

Measuring Competency Development in Objective Evaluation of Videofluoroscopic Swallowing Studies

Natrah Ahmad Nordin¹ · Anna Miles¹ · Jacqui Allen^{1,2}

Received: 30 March 2016 / Accepted: 22 December 2016 / Published online: 11 January 2017
© Springer Science+Business Media New York 2017

Abstract Clinical interpretation of videofluoroscopic swallow study (VFSS) has often been criticized for its poor objectivity and inter-rater agreement. In order to address this, objective VFSS measures have been developed, reported and demonstrated to be valid and reliable. However, widespread clinical implementation is lacking. Reasons cited include lack of training and excessive time taken to perform measures. This study investigated competency development in selected standardized objective VFSS measures among speech-language pathologists (SLPs) naïve to quantitative measurement. Six novice (no VFSS experience) and four experienced (2–10 years of VFSS experience) SLPs participated in 4 h of training in how to perform selected objective VFSS measures including pharyngeal constriction ratio, maximum pharyngoesophageal opening, pharyngoesophageal opening duration, airway closure duration and total pharyngeal transit time. Each week for eight weeks, participants were asked to independently measure and report three VFSS of patients affected by stroke. By week 8, all SLPs, irrespective of prior experience level, were able to achieve 80% accuracy in measures in comparison to the consensus of three expert clinicians. SLPs' mean time for completion reduced from 50 min in week 1–25 min in week 8. Inter-rater agreement for measures improved across the eight-week period (range ICC = −31.05 to .60 in week 1 to ICC = .71 to .98 in

week 8). There was high agreement in location of impairment and consequent management recommendations by week 8. In conclusion, SLPs can reliably learn and incorporate objective VFSS measures within a reasonable time frame. Level of experience has limited influence on the learning curve.

Keywords Dysphagia · Deglutition · Deglutition disorders · Videofluoroscopic study of swallowing · VFSS · Modified barium swallow · MBS · Competency · Inter-rater reliability

Introduction

Use of videofluoroscopic study of swallowing (VFSS) by speech-language pathologists (SLPs) as an instrumental dysphagia assessment is well known [1]. VFSS is a dynamic, real-time, two-dimensional assessment of swallowing [2]. Despite the widespread clinical use of VFSS, interpretation is often subjective and inter-rater agreement for diagnosis and treatment recommendations is highly variable across clinicians [3–5]. Clinician agreement, or inter-rater reliability, is improved by training clinicians to observe and rate diagnostic variables systematically to a criterion—typically using rating scales [6, 1, 7, 8, 9]. Hind et al. [10] investigated the judgement of trained, expert SLPs when using the criterion-based, Penetration-Aspiration Scale (PAS) [8]. After 20 min of tutoring, aspiration identification agreement was high ($K = .86$ to $.89$). Similarly, in the development of the MBSImpTM, SLPs were trained for 8 h and underwent 4 h of independent study [1]. Through standardized criterion-based scoring and coaching, agreement improved irrespective of years of experience [1]. Intra-rater and inter-rater agreement for MBSImpTM was more than 80% [1].

✉ Anna Miles
a.miles@auckland.ac.nz

¹ Speech Science, School of Psychology, The University of Auckland, Auckland, New Zealand

² Otorhinolaryngology, Waitemata District Health Board, Auckland, New Zealand

Although these studies address the issue of clinician agreement, lack of objectivity remains a challenge. Over the last fifteen years, Kendall and colleagues have published a standardized approach using objective, digital timing and displacement measures of key swallowing events including a substantial adult normative database [11]. [Appendix 1](#) provides a selection of their published measures and definitions. They have systematically defined these measures and demonstrated correlation with clinical risk such as total pharyngeal transit time and pharyngeal constriction ratio [11]. Reliability of these measures has been established across eight to ten raters in numerous studies involving normal and dysphagic patients, typically resulting in substantial agreement (intraclass correlation coefficients (ICC) .90 to .98) [11]. Clinician agreement of these objective VFSS measures was supported by a recent pilot study [12]. The authors compared VFSS interpretation agreement of SLPs pre- and post-training using objective VFSS measures. Twenty-one SLPs with a wide range of experience (6 months to 16 years) participated in the study. The study was divided into three assessment scenarios: (1) real-time view in radiology suite, (2) frame-by-frame and, (3) objective VFSS measures. Agreement across clinicians for pharyngeal phase measures increased from low agreement in scenario 1 ($K = .19$) to moderate agreement in scenario 3 ($K = .60$) following the introduction and familiarization with the objective VFSS measures [12].

Increased objectivity and agreement in VFSS interpretation has potential for clinical benefits as accuracy in rehabilitation management should also improve [13]. To encourage uptake of such objective measures, it would be helpful to advise clinicians of the expected learning curve and time requirement to become proficient with the measures. The overall aim of this study was to evaluate competency development and feasibility of SLP use of a systematic (or quantitative) VFSS measurement process in order to improve the objectivity of study interpretation and better inform care. Competency was defined as the SLPs' ability to effectively apply new knowledge and skills in learning objective VFSS measures then translating them into reporting. We measured this by comparing accuracy of measures and interpretations, inter- and intra-rater agreement. Feasibility was defined as time required to measure and report and an arbitrary judgement as to whether this would fit into a clinical workload. A second aim was to assess participating clinicians' self-reported feelings of competence and pressure when learning and using this measurement system. The following research questions were asked: Can SLPs master a selected number of objective VFSS measures within an eight-week period? Do speed, accuracy and interpretation skills improve over time? Does previous clinical experience in VFSS influence competency development? How do SLPs' perceived

competence and perceived pressure in completing the measures change over time?

Methods

This study received appropriate regional ethics approval (UAHPEC 013106) and all participants provided written informed consent.

Participants

An advertisement was published on a University social media website. Participants were selected on first-come, first-served basis. Six novice SLPs with no experience in conducting VFSS and four experienced SLPs with experience in leading VFSS were recruited. Five novice SLPs were new graduates with no previous clinical experience. One novice had three years experience working clinically without exposure to VFSS. Experienced SLPs' years of clinical experience and proficiency leading VFSS ranged from 2 to 10 years. No participants had familiarity with objective VFSS measures ([Table 1](#)). Participants were required to attend a training session in person but then completed the study from their home locality (between 5 and 1000 km from the research site).

Objective Measure Training

Participants underwent 4 h of hands-on objective VFSS measures training focused on learning five measures described in detail by Leonard and Kendall: pharyngeal constriction ratio (PCR), maximum pharyngoesophageal segment opening (PESmax), pharyngoesophageal segment opening duration (POD), airway closure duration (ACD) and total pharyngeal transit time (TPT) ([Appendix 1](#)) [11]. Participants had the opportunity to practice each measure as well as being provided with the clinical implications of their use. Participants viewed each study frame-by-frame (Quicktime Media Player, Apple Inc) and were trained to calculate displacement measures using Universal Desktop Ruler (AVPSoft Inc). Alongside the training process, they were also given weekly mentoring via email and access to the key textbook: "Dysphagia Assessment and Treatment Planning: A Team Approach" by Leonard and Kendall [11]. Mentoring included email responses to queries and brief weekly feedback regarding measures and interpretation accuracy.

Study Protocol

Twenty consecutive de-identified VFSS recordings of individuals with dysphagia due to stroke were taken from

Table 1 Demographic information

Status	Age	Gender	Years of practice	Years of VFSS practice	Clinical area of work	Competency reached (80% accuracy of measures)	Competency reached (speed less than 30 min)
Novice	24	Female	0	0	None	Week 5	Week 6
Novice	24	Female	0	0	None	Week 1	Week 7
Novice	25	Female	0	0	None	Week 1	Week 7
Novice	24	Female	0	0	None	Week 7	Week 5
Novice	23	Female	0	0	None	Week 8	Week 6
Experienced	55	Female	9	9	Urban	Week 4	Week 4
Experienced	26	Female	2	2	Urban	Week 2	Week 2
Experienced	30	Female	10	10	Urban and rural	Week 1	Week 1
Experienced	27	Female	2	2	Urban	Week 3	Week 8
Novice	27	Female	3	0	Urban	Week 5	Week 5

Accuracy in comparison to gold standard consensus of 3 expert clinicians

The University of California, Davis database. All VFSS were recorded at 30 frames per second with timing information superimposed on the fluoroscopic recording in 100ths of a second using a Horita VS-50 Video Stopwatch (Horita, CA, USA). A radio-opaque ring of known diameter was taped to the patient's chin to allow calibration for displacement measures. Selected videos ranged from 39 s to 2 min 25 s in length. In order to avoid introducing bias from recorded audio instructions, all videos were de-voiced.

Participants were sent three randomly selected videos per week. For each video, SLPs were instructed to watch the full VFSS protocol but to complete timing and displacement measures on the one 20 mL fluid swallow. Videos were sent each Monday and participants were instructed to complete the objective measures independently in their own time. Three videos per week were selected to represent a realistic caseload for a clinician running a weekly VFSS clinic. To perform analyses of measurement agreement, six randomly chosen videos were re-coded and repeated a second time during the 8-week period. A standardized interpretation sheet [penetration-aspiration scale score (PAS), five objective measures, diagnostic impression and recommendations] was used for all videos (Appendix 2). After completing the measures, participants were expected to summarize their diagnostic findings and make recommendations based solely on the VFSS measures. They were also asked to time how long it took them to complete each video. Finally, participants were asked to rate their perception of competence and pressure each week using sub-sections of the Intrinsic Motivation Inventory (IMI): Task Evaluation Questionnaire [14]. Participants were required to rate 10 statements using a seven-point Likert scale (1 = not at all true, 4 = somewhat true, 7 = very true). For example, 'I did not feel at all nervous about doing the VFSS objective

measures' and 'I felt pretty skilled at VFSS objective measures'. They were required to email their measures, interpretations, timings and questionnaires to the researchers at the end of every week. In order to encourage participation by working clinicians and limit participation time, SLPs were asked to participate for up to eight weeks or until they reached the stipulated gold standard level of speed and accuracy. Three expert clinicians completed measures on all twenty videos independently prior to the research. Inter-rater agreement between expert clinicians was ICC .92 (confidence interval (CI) .87 to 1.0). Due to their high level of inter-rater agreement, the three sets of expert ratings were used as the gold standard against which study participants were compared. For experts, the time taken to complete one video was on average 20 min (range 18–24). Gold standard was set at 80% accuracy in comparison to the consensus of the three expert clinicians [5] within a 30-min completion time.

Data Analysis

Data were inputted into Excel (Microsoft Corporation, Santa Rosa, CA, USA) by the primary researcher blinded to participant number or week number. Statistical analyses were completed using SPSS version 22 (SPSS, IL, USA). Speed, measures, interpretations and perceived competence and pressure ratings were compared across weeks and experience levels. In order to explore agreement in diagnostics and management, two elements of the VFSS were selected for analysis. Diagnostic impression was coded for identification of pharyngeal constriction and pharyngoesophageal segment impairments to reflect the two objective pharyngeal displacement measures (PCR and PESmax) taken by participants. The management recommendations were also coded to identify exercises specifically known to focus on pharyngeal constriction (masako, effortful

swallow) and pharyngoesophageal segment opening (Shaker head lift, Mendelsohn).

Inter-rater and intra-rater reliability was conducted using intraclass correlation coefficient (ICC) to examine agreement among participants over time. The ICC values were interpreted using Kappa criteria of Landis and Koch [15] as they are considered comparable [16]: .01 indicates poor agreement, .01 to .21 slight agreement, .21 to .40 fair agreement, .41 to .60 moderate agreement, .61 to .80 substantial agreement and .81 to 1.00 almost perfect agreement [15]. Speed, measurements and interpretations were compared to the gold standard consensus to calculate percentage agreement. Participant percentage accuracy was defined as within one standard deviation from the experts—gold standard. Relationships between speed, accuracy, perceived competence and pressure were explored using linear regression. Friedman test was carried out to examine the differences in Task Evaluation Questionnaire scores over weeks for perceived competence and pressure domains.

Results

All ten participants completed the entire study. No drop out occurred.

Demographic Data

Participant characteristics are presented in Table 1. One experienced SLP stopped measures at week 3 (participant 8) and another at week 5 (participant 6) having reached gold standard. The six novice SLPs and the two other experienced SLPs chose to complete all eight weeks irrespective of reaching gold standard on earlier weeks (Table 1).

Speed

Across participants, the mean completion time for one video was 50 min in week 1 (range 28.3–62.3, SD 15.0) and reduced to 25 min by week 8 (range 15.0–33.3, SD 6.65, $p < .001$) (Fig. 1). Experienced SLPs reduced their time for completion earlier than novice SLPs. However, there was no significant difference in overall duration between groups at initial ($p = .84$) or final week ($p = .99$).

Measure Agreement

Inter-rater agreement improved across weeks for PCR, PESmax and POD (Table 2). However, ACD agreement decreased from moderate agreement (ICC = .51) to fair agreement (ICC = .22). In comparison, PAS score achieved almost perfect agreement for every week (week 1

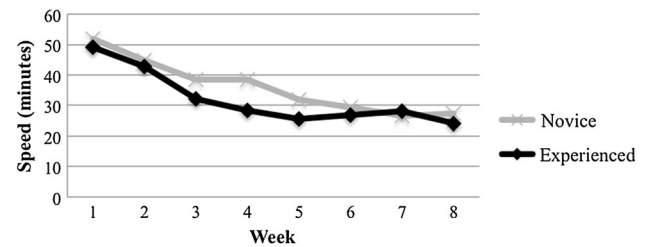


Fig. 1 Mean speed to complete measures across weeks

ICC = .91, 95% CI .58 to 1.0; Week 8 ICC = .95, 95% CI .80 to 1.0). Across participants, intra-rater agreement also increased over time (week 1 to week 3 = ICC range .09 to .986; week 3 to week 7 = ICC range .959 to .994) with no difference between experienced and inexperienced SLPs.

Accuracy of Measures

Accuracy was similar across experience groups (Fig. 2). At week 3, the accuracy was more than 80% for all measures with exception of PCR, where 80% agreement was achieved at week 5 (Table 3). ACD remained variable across weeks. As accuracy of the measurement improved, time for completion reduced ($R^2 = .688$, $p < .05$).

Perceived Competence and Pressure

There was a statistically significant increase in perceived competence among the participants across weeks ($\chi^2 (7) = 50.54$, $p < .001$) and a statistically significant decrease in pressure ($\chi^2 (7) = 37.75$, $p < .001$) (Figs. 3, 4). Experienced SLPs had higher perceived competence between week 1 and week 5. However, by week 6, there was no difference between experience levels. As accuracy in measurement increased, perceived competence increased ($R^2 = .575$, $p = .12$) and perceived pressure decreased ($R^2 = .498$, $p = .20$). As time to complete reduced, perceived competence increased ($R^2 = .957$, $p = 3.94$) and perceived pressure decreased ($R^2 = .9134$, $p < .001$).

Diagnosis and Rehabilitation Recommendations

Mean percentage agreement of diagnosis and management recommendations are displayed in Table 4. By week 8, mean percentage agreement for diagnosis and management recommendations ranged from 66.67 to 100%.

Discussion

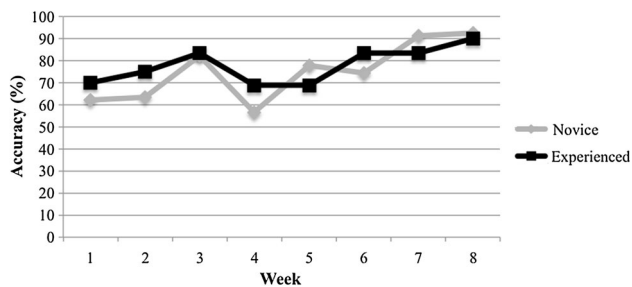
This research programme explored competency development for learning and undertaking objective VFSS measures in both novice and experienced SLPs. Comparisons

Table 2 Inter-rater agreement for measures across time

Measure	Week 1			Week 3			Week 8		
	ICC	LCI	UCI	ICC	LCI	UCI	ICC	LCI	UCI
PCR	.535	−.407	.987	.951	.786	.999	.864	.478	.996
PESmax	.597	−.177	.988	.992	.968	1.000	.962	.831	.999
POD	−.31.05	.0	−.177	.947	.782	.999	.977	.896	.999
ACD	.511	−1.718	.988	.526	−1.159	.988	.222	−1.988	.979
TPT	.323	−.981	.977	.663	−.192	.991	.708	−.316	.993

Bold values are purely distinguishes ICC from confidence intervals

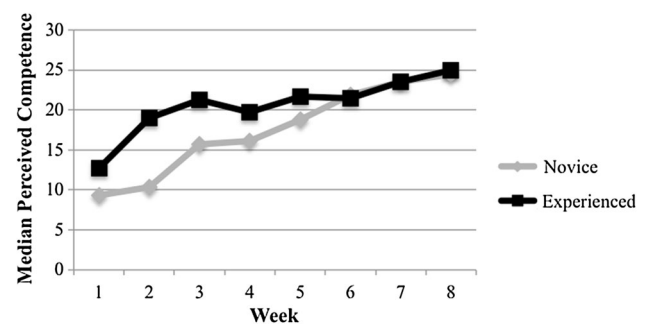
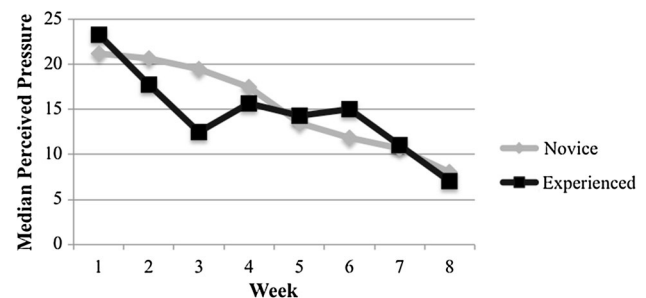
ICC intraclass correlation coefficient, LCI lower confidence interval, UCI upper confidence interval

**Fig. 2** Mean percentage accuracy of measures across weeks

were based on recorded pharyngeal timing and displacement measures of 20 mL fluid swallows. Speed and accuracy are often considered the simplest measures for evaluating skill training [17]. With training, and eight-weeks of practice, both novice and experienced SLPs achieved competency as judged by time to complete, accuracy and inter-rater agreement.

Speed

In order for objective VFSS measures to be feasible in clinical practice, measures need to be achievable within the time constraints of a busy caseload. This requires both the

**Fig. 3** Median perceived competence across weeks**Fig. 4** Median perceived pressure across weeks**Table 3** Mean percentage accuracy of measures across participants

Week	Mean percentage accuracy for measures					
	PCR	PESmax	POD	ACD	TPT	Overall
1	60	73.33	86.67	56.67	50	65.33
2	73.33	66.67	63.33	80	60	68.66
3	66.67	100	83.33	83.33	80	82.66
4	62.96	83.33	74.07	48.19	62.96	84.17
5	88.89	83.33	70.37	74.07	70.37	84.17
6	83.33	83.33	91.67	54.17	75	76.66
7	83.33	80	79.17	79.17	79.17	83.33
8	91.67	82.66	87.5	79.17	70.83	82.5

Bold 80% and above

Table 4 Mean percentage agreement in diagnosis and management recommendations across weeks

	Percentage (%) agreement					
	W1		W3		W8	
	Mean	Range	Mean	Range	Mean	Range
Diagnosis						
PES						
Novice	83.33	33.33	83.33	33.33	83.33	33.33
Experienced	77.78	33.33	77.78	33.33	66.67	50
Pharynx ^b						
Novice	88.89	33.33	83.33	33.33	88.89	33.33
Experienced	88.89	33.33	88.89	33.33	100	–
Management recommendations						
PES ^a						
Novice	83.33	33.33	66.67	–	72.22	50
Experienced	78.45	33.33	77.78	33.33	83.33	50
Pharynx						
Novice	94.44	16.67	88.89	16.67	77.78	33.33
Experienced	88.89	33.33	88.89	33.33	83.33	50

Bold 80% and above

^a PES interventions included Shaker or Mendelsohn exercise

^b Pharyngeal strength included effortful swallow or Masako exercise

speed of performing the measures, undertaking analysis and reporting, as well as duration of the learning curve for individual SLPs to reach reasonable levels of accuracy and to feel competent with measures, to be realistic. Within six weeks, new graduate SLPs had achieved the same speed of completion measuring five measures of swallowing as experienced SLPs even though experienced clinicians increased their speed of analysis more quickly. This suggests that exposure to VFSS clinically does influence the learning curve of SLPs incorporating objective measures for the first time. Comprehensive knowledge in anatomic structures is essential for interpreting VFSS. Experience is well established as advantageous for learning [18, 19] with experience enhancing ones' ability to reflect on action while practicing [20, 21]. Faster recall is also achieved with increased experience [19, 22, 23], and it is reasonable to expect experienced SLPs to make quicker decisions regarding interpretation of findings and recommendations compared with new graduates. It must be noted that although participants reported on the full VFSS including penetration/aspiration and oral parameters, the focus of this study was accuracy of five specific pharyngeal timing and displacement measures. Participants had little clinical information about each patient. Speed of completion may be greater if further clinical decision-making is needed or if additional measures of swallowing kinematics were to be measured. This would need to be factored into the clinical setting.

Measure Agreement

This research provides preliminary information regarding SLP agreement in objective VFSS measures following 4 h of training. The overall agreement of measures increased over time. All measures achieved substantial to almost perfect agreement within 8 weeks, except for ACD. Clinician agreement of objective VFSS measures yielded higher inter-rater agreement than that found by many previous researchers using the more traditional observational approach [6, 7]. Similar to previous study, looking at group discussion and training [24, 25], the findings of this study show promising results using an off-site mentoring technique following initial face-to-face training.

Poor agreement in airway closure duration throughout the study period sits in contrast to the other measures. Poor screen resolution to detect subtle grey-scale changes is a possible explanation as participants often verbalized difficulty visualizing the laryngeal vestibule during this measure [25, 26]. Standardizing VFSS resolution and optimizing screen viewing by control of ambient light, adequate warm-up of the viewing screen and screen quality may improve measure accuracy.

Accuracy

Previous clinical experience with VFSS did not appear to influence SLPs' ability to achieve accuracy in objective

VFSS measurements. In fact, two novice SLPs achieved 80% accuracy in measures in week 1 following only four hours training. These data support the findings by Logemann and colleagues [27] that SLPs can achieve specific clinical skills with training irrespective of experience.

Diagnosis and Recommendations

Arguably one of the most important clinical questions is whether high accuracy and clinician agreement lead to high agreement in diagnostic impressions and treatment decisions. Although this was not the aim of this study, preliminary analyses suggest that participants did agree in diagnosis and treatment options for both impairments of pharyngeal strength and pharyngoesophageal segment opening. These positive results are in accordance with those obtained by Miles and colleagues [12], where agreements in interpretation and management decisions were high when objective VFSS measures were used for study interpretation. Interestingly, there was no difference between novice and experienced SLPs' agreement of diagnostic impression and management decision, suggesting that objectivity refutes the reliance on experience in clinical decision-making. Evidence-based, objective measures perhaps provide confidence to novice clinicians in developing treatment recommendations. The researchers planned to code participants' recommendations for oral intake as: normal, modified, nil-by-mouth. However, heterogeneity of responses precluded any statistical analysis of these responses, which provided only qualitative information. Participants' responses were rarely decisive (for example, "*nil-by-mouth but tastes of thick fluid and puree as tolerated*", "*thick fluids or perhaps thin fluids with a straw might work*"). This perhaps reflects the difficulty participants had in management decisions without adequate clinical information. A further more standardized study has been designed to address recommendations related to quantitative measures.

Perceived Competence and Perceived Pressure

As participants progressed through the eight-week study, perceived competence increased, while perceived pressure decreased with the majority of change occurring within the first three weeks. In the adult education literature, trainees often start with feeling the pressure of being externally assessed, fear of revealing low competence or differing from peers [18]. As trainees receive positive feedback from a mentor, they perform better [28, 29]. In fact, there is evidence that stress may be a catalyst for greater performance [29]. Again, the value of off-site (email or telehealth based) mentoring is promising.

Limitations

To the authors' knowledge, this is the first study to compare the development of competency between novice and experienced SLPs utilizing objective VFSS measures. Thus, there are no gold standard parameters that can be used to define competency attainment. The gold standard of 80% accuracy was based on previous benchmarking work [5] and 30-min completion time was solely based on what the researchers deemed to be clinically feasible within a hospital SLPs' workload. This research may act as the baseline to identify the exact competency standard for SLPs who are learning objective VFSS measures in the future.

One limitation of the data was that with only three videos measured per week, one difficult case could skew the data considerably. The researchers ensured that high-quality videos were sent to participants; however, the participant SLPs' personal computers might have inadequate processing power or screen resolution [30].

The limited number of participants may have influenced reliability measures and led to wide confidence intervals. Caution must be applied, especially as numbers of experienced SLPs decreased over the weeks, as participants reached gold standard. This can lead to normality violation and negatively affect ICC results [31, 32]. A fixed sample size could improve the power of future studies [33]. An alternative approach to strengthen the results would be to reduce variability through stricter data collection sheets with tick box responses rather than open-ended responses. Open-ended responses were chosen in this study to avoid priming of participants. However, criterion-based diagnostic impressions and rehabilitation recommendations could be added to the objective VFSS measures to reduce divergence.

Additionally, without adequate patient history information, diagnostic impression and recommendations were based solely on interpretation of the objective VFSS measures in this study. In reality, judgments influencing treatment recommendations are made based on all aspects of patient history, examination and co-related conditions. Furthermore, other measures of swallowing would likely be utilized in clinical practice including bolus flow and timing measures and these may have altered treatment recommendations in this study, had they been included.

Conclusions

SLPs can learn and incorporate objective VFSS measures within a feasible time frame. Speed to completion, inter-rater agreement and accuracy improved over an eight-week

time period irrespective of prior VFSS experience levels. As perceived competence with measures increased over time, pressure of the new tool receded. Most importantly, these data demonstrate good agreement among SLPs on VFSS interpretation when utilising objective VFSS measures. With the use of objective, reproducible measures and strong agreement in diagnostic interpretation and recommendations, more specific standardized care can be recommended. This should, in turn, lead to better ability to share data in a standardized fashion and hopefully better patient outcomes. This study can inform SLP departments when developing training programmes to facilitate the implementation of objective VFSS measures in future.

Acknowledgements Thank you to all the speech-language pathologists who participated in the research. The authors would like to acknowledge Dr. Rebecca Leonard and thank her for providing guidance and the videos for this study.

Compliance with Ethical Standards

Conflicts of interest The authors have nothing to declare. This work has not been published elsewhere. All authors have contributed to the study and are accountable for accuracy and integrity. The authors agree to the copyright conditions of this journal.

Appendix 1

See Table 5.

Table 5 VFSS objective measures(Adapted from [11])

Measure	Definition
Timing (s)	
Total pharyngeal transit time (TPT)	Represents the total time of the bolus passage through the pharynx, from when the bolus head passes the posterior nasal spine (B1), to the time at which the bolus tail completely clears the PES (BP2) <i>Total pharyngeal transit time = BP2–BP1</i>
Airway closure duration (ACD)	The duration of total airway closure, from the approximation of the elevated arytenoids with the down folding epiglottis (AEc), to the first frame in which the epiglottis has returned to its pre-swallow position (Em) <i>Airway closure time = Em–AEc</i>
PES opening duration (PESdur)	The duration of PES opening from the first frame in which it opens (Pop), to when it closes behind the bolus tail (Pcl) <i>PES opening time = Pcl–Pop</i>
Displacement measures (cm)	
Maximum pharyngeal area	Measured when the pharynx is at rest, either prior to or following a swallow. The pharyngeal area is outlined by the posterior pharyngeal wall extending from the mid-portion of the tubercle of the atlas to the top of the arytenoid cartilages, anteriorly over the arytenoids to outline the epiglottis, valleculae and tongue base up to the soft palate [34]
Maximum pharyngeal constriction	The same pharyngeal area as outlined in the maximum pharyngeal area was measured again, but at the point of maximum constriction during a swallow
Pharyngeal constriction ratio (PCR)	The ratio of pharyngeal area at maximum constriction to the area of the pharynx at rest <i>PCR = Maximum pharyngeal area divided by maximum pharyngeal constriction</i>
PES max opening (PESmax)	The width of the pharyngoesophageal segment was measured at the point of maximum opening during the swallow

Appendix 2

See Table 6.

Table 6 Standardized interpretation sheet

Time started:

Diagnostic Summary (fill in the blanks)

20ml Fluid Bolus Displacement Measures			Norms (2SD)				WNL?
			Male <65	Male >65	Female <65	Female >65	
Pharyngeal constriction ratio (PCR)	Pharynx Area - Hold (cm ²)						
	Area - max constrict						
	max constrict/hold =		.09 (.16)	.09 (.16)	.09 (.16)	.09 (.16)	
Max Hyoid Displacement			2.40 (1.36)	2.48 (2.04)	1.81 (1.46)	2.07 (.45)	
Max PES Opening (PESmax)			.90 (.55)	.80 (.40)	.90 (.55)	.80 (.40)	
20ml Fluid Bolus Duration Measures			<65yrs	>65yrs			
PES Open Duration	Pcl						
	Pop						
	Pcl - Pop =		.50 (.21)	.64 (.21)			
Airway Closed Duration	Aes						
	Em						
	AE _{close}						
	Em - AE _{close} =		.72 (.50)	.85 (.93)			
Total Pharyngeal Transit	B1						
	BP1						
	BP2						
	BP2 - B1 =		.87 (.50)	1.19 (.62)			
20ml Fluid Bolus Coordination of Timing Measures							
Aes before Pop?							
AE _{close} at least 0.1 sec prior to BP1?							
Pop before or same time as BP1?							
BP2 before or same time as Pcl?							
Maximum Penetration/ Aspiration Scale Score (1-8):							

Pati

ent presents with _____ dysphagia. Characterized by (physiological impairments) _____ Resulting in (symptoms – residue, aspiration) _____

Recommendations for oral intake

Recommendations for SLP management:

Time finished:

References

- Martin-Harris B, Brodsky MB, Michel Y, Castell DO, Schleicher M, Sandidge J, et al. MBS measurement tool for swallow impairment—MBSImp: establishing a standard. *Dysphagia*. 2008;23(4):392–405. doi:10.1007/s00455-008-9185-9.
- Bingjie L, Tong Z, Xinting S, Jianmin X, Guijun J. Quantitative videofluoroscopic analysis of penetration-aspiration in post-stroke patients. *Neurol India*. 2010;58(1):42–7. doi:10.4103/0028-3886.60395.
- Dyer JC, Leslie P, Drinnan MJ. Objective computer-based assessment of valleculae residue—is it useful? *Dysphagia*. 2008;23(1):7–15. doi:10.1007/s00455-007-9088-1.
- Huang Y, Tseng F, Tsai S, Lin C, Chou Y, Lew HL. The use of the gaussian curve fitting method for scintigraphic measurements of the swallowing process in healthy subjects: implications for evaluation of dysphagia. *Arch Phys Med Rehabil*. 2006;87(11):1526–9. doi:10.1016/j.apmr.2006.08.324.
- Wilcox F, Liss JM, Siegel GM. Interjudge agreement in videofluoroscopic studies of swallowing. *J Speech Language Hear Res*. 1996;39(1):144–52. doi:10.1044/jshr.3901.144.
- Kuhlemeier K, Yates P, Palmer J. Intra- and interrater variation in the evaluation of videofluorographic swallowing studies. *Dysphagia*. 1998;13(3):142–7. doi:10.1007/PL00009564.
- McCullough GH, Wertz RT, Rosenbek JC, Mills RH, Webb WG, Ross KB. Inter- and intra-judge reliability for videofluoroscopic swallowing evaluation measures. *Dysphagia*. 2001;16(2):110–8. doi:10.1007/PL00021291.
- Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL. A penetration-aspiration scale. *Dysphagia*. 1996;11(2):93–8. doi:10.1007/BF00417897.
- Stoeckli SJ, Huisman TA, Seifert BA, Martin-Harris BJ. Interrater reliability of videofluoroscopic swallow evaluation. *Dysphagia*. 2003;18(1):53–7. doi:10.1007/s00455-002-0085-0.
- Hind JA, Gensler G, Brandt DK, Gardner PJM, Blumenthal L, Gramigna GD, et al. Comparison of trained clinician ratings with expert ratings of aspiration on videofluoroscopic images from a randomized clinical trial. *Dysphagia*. 2009;24(2):211–7.
- Leonard R, Kendall K. *Dysphagia assessment and treatment planning: a team approach*. 3rd ed. San Diego: Plural Publishing; 2014.
- Miles A, Ting A, Leonard R, Allen J. Applying objectivity to VFS interpretation: objective measures of swallowing physiology. *Dysphagia*. 29. Platform presentation at Dysphagia Research Society Meeting and Post-Graduate Course, March 5–8, 2014, Nashville; 2014. p. 743.
- Leonard RJ, Kendall KA, McKenzie S, Gonçalves MI, Walker A. Structural displacements in normal swallowing: a videofluoroscopic study. *Dysphagia*. 2000;15(3):146–52. doi:10.1007/s004550010017.
- Ryan RM, Mims V, Koestner R. Relation of reward contingency and interpersonal context to intrinsic motivation: a review and test using cognitive evaluation theory. *J Pers Soc Psychol*. 1983;45:736–50.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74. doi:10.2307/2529310.
- Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Measur*. 1973;33(3):613–9. doi:10.1177/001316447303300309.
- Tsuda S, Scott D, Doyle J, Jones DB. Surgical skills training and simulation. *Curr Probl Surg*. 2009;46(4):271–370.
- Daley BJ. Novice to expert: an exploration of how professionals learn. *Adult Educ Q*. 1999;49(4):133–47. doi:10.1177/074171369904900401.
- Dreyfus SE, Dreyfus HL (1980) A five-stage model of the mental activities involved in directed skill acquisition (No. ORC-80-2). California Univ Berkeley Operations Research Center. <http://www.dtic.mil/dtic/tr/fulltext/u2/a084551.pdf>. Accessed 1 Jan 2016.
- Eraut M. *Developing professional knowledge and competence*. London, Washington D.C.: The Falmer Press; 1994.
- Wainwright SF, Shepard KF, Harman LB, Stephens J. Novice and experienced physical therapist clinicians: a comparison of how reflection is used to inform the clinical decision-making process. *Phys Ther*. 2010;90(1):75–88. doi:10.2522/ptj.2009077.
- Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *Am J Surg*. 2003;185(2):146–9. doi:10.1016/S0002-9610(02)01213-8.
- Unsworth CA. The clinical reasoning of novice and expert occupational therapists. *Scand J Occup Ther*. 2001;8(4):163–73. doi:10.1080/110381201317166522.
- Palmer A, Graville D. Improving the reliability of interpretation of the modified barium swallow study. Paper presented at the Sixteenth Annual Dysphagia Research Society Meeting, Isle of Palms; 2008.
- Scott A, Perry A, Bench J. A study of interrater reliability when using videofluoroscopy as an assessment of swallowing. *Dysphagia*. 1998;13(4):223–7. doi:10.1007/PL00009576.
- Burns CL, Keir B, Ward EC, Hill AJ, Farrell A, Phillips N, et al. A dynamic image quality evaluation of videofluoroscopy images: considerations for telepractice applications. *Dysphagia*. 2015;30(4):473–81. doi:10.1007/s00455-015-9626-1.
- Logemann JA, Lazarus CL, Keeley SP, Sanchez A, Rademaker AW. Effectiveness of four hours of education in interpretation of radiographic studies. *Dysphagia*. 2000;15(4):180–3. doi:10.1007/s004550000025.
- Deci EL, Vallerand RJ, Pelletier LG, Ryan RM. Motivation and education: the self-determination perspective. *Educ Psychol*. 1991;26(3–4):325–46. doi:10.1080/00461520.1991.9653137.
- Struthers CW, Perry RP, Menec VH. An examination of the relationship among academic stress, coping, motivation, and performance in college. *Res High Educ*. 2000;41(5):581–92. doi:10.1023/A:1007094931292.
- Peladeau-Pigeon M, Steele CM. Technical aspects of a videofluoroscopic swallowing study. *Can J Speech Lang Pathol Audiol*. 2013;37:216–26.
- Eng J. Sample size estimation: how many individuals should be studied? 1. *Radiology*. 2003;227(2):309–13. doi:10.1148/radiol.2272012051.
- Ionan AC, Polley MC, McShane LM, Dobbin KK. Comparison of confidence interval methods for an intra-class correlation coefficient (ICC). *BMC Med Res Methodol*. 2014;14(1):121. doi:10.1186/1471-2288-14-121.
- Zou G. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Stat Med*. 2012;31(29):3972–81. doi:10.1002/sim.5466.
- Yip H, Leonard R, Belafsky PC. Can a fluoroscopic estimation of pharyngeal constriction predict aspiration? *Otolaryngol Head Neck Surg*. 2006;135(2):215–7.

Natrah Ahmad Nordin BSc

Anna Miles PhD

Jacqui Allen MB, FRACS