulation was adopted and included standardised patient interviews, group workshops, and case-based discussions in addition to a ‘real client’ screening assessment. Within interviews, students obtained a case history from, and provided management information to, standardised patients who portrayed the role of a parent or grandparent of a child presenting with a speech delay. The simulated clinical placement included three-hour clinic sessions held once a week for 12 weeks. Students worked in groups of six with a clinical educator. The structure of the sessions is detailed in table 1. The learning objectives formulated for this simulated clinical placement had a focus on the development of foundation clinical skills of communication, interviewing, professionalism, and management of the case history process.

INSERT TABLE 1 ABOUT HERE

Development of Assessment

An expert group of experienced academic staff (n=3) and clinical educators (n=6) engaged in a group discussion to determine essential features and components of COMPASS[®] which were relevant in the development of a student assessment for the simulated clinical placement. In addition, the specific and unique learning objectives and clinical activities embedded within the simulated clinical placement were considered. Of particular importance was the requirement that the assessment process would provide feedback to students throughout the placement and would articulate with COMPASS[®] which would be used in later workplace placements. The following features of COMPASS[®] were determined to be applicable to the simulated learning context: the structure of units of competency and elements within the units; the visual analogue scale (VAS) to rate students’ competency; the provision of behavioural descriptors to describe levels of performance on

each competency; the overall competency rating at midway (formative) and end (summative) placement; and the inclusion of an ‘at risk’ statement indicating when supportive action was required to assist students in developing the required level of performance for each unit.

The expert group agreed that seven of the 11 units of competency assessed by COMPASS[®] were relevant to the learning objectives for the simulated clinical placement: four professional competencies (Reasoning, Communication, Life-long Learning and Professionalism) and three occupational competencies (Assessment, Analysis and Interpretation, and Planning of Speech Pathology Intervention). CBOS competency units 4, 5, 6, and 7 (Speech Pathology Intervention, Planning, Maintaining and Delivering Speech Pathology Services, Professional, Group and Community Education, and Professional Development) were not considered suitable for inclusion as the simulated clinical placement did not provide opportunity for students to be assessed on those units. For example, students did not participate in any intervention with clients. Due to the structured nature of the program, they did not engage in service delivery decisions, nor did they create networks within or outside of their clinic. In addition, discussion within the expert group established that not all individual elements for each COMPASS[®] unit of competency were relevant to the learning activities of the current placement. Four to seven placement-specific elements were identified for each of the seven units (total of 36 elements).

The VAS used in COMPASS[®] was agreed to be a valuable way of recording the development of students’ competency. The VAS scale of the paper version of COMPASS[®] is a single line bounded by two upright marks. ‘Novice’ level is indicated at the left end, ‘Intermediate’ around the middle and ‘Entry level’ at the right.

As the simulated clinical placement was the students’ first placement, they were required to demonstrate a minimum of ‘Novice’ level of competency at the end of the placement. The VAS for the simulated clinical placement assessment needed to be extended

below Novice to capture performances below this level for both formative and summative purposes. This important transition from pre-Novice to Novice was marked with a small upright line at the 19mm point, such that Novice and above was from 19mm onwards, with a total line length of 105 mm. The VAS included a right-pointing arrow at its right extremity to indicate the development of student competency beyond Intermediate level. The field trial of COMPASS[®] found that the length and type of representation of the VAS (paper or online) did not affect the way in which it was used for rating (McAllister, 2006).

Inclusion of a pre-Novice category required the development of behavioural descriptors to guide clinical educators' ratings. These were developed based on discussion with the expert group and drew upon the three elements that underpin the behavioural descriptors in COMPASS[®]: transforming knowledge into practice, dealing with complexity, and level of independence (McAllister et al., 2011). The COMPASS[®] descriptors for Novice and Intermediate performances on the competencies were included for clinical educators to rate students whose skills were more highly developed.

In addition to the VAS, consistency of performance on each element was recorded in order to provide formative feedback. Novice students have a propensity to be under-confident and anxious in their interaction with clients (Benner, 2001; Chan, Carter, & McAllister, 1994) and variable in their performance across different competency areas (McAllister et al., 2011). It was postulated that these characteristics would contribute to inconsistency of performance for students undertaking their first clinical placement and that this may be an important characteristic to determining whether a student passed or failed the simulated clinical placement. A categorical Consistency scale was provided to record the consistency of the student's performance on each element as 'not applicable', 'rarely', 'mostly' and 'consistently'. An opportunity for comments from both clinical educator and

student was provided. The final assessment format was called the Assessment of Foundation Clinical Skills (AFCS). The AFCS may be obtained by contacting the first author.

Participants

University employed clinical educators and speech-language pathology students enrolled in the simulated clinical placement at an Australian University in 2009 and 2010 consented to participate (student n= 130, 127 women and three men; clinical educator n=18, all women). Students were enrolled in the simulated clinical placement in the first semester of the second year of their four year speech-language pathology program. Students were aged from 18 to 47 years with a mean age of 20 years. Clinical educators had an average of 18 years of clinical experience (range of 6 to 28 years) and an average of eight years of experience as a clinical educator (range of 0 to 28 years).

Procedure

Ethical clearance was obtained from the Behavioural and Social Sciences Ethical Review Committee of The University of Queensland, Australia. The structure and components of the AFCS were explained to students prior to their simulated clinical placement. Clinical educators attended a two hour meeting in which the placement structure and assessment were discussed. The AFCS was introduced to both educators and students prior to the start of the clinical placement with attention drawn to each of the units and elements and their application to the placement structure, each of the assessment rating components, and the behavioural descriptors. A demonstration of rating with reference to specific examples was provided.

Formative feedback was provided to students during each session of the simulated clinical placement. Assessment ratings on the AFCS were undertaken at the midway point (week 5) and the end point (week 12) of the placement. At each rating point, clinical educators were asked to consider students' performances within all components of the

simulated clinical placement that had been undertaken to that point (see table 1). Clinical educators first used the Consistency scale to rate each student's performance on each element of each unit. They then rated each student's overall performance for each competency unit on the VAS. Following completion of the 'at risk' statement and the overall competency rating, comments were provided by the clinical educator and then by the student. At the completion of the placement, the AFCS for each student was collected, de-identified and entered into a data file for statistical analysis. As the assessment process for the mid and final assessment was identical and conducted on two separate occasions, data was combined to provide an overall picture of students' development of competency and clinical educators' use of the AFCS.

Data Analysis

Rasch analysis (Rating Scale Model) (Bond & Fox, 2007) was used during validation of COMPASS[®] (McAllister, 2006) and therefore, in consideration of its shared features, was suitable for analysis of the AFCS. Rasch analysis has previously been used in the fields of health and social sciences in validating assessment tools (e.g., Baylor et al., 2011; Beglar, 2010; Lim, Rodger, & Brown, 2009).

Rasch analysis has particular utility for validation of performance assessments in a number of areas. Firstly, it is a sample-independent analysis that compares a set of observed data with a prediction of how the data should be represented in quality assessment (Bond & Fox, 2007). This process allows the examiner to determine if specific items on the assessment do not adhere to expectations, for example, are rated more difficult than would be expected. It also highlights examinees whose performance is variable and unpredictable (McAllister et al., 2006). Secondly, its use in validation of a rating scale enables demonstration of developmental change (McAllister et al., 2010). The performance of a group of people on a rating scale can be classified into distinct functional categories that

represent meaningful and equal sized increases in performance levels. This allows for identification of increases in performance over time. Finally, Rasch analysis can identify if the assessment samples a single underlying or latent trait (Baylor et al., 2011), in this case students' clinical competency. Specific analysis procedures used in this study were modelled on the procedure conducted by McAllister (2006) in the validation of COMPASS[®].

Visual Analogue Scale analysis. Clinical educator ratings on the VAS were converted into numerical scores (millimetre measurements; mm) according to their measurement from the start of the VAS. The start of the category for a 'Novice' (passing) level of performance was represented by a vertical line 19mm from the start of the VAS. The first step in analysis of the VAS was to determine functional categories to represent meaningful levels of performance on each of the competencies To be deemed a 'functional category', data were subjected to the rule-based procedures outlined by Linacre (2002) and followed by McAllister (2006).

A systematic procedure of coding and recoding data was used to establish functional categories. Five iterations of this process were required to achieve adherence to Linacre's (2002) requirements. Each iteration and the degree to which the resultant categories met Linacre's (2002) eight criteria are outlined in table 2. As a result of the analysis, five functional rating categories were established along the VAS. Rasch analysis was then continued to evaluate the use of the VAS for each competency by clinical educators. The process of analysis followed the procedure of Bond and Fox (2007) and determined;

- Unidimensionality: fit statistics, reported as infit and outfit mean squares, provide information on how the data fit the rating scale model and whether they suggest that competencies rated represent a single construct, in this case, foundation clinical skills.
- Item difficulty: which competencies students found most and least difficult.

- Person reliability: the likelihood of the person ordering being replicated if the students were given another parallel set of competencies to be rated on that relate to the same underlying construct of competency.
- Item reliability: the likelihood of the competencies being rated in same way if they were used with another group of students.

INSERT TABLE 2 ABOUT HERE

Consistency categorical scale analysis. Analysis of the Consistency scale was carried out in the same manner as for the VAS. Firstly, Linacre’s (2002) guidelines were used to determine how each of the three categories functioned (‘rarely’, ‘mostly’, ‘consistently’). Secondly, analysis determined the fit statistics, item difficulty, person reliability and item reliability of each of the 36 elements of competency rated.

Results

Visual Analogue Scale (VAS)

Categories. Details of the five functional categories are recorded in table 3. VAS measurements represented by each category were as follows: Category 1 - 0 to 18mm; Category 2 - 19mm to 21mm; Category 3 - 22mm to 28mm; Category 4 - 29mm to 38mm; and Category 5 - 39mm and above. Each category had over 10 observations. Observations had a unimodal peak with a pivot point at category 3, as opposed to regular distribution. This is in line with Linacre’s (2002) suggestion that a central, unimodal peak reflects a “substantively meaningful distribution” (p. 5). The observed averages for the five categories advanced monotonically. The infit and outfit measures were all under 2.0. Step calibration advanced. Category measures advanced by 3.07, 4.37, 4.59, and 3.29. Coherence levels for all five categories were acceptable.

INSERT TABLE 3 ABOUT HERE

Fit statistics. The infit mean squares for all competencies rated ranged between 0.89 and 1.11. The outfit mean squares for all competencies ranged between 0.76 and 1.11. These results (reported in table 4) are in accordance with Bond and Fox's (2007) recommended guidelines that infit and outfit mean squares for an assessment based on rating performance should range from 0.6-1.4 if the competencies rated are to be considered as sampling one construct.

Item difficulty. Analysis determined which competencies were more difficult than others according to how likely students of various ability levels were to be rated high/low on each competency. The item difficulty was represented by the competency's measure, with a higher measure indicating a more difficult competency and a lower measure indicating an easier competency. Table 4 reports the competency measures and their equivalent standard error. Unit 7, CBOS Competency 3.0 Planning of Speech Pathology Intervention, was the most difficult competency for students to achieve a high rating on with a measure of 1.33 and standard error of 0.18. Unit 4, Generic Professional Competency (GPC) 4.0 Professionalism, was the easiest competency with a measure of -1.73 and a standard error of 0.14.

INSERT TABLE 4 ABOUT HERE

Person ability, person reliability and item reliability. A large spread of person ability was observed for the AFCS, with a range from 10.18 to -10.03. Linacre (2002) suggested that a person ability range as small as 6 is satisfactory. The AFCS had a person reliability of 0.95 and an item reliability of 0.98, with acceptable levels for both measures being 0.80 (McAllister, 2006).

Consistency Categorical Scale

Rasch analysis was undertaken on the Consistency scales of the AFCS. This analysed clinical educators' ratings of students' performance with reference to the 36 specific skill elements within the seven competency units.

Categories. Table 5 details the category information for the three categories (rarely, mostly, consistently) with reference to Linacre's (2002) guidelines for functional categories. The requirement of each category having over 10 observations was met. Observations were not regularly distributed as the observed count in each category differed by over 2000 observations. Clinical educators rated the majority of students as 'consistently' (category 3), skewing the observed count. The observed averages advanced monotonically. The outfit mean square of Category 1 was 2.01 and therefore above the recommended level. Categories 2 and 3 adhered to the guideline with outfit mean squares of 1.28 and 0.95 respectively. Step calibration advanced. Category measures advanced by 2.97. Coherence levels were all above 48%.

INSERT TABLE 5 ABOUT HERE

Fit statistics. Infit mean squares for the ratings of all competencies fell within the range proposed by Bond and Fox (2007). The outfit mean squares of six competencies fell outside this range. Five competencies were above the stipulated level: GPC Unit 4.0 Professionalism, element c. 'Maintains professional appearance (including suitable dress and hair) and acts in a professional manner' (outfit mean square of 4.39), element a. 'Meets agreed deadlines for all clinical tasks' (outfit mean square of 2.42), element b. 'Gains informed consent and maintains confidentiality' (outfit mean square of 2.40) and element e. 'Follows policies and procedures of the clinic, especially in relation to use of furniture, equipment and resources' (outfit mean square of 2.18) and GPC unit 2.0 Communication, element g. 'Other written records (e.g. Clinical workbook tasks) are clear and complete'

(outfit mean square of 3.02). GPC unit 3.0 Lifelong learning element c. 'Adapts behaviour to address learning goals' was lower than recommended (outfit mean square of 0.51).

Item difficulty. CBOS Competency Unit 3.0 Planning of Speech Pathology Intervention element b. 'Discusses rationale for various therapy approaches', was the most difficult competency with a measure of 2.74 and standard error of .19. GPC Unit 4.0 element d. 'Demonstrates and maintains respect and consideration for clients, peers and staff', was the least difficult competency with a measure of -3.32 and standard error of 0.32.

Person ability, person reliability and item reliability. The person reliability for the Consistency scale was 0.90 with person ability ranging from 7.28 to -1.56 (range of 8.84). The item reliability based on the total competency was 0.97.

Discussion

The statistical evaluation of the competencies rated and assessment processes (global VAS ratings and categorical ratings of performance consistency) comprising the Assessment of Foundation Clinical Skills indicated that these components have good content, substantive, and structural validity and generalisability (Messick, 1995) for the assessment of speech-language pathology students within a simulated clinical placement. The development and implementation of appropriate assessments of simulated learning experiences for speech-language pathology students is in its infancy and the assessment content and processes of the AFCS show promise as a strategy for authentic assessment. The description of a pre-Novice level of performance is a new addition to rating student competency in speech-language pathology, and was found to operate as a valid category in the context of this simulated placement. The AFCS allows for formative feedback and accurate assessment of students' clinical skills at early stages of clinical curricula, whilst also sharing the conceptual framework of COMPASS®, offering students valuable continuity in competency assessment

processes. While the pre-Novice behavioural descriptor and other specific features of the AFCS were supported, some components require further investigation before their validity can be assured.

Content Validity

The inclusion of units and elements of competency which were relevant and representative of foundation clinical skills was determined by expert, professional judgement, based on COMPASS[®] units and elements (McAllister, 2006). Fit statistics established whether each competency unit and element in the AFCS contributed to its overall construct of foundation clinical skills in a meaningful way (Bond & Fox, 2007). Fit statistics for the seven AFCS competency units rated on the VAS fell within the guidelines stipulated by Bond and Fox (2007) and therefore, could be considered relevant and representative (see table 4).

Infit measures for the Consistency scale indicated that all competency elements conformed to Bond and Fox's (2007) guidelines and that outfit measures for 30 of the 36 elements also met requirements. Infit measures are considered more important than outfit measures which may "have no practical implications at all" (Bond & Fox, 2007, p. 240). With regard to aberrant outfit scores, one element (GPC unit 3.0 Lifelong Learning element c. 'Adapts behaviour to address learning goals') was below the guidelines and therefore represented overfit, or an element which failed to function independently of other elements (Bond & Fox, 2007). It is possible that this element's rating was influenced by that of the following element ('Actively seeks to extend and integrate learning') and may therefore constitute a duplication. An alternative view, however, is that overlap in element content allows for additional opportunity for clinical educator rating and consequently, more informed judgment of performance in the overall unit of competency.

Fit statistics for the other five elements fell above the stipulated guidelines, thereby representing underfit or competencies which clinical educators rated more variably or

unpredictably and suggest these competencies were difficult to interpret (Bond & Fox, 2007). Attention to such competencies is imperative as “underfit degrades the quality of the ensuing measures” (Bond & Fox, 2007, p. 240). A possible contributor to this inconsistency was clinical educator irregularity in using the Consistency scale (further discussed in relation to structural validity below). A consequence was that the ratings of students’ performances on these competencies did not behave in the same way as other ratings. Interestingly, four of these five elements were in the ‘Professionalism’ unit of competency, the easiest competency for students to achieve in the current study (see below). There is an apparent need for review of the relevance of each element and possible redefinition before further validation to justify their continued inclusion.

Substantive validity

Substantive validity determines whether the tasks appropriately assess a designated skill. Rasch analysis confirmed that the AFCS assessed the unidimensional trait of foundation clinical skills, with fit statistics of each unit of competency falling within the stipulated range (Bond & Fox, 2007) (see table 4). This confirms that the inclusion of both generic and professional competencies in the AFCS was justified. This study did not separately analyse students’ competency levels for each assessment event (middle and end of the simulated placement). Investigation of competency change over the placement would seek to confirm that the simulated clinical placement successfully provided students with a means by which they could develop clinical skills in an incremental and continuous manner (Benner, 2001; McAllister, 2006).

Structural Validity

Structural validity refers to the fidelity of the scoring procedures and structures incorporated into the AFCS, specifically, the inclusion of the pre-Novice level on the VAS and the use of the Consistency scale.

Visual analogue scale (VAS). The VAS was found to adequately represent the range of students' performances. Students were rated by clinical educators after multiple observations of their performance across a range of tasks, ensuring that ratings were based on sufficient evidence (Lurie, 2012; McAllister et al., 2006). Rasch analysis justified the division of the VAS scale into five functional categories of varying sizes (see table 3). This finding of an uneven spread with fewer than ten categories is similar to McAllister's (2006) findings for the way in which rating categories on the VAS in COMPASS[®] are distributed. These findings support the notion that the AFCS VAS categories are meaningful divisions between levels of competence in speech-language pathology practice in a simulated environment.

Clinical educator rating behaviour highlighted a limitation of the current VAS. In situations where educators had not observed the student demonstrate any elements of a specific unit, they either did not score the student or rated them near the lowest end of the pre-Novice level. McAllister (2006) resolved this limitation with the inclusion of a 'not applicable' box at the beginning of the COMPASS[®] VAS. It would be beneficial for future iterations of the AFCS to include such an option to maximise consistent rater behaviour and fair assessment practices.

A predetermined mark at 19mm along the VAS on the AFCS identified the change from pre-Novice level to Novice level, whereas Novice is the first or beginning point for the COMPASS[®] VAS. Category information from Rasch analysis revealed that the interval from 0-18mm represented a discrete category, that is, the lowest level of performance, indicating that clinical educators regarded the pre-Novice level of the VAS as a meaningful stage of competency development when assessing foundation clinical skills. This finding validates the inclusion of the pre-Novice section of the VAS and confirms that some students required a high degree of support to complete the structured tasks in the simulated clinical placement. Their performance at the pre-Novice level at the midway point prompted discussion of

targeted learning goals aimed at achieving Novice level by the end of the placement. Further clinical placements will then provide opportunities for students to implement the skills practised in this simulated clinical placement to manage the variety of practical and professional skills associated with workplace placements (Billett, 2011).

Consistency Categorical Scale. This scale had three pre-assigned categories: ‘rarely’, ‘mostly’, and ‘consistently’. Category information revealed that the three categories did not conform to all aspects of the guidelines outlined by Linacre (2002) and were therefore not functional categories (see table 5). The ‘consistently’ category was over-represented in clinical educator ratings, reflecting that clinical educators rated the majority of students’ performances as consistent, with half as many as ‘mostly’ consistent. The ‘rarely’ category was under-represented.

Visual inspection of ratings indicated that some clinical educators had rated a student at the lower end of a category at the midway assessment (for example, on the left side within the ‘consistently’ box), then at the higher end of the same category at the end assessment (at the right side of the same box). It appeared that educators intended to signify that the student had made improvements in skills over time but had not progressed to the next category. Furthermore, some clinical educators rated a student’s performance on the line between two categories. These rater behaviours implied that the three assigned categories did not allow for adequate representation of current performance and perceived growth in skills, resulting in clinical educators adapting the scale, effectively designating additional categories of performance. This is reported to be a potential difficulty with the use of limited rating points within a scale, with possible consequences of reduced sensitivity and reliability, and more susceptibility to a ‘collapse’ of ratings, with students with different levels of ability being rated at the same level (Smith, Wakely, De Kruif, & Swartz, 2003). Given the limitations of this rating scale, the findings regarding the validity of the elements rated for consistency and

the overall finding of good reliability indices should be treated with caution. Further discussion with clinical educators regarding their intention when using the scale as described above would have clarified rating outcomes. It may be that the overall ‘global’ ratings on the VAS provide a more meaningful representation of student performance and that clinical educators determine the rating on the VAS based on performance that is ‘mostly’ or ‘consistently’ at this level. Further research needs to be undertaken to determine how perceived consistency of performance affects competency rating decisions and, if appropriate, to establish the optimal method for rating the consistency of performance of students on these elements.

Generalisability

Generalisability, measured by item and person reliability scores, refers to the extent to which the assessment scores can be generalised to other groups of examinees and in other contexts (Messick, 1995). This study provides preliminary evidence that predicts that the AFCS has good generalisability. Item reliability scores for the AFCS indicated that the competency units and elements were likely to elicit similar performance levels if used with another group of students with similar levels of ability (Bond & Fox, 2007). Person reliability scores indicated that if this group were given an alternative clinical competency assessment measuring the same foundation clinical skills, it was probable that each student would perform at a similar level as on the AFCS compared with their peers (Bond & Fox, 2007). These high reliability scores are promising given the data was collected across two distinct cohorts of students. Furthermore, as also determined for COMPASS[®] (McAllister, 2006), the wide spread of person abilities represented in the sample indicated that the competency units were appropriate to assess the range of student ability in the current study.

It is important to note that the AFCS was designed for use in a simulated environment for a specific set of learning outcomes and activities and its validation is linked to this

specific context. Therefore it is not directly transferable or generalisable to other types of simulated learning environments or experiences. However, this research does provide preliminary support to the notion that valid assessments of simulated learning can be developed. It is likely that some components of the current study's simulated clinical placement (such as case history interviews and paediatric speech assessment administration) may be present in another foundation clinical skills clinic. In addition, quality teaching and learning practices supported by this assessment, for example, providing formative feedback to students by using the assessment at the midway point and supporting valid assessment with multiple observations of student performance, could be adopted by developers of simulated learning environments.

External Validity

External validity is the extent to which the AFCS ratings are comparable to ratings of similar assessments that measure the same construct (Messick, 1995). An alternative assessment tool does not exist and therefore it was not possible to compare the AFCS with another assessment, thereby prohibiting evaluation of external validity. However, COMPASS[®] (McAllister et al., 2006) is currently used to assess speech-language pathology students' performance in the workplace. While COMPASS[®] and the AFCS differ in their intended assessment use and in some components of their structure, Rasch analysis has confirmed that raters use the tools to rate speech-language pathology practice in a similar manner. This indicates that there is potentially some alignment between learning that occurs in simulated and workplace environments. This is a positive finding which suggests that the simulated clinical placement provides an appropriate transition to clinical tasks undertaken by students within a workplace clinic.

Similarities between the AFCS and COMPASS[®] exist in the number of identified categories, high item and person reliability measures and identification of the least difficult

unit (GPC unit 4, Professionalism). In comparison, the most difficult unit in the COMPASS[®] assessment was identified as CBOS competency unit 2, Analysis and Interpretation, while the AFCS identified CBOS unit 3, Planning of Speech Pathology Intervention as the most difficult. McAllister et al. (2011) suggested that some competencies, such as professionalism, may be more transparent for students and educators and/or more readily practised while on placement, and may therefore be easier to achieve, a trend observed in the current study. In contrast, opportunities to demonstrate others may be more limited and they may also be more difficult to describe, given their complexity (McAllister et al., 2011). In the current study, 90% of clinical educators indicated that planning of speech-language pathology intervention was not observed at the first assessment, potentially signalling a lack of opportunity for students to develop this skill throughout the placement and a subsequent lower rating at the end point.

Consequential validity

Consequential validity considers the outcomes of test use, in particular whether a test may be invalidated through misinterpretation or misuse (Messick, 1995). Some components of the AFCS warrant attention. The Consistency scale was found to be unreliable as a rating measure. In addition, competency elements of the AFCS that were rated for consistency, whilst based on those of COMPASS[®], were created specifically for the simulated clinical placement. The potential for construct underrepresentation (failure to incorporate important components of foundation clinical skills) and construct-irrelevant variance (the inclusion of competencies which are too easy or difficult) is acknowledged with Messick (1995) noting the latter to be possible in assessment in environments which simulate real-world tasks. Further research is indicated to ensure risks to validity are minimised.

Limitations and future research

There are several limitations to this study, and these highlight possible opportunities for future research. Firstly, determination of inter-rater reliability measures would add strength to the findings. Secondly, clinical educator rater behaviour may have contributed to incomplete and/or inaccurate data in some instances. Although clear rating instructions were provided, clinical educators did not consistently provide a well-defined identifiable rating at both assessment points and ratings on the Consistency scale were unpredictable. Further clinical educator training and attention to element content is recommended to minimise this lack of clarity. It would be valuable for future research to validate clinical educators' ratings of student performance on two different occasions to track the development of student competency over time. Finally, this research provides support for the use of the AFCS within this particular simulated clinical placement only. Further research is indicated to determine its validity when used in other clinical placement formats at both foundation and more advanced levels.

Conclusions

This research yielded preliminary data that supports the validity of the AFCS as a tool to assess the foundation clinical skills of speech-language pathology students in a simulated clinical placement. Specifically, designing an assessment linked to the specific learning objectives of the simulated placement and based on ratings of observed student behaviours while carrying out simulated tasks yielded useful assessment information. The addition of a pre-Novice category of performance was found to be meaningful for the simulated clinical placement and may also be of use when rating students whose performance falls below Novice level in workplace placements. Incorporating speech-language pathology competencies found to be relevant for assessment of workplace performance was supported for all units and most elements of competency. Recommendations for amendments to the AFCS include the addition of a 'not applicable' box on the VAS, and review of all elements

to ensure they are representative of foundation clinical skills. Consideration of the broader applicability of the AFCS content and process would strengthen the potential for its use in other clinical contexts.

Acknowledgements

The authors gratefully acknowledge the contribution of Danielle Aldridge in the development of the AFCS.

References

- American Speech-Language-Hearing Association (ASHA) (2009). *2005 Standards and Implementation Procedures for the Certificate of Clinical Competence in Speech-Language Pathology - Revised March 2009*. Retrieved from http://www.asha.org/certification/slp_standards/.
- Barrows, H. S. (1971). *Simulated patients (programmed patients): The development and use of a new technique in medical education*. Springfield, Illinois: Charles C. Thomas.
- Baylor, C., Hula, W., Donovan, N., Doyle, P., Kendall, D., & Yorkston, K. (2011). An introduction to Item Response Theory and Rasch models for speech-language pathologists. *American Journal of Speech-Language Pathology, 20*, 243-259.
- Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing, 27*, 101-118.
- Benner, P. (2001). *From novice to expert: Excellence and power in clinical nursing practice*. Upper Saddle River, N.J.: Prentice Hall.
- Billett, S. (2011). Integrating experiences in workplace and university settings: A conceptual perspective. In S. Billett & A. Henderson (Eds.), *Developing learning professionals, professional and practice-based learning* (Vol. 7, pp. 21-40). Dordrecht, The Netherlands: Springer.
- Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the*

- human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bradley, P. (2006). The history of simulation in medical education and possible future directions. *Medical Education*, 40, 254-262.
- Canadian Association of Speech-Language Pathologists and Audiologists (2012). *CASLPA Clinical Hours Requirements for Certification*. Retrieved from http://www.caslpa.ca/PDF/Certification/SLP_Clinical_Hours_Requirements_Details.pdf.
- Chan, J., Carter, S., & McAllister, L. (1994). Sources of anxiety related to clinical education in undergraduate speech-language pathology students. *Australian Journal of Human Communication Disorders*, 22(1), 57-73.
- Hancock, A., & Brundage, S. (2010). Formative feedback, rubrics, and assessment of professional competency through a speech-language pathology graduate program. *Journal of Allied Health*, 39(2), 110-119.
- Hill, A. E., Davidson, B. J., & Theodoros, D. G. (2010). A review of standardized patients in clinical education: Implications for speech-language pathology programs. *International Journal of Speech-Language Pathology*, 12, 259-270.
- Kramer, C., Copley, J., & Nelson, A. (2004). The effect of a pre-clinical fieldwork subject on the confidence and professional skills of undergraduate occupational therapy students. *Focus on Health Professional Education: A Multi-Disciplinary Journal*, 5(3), 7-22.
- Lane, C., & Rollnick, S. (2007). The use of simulated patients and role-play in communication skills training: A review of the literature to August 2005. *Patient Education and Counseling*, 67, 13-20.
- Le Maistre, C., & Paré, A. (2004). Learning in two communities: The challenge for universities and workplaces. *Journal of Workplace Learning*, 16(1/2), 44-52.

- Lim, S., Rodger, S., & Brown, T. (2009). Using Rasch analysis to establish the construct validity of rehabilitation assessment tools. *International Journal of Therapy and Rehabilitation, 16*(5), 251-260.
- Linacre, J. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*, 85-106.
- Lincoln, M. (2012). The diversity challenge for universities and clinical educators. *Journal of Clinical Practice in Speech-Language Pathology, 14*(1), 2-6.
- Linquist, I., Engardt, M., & Richardson, B. (2004). Early learning experiences valued by physiotherapy students. *Learning in Health and Social Care, 3*(1), 17-25.
- Lurie, S. (2012). History and practice of competency-based assessment. *Medical Education, 46*(1), 49-57.
- Lysaght, R., & Hill, A. (2010). The use of standardized patients as an alternative to fieldwork. In L. McAllister, M. Paterson, J. Higgs & C. Bithell (Eds.), *Innovations in allied health fieldwork education: A critical appraisal* (pp. 189-198). Rotterdam, The Netherlands: Sense Publishers.
- McAllister, L. (1997). An adult learning framework for clinical education. In L. McAllister, M. Lincoln, S. McLeod & D. Maloney (Eds.), *Facilitating learning in clinical settings* (pp. 1-26). Cheltenham, UK: Stanley Thornes.
- McAllister, L. (2005). Issues and innovations in clinical education. *International Journal of Speech-Language Pathology, 7*, 138-148.
- McAllister, S. (2006). *Competency based assessment of speech pathology students' performance in the workplace*. (Unpublished PhD thesis, University of Sydney, Faculty of Health Sciences). Retrieved from <http://hdl.handle.net/2123/1130>.
- McAllister, S., Lincoln, M., Ferguson, A., & McAllister, L. (2006). *COMPASS®: Competency Assessment in Speech Pathology*. Melbourne: The Speech Pathology

Association of Australia.

- McAllister, S., Lincoln, M., Ferguson, A., & McAllister, L. (2010). Issues in developing valid assessments of speech pathology students' performance in the workplace. *International Journal of Language and Communication Disorders, 45*(1), 1-14.
- McAllister, S., Lincoln, M., Ferguson, A., & McAllister, L. (2011). A systematic program of research regarding the assessment of speech-language pathology competencies. *International Journal of Speech-Language Pathology, 13*(6), 469-479.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.
- Norcini, J., & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher, 29*, 855-871.
- O'Kane, G. (2010). Developing pre-clinical skills in nutrition and dietetics students from a NSW rural university in an acute care private hospital. *Focus on Health Professional Education: A Multi-Disciplinary Journal, 11*(2), 60-75.
- Paparella-Pitzel, S., Edmond, S., & DeCaro, C. (2009). The use of standardized patients in physical therapist education programs. *Journal of Physical Therapy Education, 23*(2), 15-23.
- Sheepway, L., Lincoln, M., & Togher, L. (2011). An international study of clinical education practices in speech-language pathology. *International Journal of Speech-Language Pathology, 13*(2), 174-185.
- Sherer, C., Morris, D., Graham, C., & White, L. (2006). Competency-Based Early Clinical Education Experiences for Physical Therapy Students. *The Internet Journal of Allied Health Sciences and Practice, 4*(2). Retrieved from <http://ijahsp.nova.edu>.
- Smith, E. V., Wakely, M. B., De Kruif, R. E. L., & Swartz, C. W. (2003). Optimizing rating

scales for self-efficacy (and other) research. *Educational and Psychological Measurement*, 63(3), 369-391.

Speech Pathology Association of Australia (SPAA). (2001, 2011). *Competency-Based Occupational Standards (CBOS) for Speech Pathologists: Entry-Level (revised)*. Melbourne: The Speech Pathology Association of Australia Ltd.

Theodoros, D., Davidson, B., Hill, A., & MacBean, N (2010). *Integration of simulated learning environments into speech pathology clinical education curricula: A national approach*. Retrieved from <https://www.hwa.gov.au/sites/uploads/sles-in-speech-pathology-curricula-201108.pdf>.

Weddle, M., & Sellheim, D. (2011). Linking the classroom and the clinic: A model of integrated clinical education for first-year physical therapist students. *Journal of Physical Therapy Education*, 25(3), 68-80.

Table 1. Structure of simulated clinical placement (3 hour clinic session per week).

| Week | Session Type | Assessment |
|------|--|--|
| 1 | Clinical workshop: Introduction to clinical skills | Learning objectives of simulated clinical placement and Assessment of Foundation Clinical Skills (AFCS) introduced to students |
| 2 | Standardised patient case history interview: group practice | |
| 3 | Clinical educator role play case history interview: paired practice | |
| 4 | Standardised patient case history interview: paired model | |
| 5 | Clinical workshop: Taking data online / recording observations/ speech sound analysis. | Mid-way evaluations on the AFCS including formative feedback on interviews undertaken in weeks 2-4 |
| 6 | Clinical workshop: Speech test administration, scoring and analysis | |

- 7 Kindergarten visit: Screening assessment of a young child's speech skills
- 8 Clinical workshop: Articulation therapy planning and planning for information giving interview
- 9 Clinical educator role play information giving interview: group practice
- 10 Standardised patient information giving interview: group model
- 11 Clinical workshop: Phonological therapy planning
- 12 Assessments conducted during session

End of placement evaluations on the AFCS including summative assessment of interviews in weeks 9-10, kindergarten assessment in week 7, and performance in clinical workshop tasks in weeks 6, 8, and 11

Table 2. Summary of results of Rasch analysis in relation to Linacre's (2002) guidelines.

| Guideline | Number of categories identified along the visual analogue scale | | | | |
|---|---|-----|-----|-----|----------------------------------|
| | 9 | 7 | 4 | 5 | 5 |
| | | | | | (change in category 2 length) |
| All categories must have at least 10 observations | No | No | Yes | Yes | Yes |
| Observations are regularly distributed | No | No | No | No | Yes |
| Average measures advance monotonically | No | Yes | Yes | Yes | Yes |
| Step calibration must advance | Yes | Yes | Yes | Yes | Yes |
| Outfit mean squares less than 2.0 | No | No | No | Yes | Yes |
| Ratings imply measures and measures imply ratings | Yes | Yes | Yes | Yes | Yes |

| | | | | | |
|---------------------------------------|----|----|-----|-----|-----|
| Logits must advance by at least 1.4 | No | No | Yes | Yes | Yes |
| Logits must advance by no more than 5 | No | No | No | No | Yes |

Table 3. Visual analogue scale category information.

| Category label | Observed numbers* | Observed average | Infit MNSQ** | Outfit MNSQ** | Step calibration | Category measure | Coherence $M \rightarrow C^{\S}$ |
|----------------|-------------------|------------------|--------------|---------------|------------------|------------------|----------------------------------|
| 1 | 190 | -7.47 | 1.09 | 1.07 | None | -7.55 | 84% |
| 2 | 339 | -4.27 | 0.74 | 0.62 | -6.44 | -4.48 | 76% |
| 3 | 485 | 0.04 | 0.85 | 0.86 | -2.53 | -0.11 | 78% |
| 4 | 290 | 3.65 | 1.24 | 1.14 | 2.30 | 4.48 | 68% |
| 5 | 104 | 7.33 | 1.08 | 1.08 | 6.66 | 7.77 | 71% |

*Number of ratings recorded in each category. Recorded numbers include all entered data for 130 students' mid and end scores.

**Mean-square. \S Measure implies Category.

Table 4. Visual analogue scale competency information.

| Item (competency unit)* | Measure (logits) | Measure (standard error) | Infit MNSQ** | Outfit MNSQ** |
|---|------------------|--------------------------|--------------|---------------|
| Planning of Speech- Language Pathology Intervention | 1.33 | 0.18 | 0.89 | 0.82 |
| Assessment | 0.94 | 0.16 | 1.08 | 0.97 |
| Analysis and Interpretation | 0.86 | 0.17 | 0.87 | 0.76 |
| Reasoning | 0.16 | 0.15 | 0.90 | 0.86 |
| Life-long Learning | -0.75 | 0.15 | 1.08 | 1.03 |
| Communication | -0.81 | 0.15 | 0.81 | 0.76 |
| Professionalism | -1.73 | 0.14 | 1.11 | 1.11 |

**Mean-square.

Table 5. Consistency categorical scale category information.

| Category label | Observed number* | Observed average | Infit MNSQ** | Outfit MNSQ** | Structure calibration | Category measure | Coherence M→ C [§] |
|------------------|---------------------|---------------------|-----------------|------------------|--------------------------|---------------------|--------------------------------|
| 1 (rarely) | 208 | -0.32 | 1.08 | 2.01 | None | -2.97 | 48% |
| 2 (mostly) | 2054 | 1.37 | 0.98 | 1.28 | -1.85 | 0.00 | 66% |
| 3 (consistently) | 4221 | 3.90 | 0.94 | 0.95 | 1.85 | 2.97 | 83% |

*Number of ratings recorded in each category. Recorded numbers include all entered data for 130 students' mid and end scores. **Mean-square. § Measure implies Category.