Sonographer Skill Teaching Practices Survey: Development and initial validation of a survey instrument

Delwyn Nicholls, BAppSc(Rad), Grad Dip(Med US), DMU(vasc), Grad Dip Hed, Grad Cert Hed[1], **Linda Sweet**, BNg, MNg, Grad Cert Ed(Higher Ed), PhD[1], **Pawel Skuza**, BSc(Math), MEd(EducResEvalAssess)[2], **Amanda Muller**, BA (Hons), PhD, Grad Cert Education (H.Ed)[1] and **Jon Hyett**, MBBS BSc MD MRCOG FRANZCOG[3,4]

[1]Faculty of Medicine, Nursing & Health Sciences, Flinders University, Adelaide, South Australia, Australia
[2]eResearch, Flinders University, Adelaide, South Australia, Australia
[3]RPA Women and Babies, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia
[4]Discipline of Obstetrics, Gynaecology and Neonatology, Faculty of Medicine, University of Sydney, Sydney, New South Wales, Australia

Abstract

Objective: To report on the initial content development of the sonographer skill teaching practices survey (SonoSTePs) tool to measure skill teaching perceptions.

Method: The nine-step model by Sarantakos [*Social Research*, 3rd edn. London: Palgrave Macmillan; 2005. 1–464], was used to guide the initial development of the instrument, content domains, and the generated questions. The face and content validity of the instrument was established after progressing through two expert panel reviews (experts in the fields of education, statistics, and ultrasound) and two pilot tests.

Results: Results of the pilot testing identified differing discriminant variability (variation ratios) between the use of a five-point and a seven-point Likert-type frequency scale. In particular, the pilot two (P2) 30-item instrument, using a frequency seven-point Likert-type scale, resulted in good variation ratios (0.68 across 24 items; SD=0.11), demonstrating discriminatory ability between individual participant's teaching practices.

Conclusions: The content and face validity of the SonoSTePs instrument has been established using principles of questionnaire construction and development.

Keywords: clinical skills, psychomotor skills, survey instrument, teaching, ultrasound, validation.

Introduction

An ultrasound scan is a complex medical examination and sonographers need a diverse range of skills to both perform the scan and document their findings competently. The foundational skills required to physically perform and visually interpret the scan are termed psychomotor skills.[1] These can be sub-divided into the categories of visuo-motor and visuo-spatial skills that enable the sonographer to manipulate the ultrasound transducer for image acquisition while mentally constructing a three-dimensional spatial orientation of the anatomical structures.[1]

Correspondence to email dnicholls@sufw.com.au
doi: 10.1002/ajum.12011

Psychomotor skills are unique to each profession and are first taught then learnt. As a result, a number of psychomotor skill teaching models have been proposed for multidisciplinary health professionals in various clinical situations.[2–8] Universally, these models teach psychomotor skills through a stepped approach (whether the skill is simple or complex). An important aspect of psychomotor skill instruction is that the teaching practices are based upon principles aligned to the motor learning domain.[6,9–11] There are no published data that define best teaching practices or describes how sonographers' teach psychomotor skills. To our knowledge, there is no published and validated instrument to measure sonographer skill teaching practices. Therefore, it was identified that there is a need for a valid and reliable survey instrument which would provide

information about sonographer skill teaching practices. As a result, a sonographer skill teaching practices survey (SonoSTePS) was developed using published principles of questionnaire development and construction.[12–16]

This paper reports on the development and initial validation of SonoSTePS. The aims of the paper are to: (i) describe the conceptual framework and development process of the web-based SonoSTePS survey instrument, (ii) report on the iterative process used to validate the content of the survey instrument, and (iii) describe salient and relevant technical difficulties which arose during the validation process.

Materials and methods

Development of the SonoSTePS Instrument

The SonoSTePS survey instrument was developed following general guidelines for scale development established by De Vellis[17] and in particular using the nine-step approach to guide survey construction and design by Sarantakos[14] (Table 1). This model was selected to guide the questionnaire design and construction as it was systematic, logical, and able to provide safeguards to minimise major instrument design errors.

We note Sarantakos uses the words pre-test or pilot test interchangeably.[14] For the purposes of our research and survey design, we define a pilot test as an activity to test the sample frame, questions, research methods, and instruments.[14,18]

The initial step to developing a survey instrument involved determining the concept being researched, followed by identifying the central themes, called content domains, to be measured. It was initially difficult to identify the content domains relevant to teaching scanning skills due to a dearth of empirical data and published literature on the research topic. A review of literature from the psychomotor learning field and health disciplines that use psychomotor skill teaching models for clinical tasks (surgical medicine,[11,16,19–22] colonoscopy,[3] dentistry,[23] nursing[24,25], and physiotherapy [8]) identified five domains relevant to teaching psychomotor (scanning) skills in medical ultrasound. These five domains include: teach a new skill, visual exemplar, cognitive overload, immediate error correction, and skill practice. Table 2 lists the five content domains (using scales), a brief description of each domain, the literature which assisted with defining the domain, and the type of survey questions used.

The next step to developing the survey instrument entailed generating a pool of questions to explore and examine key aspects of each content domain. There are no fixed guidelines to the number of questions (items) required to represent each content domain in a survey, although, as a guideline, there should be enough questions to adequately represent the key dimensions.[17] The majority of items were derived from literature through a process of identifying the theoretical and learning principles applicable to motor skill teaching (the supporting literature is listed in Table 2). Two standalone questions were also written to elaborate and explore specific student sonography skill teaching preferences identified from one paper by Sonaggera.[43] For example, when teaching novice sonographers, scanning skill participants were asked 'When teaching a beginning student a new skill, do you scan the patient first and then follow with the student scanning after you?' The survey instrument also required a mixture of ranked questions using rating or frequency scales, closed-end and open–end questions to gather both qualitative and quantitative data.[44,45] In particular, open-end questions provided the opportunity to gather additional insights which may have been excluded by using only closed questions. Pilot one (P1) survey items were produced after culling redundant, poorly worded, and confusing questions from an initially large bank of questions.[17,46] Pilot one survey was comprised of a total of 27 items. Rating scale questions were used for three of the items. Questions 13, 14 and 16 contained 10 or less questions in each rating scale. The instrument was assembled and formatted into three key sections: (i) demographics comprised of 13 questions, (ii) psychomotor skill teaching practices and skill feedback which contained three rating scale questions and five closed/open text questions, and (iii) validation feedback which included five questions.

Table 1: Steps in questionnaire construction, design and analysis[14] [pp. 254–5].

1. Preparation-select survey type and method of administration. Search for developed survey or commence construction
2. Construct the first draft – construct survey questions around research themes
3. Self-critique
4. External scrutiny – ensure the survey is reviewed by industry experts and modify questions as required
5. Re-examine and review – where major changes are required proceed to previous step and ensure instrument undergoes external scrutiny again
6. Pilot test or pre-test – a small sample is selected to check the suitability of the survey as a whole
7. Revision-changes may ensue from pilot test. Go back to step 4. Small changes can move to step 8
8. Second pilot test
9. Formulation of the final draft

Table 2: Teaching scanning skills in clinical practice: scales and items.

Dimension	Domain/Scale	Scale description
Teaching a clinical skill	Teach new skill	The extent to which skill tutors execute skill teaching elements described by George and Dotto[2,5,6,9] *Sub scale*: Recognition of prior learning. The extent to which tutor establishes learners prior cognitive and psychomotor knowledge on skill topic[8,26–28] *Sub scale*: Simulation The extent to which tutor uses simulated patient or phantoms to teach part or whole task clinical scanning skills
	Cognitive overload	The extent to which tutors limit the quantity of information taught in any one teaching session [29–31] The extent to which tutor performs task analysis (deconstruction) prior to teaching the skill [5,16,28,29,32–35] The extent to which the tutor provides concurrent feedback during skill practice [36–39]
	Visual exemplar	The extent to which a tutor performs a silent skill demonstration to provide a visual standard of performance of skill execution[3,19,24,26,28,34,40]
	Immediate skill error correction	The extent to which tutor corrects incorrectly performed skills as they occur[5]
	Skill practise	The extent to which the tutor provides deliberate and supported practise opportunities in short skill sessions (<60 min), rather than one long session, to practice skills with feedback on performance[21,25,27,41,42]

Demographic data were sought to ascertain if skill teaching approaches were influenced by individual sonographer's professional practice, educational level, and type of educational qualification achieved. For example, in question eight we asked 'What is the highest level of qualification in ultrasound you have completed?' as we were seeking to establish the participants ultrasound qualification and response options ranged from 'On the job training with Grandfather credentialing to' to 'PhD' and 'prefer not to answer' to provide those PhD-credentialed sonographers with the option for anonymity. This was an important, as some states and territories have one sonographer in each imaging speciality with PhD credentialing. In question nine we asked 'What is the highest level of qualification in clinical health education you have completed?' Response options ranged from none to PhD. Question 10 explored whether sonographers had completed day or half day workshops to assist their teaching roles. This question was necessary because a course such as 'train the trainer' is not recognised as a qualification, yet it is a valuable course to undertake when performing a teaching role. The question asked was 'Have you completed extra training in clinical health education, such as completing 'train the trainer' course or workshops/courses conducted at national conference?' with a response option yes or no, and if yes please specify. The mix of questions in pilot one were wide reaching to garner professional practice and credentialing information to explore if educational level impacted professional skill teaching practices and behaviours.

Validation feedback

Validation feedback was twice sought from both an external expert review panel (informed consent was sought from the review panel to publish their name and salutation see Appendix 1)[47] and targeted sonographers who completed pilot questionnaires.[46,47] This vital process[46] facilitated the critical analyses of the instrument content, format, and domains throughout the developmental period. Qualitative feedback was sought from all scrutineers on the survey questions and clarity, the representativeness of the questions in relation to the research question, the survey format, and the participant information sheet.[17] Data were also collected on the time to complete the survey and any user interface or technical difficulties encountered. This information informed subsequent iterations of the instrument content and design.

Recruitment and sampling

Sonographer clinical tutors, academics, and health educators were initially invited to participate in the pilot 1 (P1) and pilot 2 (P2) pilot testing. Two types of sampling were used. The first involved identifying target participants from university web sites (purposive sampling[44,45]). The second involved contacting participants via email and then inviting them to forward the email invitation and hyperlink to other sonographer tutors or health educators who performed an academic or instruction role (snowball sampling).[44] Initially, nine emails were sent to participants in each cohort (as per well-established recommendations[12,18,48]) and follow-up email invitations were distributed to each professional cohort.

Questionnaire administration

A web-based electronic questionnaire was chosen as the method of administration. The SonoSTePS survey instrument was distributed via an email link to an online version in SurveyMonkey. There are well-known limitations of online data collections[49,50] but the benefits included national sonographer access, cost effectiveness, user-friendliness, and these outweighed the risk of poor response rates to online surveys.[18,45,48,51,52]

Ethics

Ethical approval (SBREC 5584) from the Flinders University Social and Behavioural Research Ethics Committee was obtained prior to study commencement.

Statistical analysis

All results were downloaded from http://www.surveymonkey.-com/ website onto an Excel spread sheet and then imported into SPSS (Statistical Package for Social Sciences, version 21.0.; IBM Corp., Armonk, NY, USA). Limited quantitative data analysis was performed due to small sample sizes in both pilot studies. The qualitative and quantitative data were analysed for descriptive and comparative data. Responses to open-ended questions were evaluated using content analysis.[53] This allowed the exploration of the feedback on the content, dimensions of enquiry, and usability of the instrument. Variation ratios were calculated for P1 and P2. According to Weisberg,[54] the ratio provides a measure of dispersion of participant responses across a scale for a given question. Ratio values can range for 0–1.[54] A ratio of 1 or close to 1 is desirable and indicates there is a broad range of responses across all categories for the question. Conversely, a ratio which approaches 0 indicates the scale was incapable of discriminating participant responses.

Results

Once developed, each version (P1 or P2) of the survey instrument, proceeded through, expert review consisting of four reviewers[46,47] and pilot test.[14] This rigorous process was applied and undertaken to establish content and face validity of the instrument. Between each review and pilot test the questions, question order, Likert-rating scale, and content was modified based on data garnered from feedback from the expert review panel and pilot studies. The next sections will discuss the first and second pilots in more detail.

Pilot one (P1)

Over a 6-week period, eight survey responses were received after the initial email, and a further seven responses were received following a reminder email. One response was an empty entry. No clinical health educators participated in the P1 pilot, despite being invited.

The P1 demographic data revealed that 50% of all respondents were 50–59 years old, predominantly female (71%), with a large proportion employed as university lecturers (46%). Two participants had a PhD qualification in ultrasound and a further six had a Masters qualification. Half of all participants had completed a formal qualification in clinical health education.

The Pilot One (P1) survey used a 5-point scale to measure participants' attitudes to the research question in a format similar to a Likert scale.[13,14,55] In Figure 1, the stacked bar chart frequency distribution for one 9 item rating scale question, exploring sonographer teaching practice behaviours, illustrates the concentration of responses across two rating scales. The

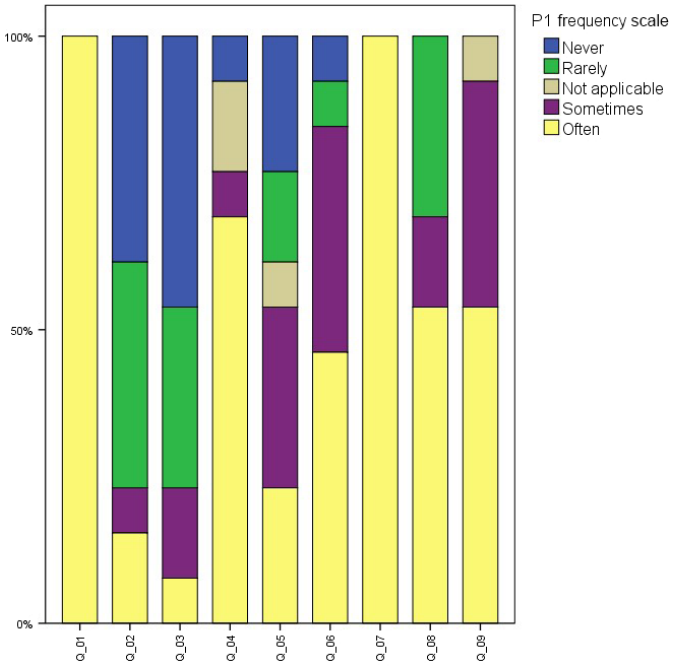

Figure 1: Stacked bar chart with responses to the questions related to teaching new skills from P1 pilot study which used Likert 5-point rating scale (n = 14). Graph purposely does not report the names of the items, thus focusing readers attention to the concentration of responses for often or always.

distribution of the responses to the 5-point scale, of attitudinal questions assessing sonographers' skills teaching and feedback practice, indicated a considerable clustering of responses for most of the questions (18 out of 24). For all these items, only two response categories accounted for over 80% of all pilot data available. Furthermore, four questions had a single category selected by over 90% participants. The average of variation ratio across all 24 items contained in the three rating scale questions was 0.38 (SD = 0.21), and this result indicates a limited variability and discrimination capability for P1 items. The 5-point scale was therefore modified to a 7-point scale as recommended by Vagias.[56] The subsequent Pilot 2 (P2) used a scale ranging from 1 (never) to 7 (always) for data collection.[57]

The qualitative survey feedback received at the P1 stage of the validation process focussed on survey flow and length, question clarity, and administration of the online survey tool. A descriptive content analysis of the qualitative feedback identified three categories. These were broadly grouped into user interface, technical issues with online survey, and survey content. Regarding user interface, one respondent replied 'Would be good to have a completion bar % across top of survey so you know how far to go' and 'radio buttons instead of yes/no written responses'. Another respondent stated 'I found the survey easy to navigate'. Respondents replied with contrasting feedback regarding sufficient room for open-ended questions. One respondent replied 'sufficient room' while four replied 'could

do with more room' and '. . .more space might be useful'. Two respondents gave feedback on the survey content. One respondent replied 'there is no assessment of skills. Maybe something could be included around the assessment/expectations of skill development for students' and another respondent suggested including content on simulated learning: '. . .It might have been appropriate initially to syphon off the lecturers into an extra feedback area for simulation teaching with some appropriate questions'.

As a result of the P1 feedback, modifications were made to the explanatory letter to participants invited to participate in P2 validation process. Participants were advised that the survey was not exploring assessment of skill or competence. An optional three questions were included on the use and role of simulation to teach scanning skills and this was added as a sub dimension to 'teach new skill'. Also, the expert panel identified the need for a definition of simulation and examples of simulated learning aids in order for participants to understand and answer the optional questions about the use of simulation to teach scanning skills. Both Schaeffer *et al.*[15] and Sarantakos[44] highlight the need to define all technical terms to minimise poor or non-response bias when constructing questions for a survey prudent. The P2 survey was modified to incorporate these three questions and expert panel suggestions and these extended the instrument to 30 items. The rating scale items were reduced from 25 in P1 to 24 in P2 after the removal of a question exploring whether simulated aides were used to teach psychomotor skills. Four questions on this topic area would have been an excessive number.

The median time to complete the survey was 20 min, with a range of 10–75 min. An outlying value of 75 min was recorded as a result of encountering technical difficulties to complete the survey. Furthermore, another two participants reported technical errors which were corrected and did not impact the completion time.

Pilot two (P2)

Over an 8-week period, 14 survey responses were received after the initial email and a further five responses were received following a reminder email. Nineteen sonographers participated in the P2 validation of the survey. No clinical health educators participated, despite being invited.

Analysis of the demographic information showed that almost half of participants were over 50 years old, with 84% of them being females. The most represented group, in regard to professional role, were clinical sonographers (37%). The majority of the cohort (61%) was employed as general sonographers in public hospitals. One-third had completed an additional clinical health qualification.

The 5-point scale used in P1 survey was adjusted to a seven-point frequency scale (never-always) rating scale in P2. In order to acquire more meaningful data regarding sonographer teaching practices and behaviours frequency adverbs[58][p. 255] were accompanied by a frequency per cent range (never: 0–2% of time, rarely: 3–19% of time). These strategies we hypothesised would overcome both the described limitations using the 5-point scale, and the potential ambiguity of using word responses which according to Dilmann *et al.*[59] means something different to each participant. The average variation ratio across the 24 items contained in the three rating scale questions was 0.68 (SD = 0.11) and this indicates that the discrimination capability of items from the P2 scale has improved. In Figure 2, the P2 frequency distribution for the same 9-item rating scale question, using a 7-point scale illustrates a dispersion of responses across all rating scales. The content analysis of the qualitative survey feedback received in the P2 validation process identified two categories. These are broadly grouped into question clarity and technical issues with online survey. Two respondents had difficulty interpreting one of the questions containing more than one variable. Kumar[45][p. 154] explains that an ambiguous question is 'one that contains more than one meaning and that can be interpreted differently by different respondents'. This question has since been reviewed and rewritten. All respondents identified that there was enough room to complete the open text questions.

The time to complete the survey ranged from 10 to 30 min with the median value being 15 min. Similar to the P1 pilot, two participants experienced technical difficulties, which were dealt with promptly.

Discussion

The aim of our research was to undertake initial development and validation of a survey instrument which would be capable

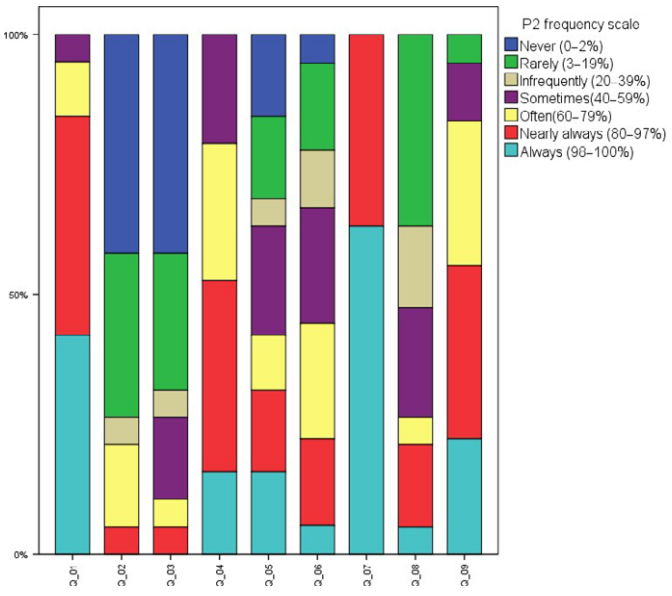

Figure 2: Stacked bar chart with responses to the questions related to teaching new skills from P2 pilot study which used Likert 7-point rating scale (n = 19). Items correspond to the ones in the Figure 1.

of identifying and measuring sonographer skill teaching practices. The survey instrument development and validation model published by Sarantakos[44] provided a framework with which to guide construction and survey content. Applying these steps resulted in the instrument proceeding through two pilot tests and expert review. The results of both pilot tests allowed the development of a measurement instrument, labelled Sonographer Skill Teaching Practices Survey (SSTPS) and subsequently named SonoSTePS. There are a few main points worth mentioning in this discussion, as will be seen in the comments in the following sections on demographics, expert panel review, refining the survey content, and Likert rating vs. frequency scale.

Demographics

The demographic profile of the pilot cohorts completing the survey indicated their adequate representativeness of the broader profession in Australia, which is female dominant. Currently, the female–male ratio of practicing sonographers is 3:1. The Australian Sonography Accreditation Registry (ASAR) reported the in 2012 there were 3380 (76%) females sonographers and 1080 (24%) male sonographers.[60] Our study data demonstrated similar female/male percentages, P1 (71% and 19%) and P2 (84% and 16%). Furthermore, of the academic sonographer cohort, approximately 50% of the P1 and 67% of the P2 cohort had completed additional qualification in clinical health education; therefore, we hypothesise this cohort had the expertise to review the survey content.[13][p. 214] It is of note that only sonographers completed P1 and P2 surveys, although we had invited nine clinical health education academics with niche educational knowledge and expertise to review the instrument. This void was filled by the expertise of the expert review panel and we suggest was not detrimental to the development of the survey instrument.

Expert panel review

Both Lynn[46] and Wetzel[47] suggest at least three panel members are required to critically review the instrument content. We selected four panel members to review each pilot study and their comments lead to question restructuring, rewording, brevity, and placement within the survey. One expert panel member identified the necessity to include a definition of simulated learning in the P2 survey, while another identified the use a seven-point of frequency scale would assist with response discrimination.

Refining the survey content

The validation process involved the evaluation and analysis of the five content domains and the appropriateness of the rating or frequency type-Likert scale in P1 and P2. As a result of the iterative process to validate the content domains of the survey, the survey was revised. A further subgroup of 'teach new skill' was modified to incorporate a section on the use of simulation to teach scanning skills. The inclusion of these questions in the P2 survey meant the stand-alone rating scale question on simulated learning, was now redundant and removal resulted in 24 items. Anecdotally, simulation is widely used in medical ultrasound imaging in Australia to teach foundational scanning skills. However, at the time of the survey instrument's development (2012), a paucity of profession-specific literature made the exploration of the sub-theme difficult. The use of current and representative content domains is a crucial step in the development of a validated survey instrument. Ensuring the instrument authentically and wholly represents the concept being explored and measured is an important step towards establishing content validity. Content validity is reported by Lynn in seminal literature as 'the determination of the content representativeness or content relevance of the elements/items of an instrument. . .'[46][p. 382]. Polit and Beck additionally highlight the need for an instrument to have 'an appropriate sample of items for the construct being measured'. Australian Sonography Accreditation Registry[61][p. 489]. Both Lynn,[46] and Polit and Beck[61] concur that a research instrument must be assessed for content validity prior to use. P1 and P2 were an attempt to achieve this aim.

Likert rating vs. frequency scale

Dilmann et al.[59] explain that a 5-point scale using a Likert design is one method to measure participants' attitudes, opinions and behaviours to a research question. The P1 survey used a 5-point scale ranging from 1 (strongly agree) to 5 (strongly disagree), or 1 (often) to 5 (never). Participants were able to select not applicable (N/A). This option was located between rarely and often, and had the potential for ambiguous interpretation by respondents. This was a design error and corrected in the P2 pilot.

The responses to the 7-point frequency scale ranging from 1 (never) to 7 (always) for questions related to general skill teaching practices were plotted into a stacked bar chart (see Figure 2). The frequency responses of skill teaching and feedback practices exhibit a broader distribution across all response categories when compared to the questions used in the 5-point scale in P1. The average of the variation ratios across all 24 P2 items was 0.68 (SD = 0.11). While it is acknowledged that a direct comparison between the average variation ratios from P2 and P1 cannot be undertaken (given that the wording of some items has changed and pilot samples are small and differ in regard to some of the demographic characteristics), it can be argued that items from P2 are more variable.

The decision to use a 7-point frequency scale with unequal anchor points was based upon two factors. First, the 5-point scale used in the P1 was deemed incapable of satisfactorily discriminating frequency responses across categories. Second, the results of P2 pilot analyses, as well as the literature, support the use of a 7-point frequency scale as having the ability to discriminate more efficiently.[62] This use of a quantified frequency

scale was an attempt to glean nuanced perceptions of skill teaching practices. Indeed, it seems that this aim was achieved. Unlike the 5-point scale in P1, the participants' responses to perceptions of skill teaching practices were dispersed across the seven rating scales in the P2 7-point frequency scale (see Figure 2).[62]

Technical difficulties were encountered while using the web-based survey instrument. The errors were corrected through email correspondence quickly after the initial survey was dispersed.[48] Undertaking a small pre-test of the instrument prior to commencing pilot testing would have revealed these programming errors[44,45][p. 266]. Schleyer & Forrest[18] [p. 419] identify that one purpose of a small pre-test is to test the user interface, usability of the instrument, and detect programming errors prior to distribution of the instrument.

Limitations

Due to the nature of pilot studies, the small sample sizes of P1 and P2 limited the possibility of more formal statistical assessment of changes to the items variability between P1 and P2, although the demographic characteristics of both P1 and P2 samples were relatively comparable. As a result, the utility of the survey instrument will need further testing and refinement on a larger population.

Summary

The P2 survey instrument evolved from a 27-item to a 30-item questionnaire. Between P1 and P2, the survey content and Likert scales were changed, and this improved the dispersion and distribution of responses to probing teaching practice questions. Further research is required to perform basic exploratory psychometric statistical analysis of the measurement instrument using a sample number of at least 300 participants.[17] These processes are critical to the development of a robust instrument which is able to withstand critical review of instrument content, item clarity, and relevance.[17] To avoid the program and access errors encountered with web-based surveys, we suggest performing a survey instrument pre-test prior to dispersal, to mitigate user interface errors.[18]

Acknowledgements

We thank members of the expert review panel: Marilyn Baird PhD; Linda Sweet PhD; Sue Campbell-Westerway PhD and Ann Quinton PhD for their assistance with the external review of the survey instrument.

Disclosure

The authors have declared no conflicts of interest.

References

1 Nicholls D, Sweet L, Hyett J. Psychomotor skills in medical ultrasound imaging: an analysis of the core skill set. *J Ultrasound Med* 2014; 33: 1349–52.
2 Fitts PM. Factors in complex skill training. In: GR. ed. Training, research and education. Pittsburgh: University of Pittsburgh Press; 1962. 177–97.
3 Raman M, Donnon T. Procedural skills education – colonoscopy as a model. *The Canadian Journal of Gastroenterology* 2008; 22(9): 767–70.
4 Poole J. Application of motor learning principles in occupational therapy. *Am J Occup Ther* 1991; 45(6): 531–7.
5 George J, Doto F. A simple five-step method for teaching technical skills. *Fam Med* 2001; 33: 577–8.
6 Simpson E. The classification of educational objectives, psychomotor domain. Urbana, IL: University of Illinois; 1966. 45 pp.
7 Walker M, Peyton JWR. Teaching in theatre. Teaching and learning in medical practice. Manticore Europe: Rickmansworth; 1998.
8 Rose M, Best D. Transforming practice through clinical education, professional supervision and mentoring. Edinburgh: Elsevier-Churchill Livingstone; 2005.
9 Fitts PM, Posner MI. Human performance. Belmont, CA: Brooks/Cole; 1967.
10 Wise SP, Willingham DT. Motor skill learning. In: Squire LR, editor. Encyclopedia of neuroscience, Vol. 3. Oxford: Academic Press; 2009. 1057–66.
11 Walker M., Peyton R. Teaching in theatre. In: Peyton R, editor. Teaching and learning in medical practice. Rickmansworth, UK: Manticore Europe Limited; 1998. 171–80.
12 Johnson B, Christensen L. Educational research – quantitative, qualitative, and mixed approaches, 4th edn. Los Angeles, CA: Sage; 2012. 1–621.
13 Creswell J. Educational research. Planning, conducting, and evaluating quantitative and qualitative research, 3rd edn. Upper Saddle River, NJ: Pearson Prentice Hall; 2008.
14 Sarantakos S. Social Research. Third ed. London: Palgrave Macmillan; 2005. 1–464.
15 Schaeffer NC, Dykema J. Questions for surveys: current trends and future directions. *Public Opin Quart* 2011; 75(5): 909–61.
16 Hamdorf JM, Hall JC. Acquiring surgical skills. *Br J Surg* 2000; 87: 28–37.
17 De Vellis RF. Scale development theory and applications, 3rd edn. Los Angeles, CA: Sage; 2012. 205 p.
18 Schleyer TK, Forrest JL. Methods for design and administration of Web-based survey. *J Am Med Inform Assoc* 2012; 7(4): 416–25.
19 Kovacs G. Procedural skills in medicine: linking theory to practice. *J Emerg Med* 1997; 15(3): 387–91.
20 Lake FR, Hamdorf JM. Teaching on the run tips 5: teaching a skill. *Med J Australia* 2004; 181(6): 327–8.
21 Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative 'bench station' examination. *The American Journal of Surgery* 1997; 173(3): 226–30.
22 Grantcharov T, Reznick R. Teaching rounds: teaching procedural skills. *Br Med J* 2008; 336: 1129–31.
23 Virdi MS, Sood M. Effectiveness of a five-step method for teaching clinical skills to students in a dental college in India. *J Dent Educ* 2011; 75(11): 1502–6.
24 DeYoung S. Teaching strategies for nurse educators. Upper Saddle River, NJ: Prentice Hall; 2003.

25 DeBourgh GA. Psychomotor skills acquisition of novice learners. A case for contextual learning. *Nurse Educator* 2011; 36(4): 144–9.
26 Dent JA, Harden RM. A practical guide for medical teachers, 3rd edn. Edingurgh: Churchill Livingstone; 2009. 435 pp.
27 Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993; 165(3): 358–61.
28 Castanelli DJ. The rise in simulation in technical skills teaching and the implications for training novices in anaesthesia. *Anaesth Intensive Care* 2009; 37(6): 903–10.
29 Phipps D, *et al.* Human factors in anaesthetic practice: insights from a task analysis. *Br J Anaesth* 2008; 100(3): 333–43.
30 Van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. *Med Educ* 2010; 44(1): 85–93.
31 Young JQ, *et al.* Cognitive load theory: implications for medical education: Amee guide no. 86. *Med Teach* 2014; 36(5): 371.
32 Razavi SM, Karbakhsh M, Panah Khahi M, Dabiran S, Asefi S, Shahrak GH, *et al.* Station-based deconstructed training model for teaching procedural skills to medical students: a quasi-experimental study. *Advances in Medical Education and Practice* 2010; 1(1): 17–23.
33 Sullivan ME, *et al.* The use of cognitive task analysis to improve the learning of percutaneous tracheostomy placement. *Am J Surg* 2007; 193(1): 96–9.
34 Lammers RLMD, *et al.* Teaching and assessing procedural skills using simulation: metrics and methodology. *Acad Emerg Med* 2008; 15(11): 1079.
35 Jabbour N, *et al.* Psychomotor skills training in pediatric airway endoscopy simulation. *Otolaryngology – Head and Neck Surgery* 2011; 145(1): 43–50.
36 Kantak SS, Winstein CJ. Learning–performance distinction and memory processes for motor skills: a focused review and perspective. *Behav Brain Res* 2012; 228(1): 219–31.
37 Walsh CM, *et al.* Concurrent versus terminal feedback: it may be better to wait. *Acad Med* 2009; 84(10 Suppl.): S54–7.
38 Winstein CJ, Pohl PS, Lewthwaite R. Effects of physical guidance and knowledge of results on motor learning: support for the guidance hypothesis. *Res Q Exerc Sport* 1994;65(4): 316–23.
39 Winstein CJ. Knowledge of results and motor learning – implications for physical therapy. *Phys Ther* 1991; 71(2): 140–9.
40 Bjork IT. Changing conceptions of practical skill and skill acquisition in nursing education. *Nursing Inquiry* 1997; 4(3): 184–95.
41 Pendleton D, Schofield T, Tate P. The consultation: an approach to learning and teaching. A method for giving feedback. Oxford: Oxford University Press; 1984.
42 Ericsson KA, Krampe RT, Tesch-Romer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993; 100: 364–406.
43 Sonaggera T. Sonography clinical education. *Journal of Diagnostic Medical Sonography* 2004; 20(5): 356–60.
44 Sarantakos S. Social research, 4th edn. New York, NY: Palgrave Macmillan; 2013. 509 pp.
45 Kumar R. Research methodology a step-by-step guide for beginners, 3rd ed. London:Sage Publications; 2011.1–415.
46 Lynn MR. Determination and quantification of content validity. *Nurs Res* 1986; 35(6): 382–6.
47 Wetzel A. Factor analysis methods and validity evidence: a review of instrument development across the medical education continuum. *Acad Med* 2013; 7(8): 1060–9.
48 Raffi F, Shaw RW, Amer SA. National survey of the current management of endometriomas in women undergoing assisted reproductive treatment. *Hum Reprod* 2012; 1–8.
49 Buchanan EA, Hvizdak EE. Online survey tools: ethical and methodological concerns of human research ethics committees. *Journal of Empirical Research on Human Research Ethics* 2009; 4(2): 37–48.
50 Evans JR, Mathur A. The value of online surveys. *Internet Research* 2005; 15(2): 195–219.
51 Zheng K, *et al.* Development and validation of a survey instrument for assessing prescribers' perception of computerized drug–drug interaction alerts. *J Am Med Inform Assoc* 2011; 18(Suppl 1): i51–61.
52 Braithwaite D, *et al.* Using the Internet to conduct surveys of health professionals: a valid alternative? *Fam Pract* 2003; 20(5): 545–51.
53 Saldana J. The coding manual for qualitative researchers. London: SAGE Publications; 2009.
54 Weisberg HF. Variation ratio. In: Lewis-Beck M, Bryman A, Futing Liao T, editors. Encyclopedia of social science research methods. Thousand Oaks, CA: SAGE Publications; 2004. 1178.
55 Kaye BK, Johnson TJ. Research methodology: taming the cyber frontier: techniques for improving online surveys. *Social Science Computer Review* 1999; 17(3): 323–37.
56 Vagias WM. Likert-type scale response anchors. Clemson, SC: Clemson University: Clemson International Institute for Tourism and Research Development, Department of Parks, Recreation and Tourism Management; 2006.
57 Beckstead JW. On measurements and their quality. Paper 4: verbal anchors and the number of response options in rating scales. *Int J Nurs Stud* 2014; 51(5): 807–14.
58 Leitz P. Research into question design. A summary of the literature. *International Journal of Market Research* 2009; 52(2): 249–72.
59 Dillman D, Smyth J, Christian L. Internet, phone, mail, and mixed-mode surveys. The tailored design mode, 4th edn. Hoboken, NJ: John Wiley & Sons, Inc.; 2014; 509 pp.
60 Australian Sonography Accreditation Registry. ASAR Accredited Qualified Sonographers – Gender 2008-2012, 2014 (accessed 8/9/2014).
61 Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health* 2006; 29(5): 489–97.
62 Streiner DL, Norman GR. Health measurement scales – a practical guide to their development and use, 4th edn. New York, NY: Oxford University Press; 2008. 431 pp.

Appendix 1 Survey review panellists

Panel member	Affiliation	Pilot 1	Pilot 2
A/Prof Linda Sweet	Flinders University Education	Yes	Yes
Prof Marilyn Baird	Monash University Education	Yes	No
Associate Professor Sue Campbell-Westerway	Charles Sturt University Ultrasound	Yes	Yes
Name withheld	Sydney University Statistics	Yes	No
Dr Ann Quinton	Sydney University Ultrasound	No	Yes
Pawel Skuza	Flinders University Statistics	No	Yes