

Reliability of the ECHOWS Tool for Assessment of Patient Interviewing Skills

Jill S. Boissonnault, Kerrie Evans, Neil Tuttle, Scott J. Hetzel, William G. Boissonnault

Background. History taking is an important component of patient/client management. Assessment of student history-taking competency can be achieved via a standardized tool. The ECHOWS tool has been shown to be valid with modest intrarater reliability in a previous study but did not demonstrate sufficient power to definitively prove its stability.

Objective. The purposes of this study were: (1) to assess the reliability of the ECHOWS tool for student assessment of patient interviewing skills and (2) to determine whether the tool discerns between novice and experienced skill levels.

Design. A reliability and construct validity assessment was conducted.

Methods. Three faculty members from the United States and Australia scored videotaped histories from standardized patients taken by students and experienced clinicians from each of these countries. The tapes were scored twice, 3 to 6 weeks apart. Reliability was assessed using interclass correlation coefficients (ICCs) and repeated measures. Analysis of variance models assessed the ability of the tool to discern between novice and experienced skill levels.

Results. The ECHOWS tool showed excellent intrarater reliability (ICC [3,1]=.74-.89) and good interrater reliability (ICC [2,1]=.55) as a whole. The summary of performance (S) section showed poor interrater reliability (ICC [2,1]=.27). There was no statistical difference in performance on the tool between novice and experienced clinicians.

Limitations. A possible ceiling effect may occur when standardized patients are not coached to provide complex and obtuse responses to interviewer questions. Variation in familiarity with the ECHOWS tool and in use of the online training may have influenced scoring of the S section.

Conclusion. The ECHOWS tool demonstrates excellent intrarater reliability and moderate interrater reliability. Sufficient training with the tool prior to student assessment is recommended. The S section must evolve in order to provide a more discerning measure of interviewing skills.

J.S. Boissonnault, PT, PhD, WCS, Department of Physical Therapy and Health Care Sciences, The George Washington University School of Medicine and Health Sciences, 2000 Pennsylvania Ave NW, Suite 227, Washington, DC 20006 (USA). Dr Boissonnault was affiliated with the Department of Orthopedics and Rehabilitation, Physical Therapy Program, University of Wisconsin–Madison, Madison, Wisconsin, at the time of the study. Address all correspondence to Dr Boissonnault at: jillboiss@email.gwu.edu.

K. Evans, PT, PhD, School of Allied Health Sciences, Menzies Health Institute Queensland, Griffith University–Gold Coast Campus, Gold Coast, Australia.

N. Tuttle, PT, PhD, School of Allied Health Sciences, Menzies Health Institute Queensland, Griffith University–Gold Coast Campus.

S.J. Hetzel, MS, Department of Biostatistics and Medical Informatics, University of Wisconsin–Madison.

W.G. Boissonnault, PT, DPT, DHSc, FAAOMPT, FAPTA, Executive Vice President, Professional Affairs, American Physical Therapy Association, 1111 N Fairfax St, Alexandria, Virginia. At the time of the study, Dr Boissonnault was affiliated with the Department of Orthopedics and Rehabilitation, Physical Therapy Program, University of Wisconsin–Madison.

[Boissonnault JS, Evans K, Tuttle N, et al. Reliability of the ECHOWS tool for assessment of patient interviewing skills. *Phys Ther.* 2016;96:443–455.]

© 2016 American Physical Therapy Association

Published Ahead of Print:

August 27, 2015

Accepted: August 18, 2015

Submitted: March 23, 2015



Post a Rapid Response to this article at:
ptjournal.apta.org

As noted in the *Guide to Physical Therapist Practice 3.0*¹ of the American Physical Therapy Association (APTA), patient interviews or patient histories comprise a major component of the initial examination process during nearly every patient/client encounter. Additionally, APTA lists history taking and communication competencies as part of the *Minimum Required Skills of Physical Therapist Graduates at Entry Level*.² In a recent editorial in the *Journal of Physical Therapy Education (JOPTE)*,³ Jan Dwyer and Laurita Hack presented the “JOPTE Editorial Board Recommendations for an Educational Research Agenda.” It included the following items, among others:

- Develop valid outcomes tools for all aspects of physical therapy education, including readiness to practice (ie, professionalism, clinical outcomes).
- Develop tools to assess with good discrimination to properly measure student trajectory with respect to patient outcomes.
- Develop measurement tools to measure student readiness for clinical education.

This list speaks to the need for creation of student assessment tools, including those for appraisal of competency in patient interviewing skills. In Australia, the Australian Physiotherapy Council includes collecting client information via effective communication skills sensitive to individual needs and diversity as a core competency for physical therapist practice standards.⁴ Refshauge and Gass stated that taking an effective patient history “is arguably the most important part of the examination process because it is from this that we decide the nature of the patient’s problem and the possible interventions that might consequently be used.”^{5(p117)} Davis⁶ argues for a “helping interview” based on a “healing attitude” of the practitioner. The healing attitude emanates from a practitioner who assists patients in resuming some control over their situation, listens and attends, communicates well, both verbally and nonverbally, and avoids judgment of the patients or their situation.⁶ Patient interviewing skills are initially

learned in the didactic component of professional physical therapist education programs and then honed during clinical education experiences.

Components of a physical therapist’s patient interview are similar to those of other health care disciplines, including investigation of the patient’s chief complaint, medical and surgical history, and medications. For physicians, the result may be diagnosis of a pathological process, whereas for the physical therapist, the outcome may be diagnosis of a movement disorder, a patient referral to a physician, or identification of biomechanical factors associated with loss of function. There are categories of information that physical therapists tend to emphasize more than physicians, such as environmental factors (eg, whether a patient’s living environment has stairs or floor coverings) and detailed postural, movement, and daily function assessment, that assist in goal setting and intervention strategies.¹ In summary, the physical therapy interview findings provide important information concerning the unique circumstances and context of the patient’s condition and the therapist’s understanding of the condition and ultimately assist in the development of a diagnosis, a prognosis, precautions and contraindications to treatment, and a plan of care.^{7,8}

Communication and establishment of rapport between practitioner and clinician are at the center of effective interviewing skills. Patient communication has been noted to be of particular importance in promoting effective patient outcomes and efficient patient encounter time management.⁹⁻¹¹ Effective interviewing skills include attentive listening, attention to cultural congruency, and ability to convey information at appropriate medical literacy levels.¹² Educators routinely use checklists, grading rubrics, and other formative and summative tools to assess students on a variety of competencies and skills.¹³ Although there are multiple assessment tools in the medical literature to score medical student and physician history taking skills, until recently there has been no validated tool in the literature created for the purpose of assessment of physical therapist

student competence in patient interviewing.¹⁴

The ECHOWS tool and the “Guide to ECHOWS” were developed at the University of Wisconsin–Madison in 2010 (see Appendixes 1 and 2). The ECHOWS tool consists of 2 sections: the ECHOW section and the summary of performance (S) section. The ECHOW section comprises the following elements: E (establishing rapport), C (chief complaint), H (health history), O (obtain psychosocial perspective), and W (wrap-up). The S section is more of a skills assessment, whereas the ECHOW section is more a listing of necessary components. For the ECHOW section, the specific aspects of each element are scored 0 (not observed) or 1 (observed), with a maximum score of 22 points. As the S section is more complex and likely requires skill and perspective of a physical therapist versus a student, the 10 items related to overall performance are scored on an ordinal scale as 0 (needing improvement), 1 (satisfactory), or 2 (superior), with a maximum score of 20 points. The ECHOWS tool scored as a whole has a potential total of 42 points. Pilot work established content validity and preliminary intrarater reliability but did not establish interrater reliability or construct validity.¹⁴ The purposes of this study were: (1) to confirm intrarater reliability and establish interrater reliability of the ECHOWS tool, (2) to determine the limits of agreement for the ECHOWS tool, and (3) to determine whether there are differences in history taking scores between novice and experienced practitioners using standardized patient (a person carefully recruited and trained to take on the characteristics of a real patient) encounters.

Method

Participants

Participants included entry-level students (novice practitioners), experienced clinicians, and faculty who served as the evaluators. Students were recruited from the University of Wisconsin–Madison (UW–Madison) Doctorate in Physical Therapy (DPT) Program and the Griffith University, Gold Coast Campus (GU), Masters in Physiotherapy (MPT) Program. Both are postgraduate, entry-level programs, with the UW–Madi-

son program lasting 3 years and the GU program lasting 5 semesters over 2 years. Ten Australian and US clinicians were individually recruited. The clinicians had a minimum of 3 years of full-time-equivalent, clinical orthopedic physical therapy experience over the prior 5 years. The students and clinicians were videotaped interviewing simulated patients, all utilizing the same scenario.

Three assessors, whose roles were to review and score the aforementioned tapes, were recruited by the US and Australian researchers, respectively. In the United States, a posting of the project and a request for a notification of interest to become 1 of the 3 assessors was placed on a listserv hosted by the Section on Education of APTA. Curricula vitae were requested, and the assessors were chosen based on familiarity with history taking pedagogy, faculty status in professional physical therapy curricula, and geographic diversity. Australian faculty members from 3 different universities, who were known to the researchers as having experience in clinical education and in teaching orthopedic and musculoskeletal skills, were invited to participate in the study.

Procedure

ECHOWS tool training. Students from both programs were trained in history taking via their respective curricula, which included familiarization with the ECHOWS tool. At UW-Madison, students complete a standardized patient interview during their first year of study as part of a unit on history taking and documentation. All of the patient interviews are taped. At GU, history taking is taught in progressive iterations for each area of practice. In the first semester, when this project was undertaken, students had received generic content in history taking and content specific to seeing inpatients having orthopedic surgery. This study utilized 10 taped patient interviews from UW-Madison and 10 taped patient interviews from GU. Students from the UW-Madison DPT class of 2015 were asked to sign a consent form granting utilization of their taped standardized patient interview from their physical therapy course titled “Foundations of PT Examination, Evaluation, and Diagnosis.”

Ten student tapes were randomly selected from the 34 students who granted consent. At GU, students are not routinely videotaped using standardized patients, but all students were offered the opportunity to participate in the present study, and the first 10 volunteers were accepted and signed consent forms prior to being videotaped.

The experienced clinician cohort was given information on the ECHOWS tool and the Guide to ECHOWS as part of their preparation for conducting the interviews of the standardized patients. They were instructed to introduce themselves to the standardized patients as physical therapist students so that the videotape assessors could not distinguish them from students. The assessors were trained in the use of the ECHOWS tool via an online module. The module contained information on the organization and scoring of the tool and provided a videotaped patient interview for practice scoring purposes. The viewer scored the interview using the ECHOWS tool and then could compare his or her scoring with that of the website instructor. The practice tape instructor scoring explained scoring for each item of the tool. Copies of the ECHOWS tool and the Guide to ECHOWS were available as PDF downloads on the training site.

Assessment of recordings. Each assessor evaluated 20 student tapes and 20 clinician tapes. Assessors were free to review the tapes as many times as needed to score them and were asked to complete this task within 3 weeks. A stamped, preaddressed envelope was provided to each of the assessors to use to return the 40 ECHOWS score sheets to the research team in their country. Three weeks after the receipt of the 40 ECHOWS score sheets, the assessors were sent a second CD loaded randomly with 10 student interviews and 10 clinician interviews and a second stamped, preaddressed envelope. They were asked to complete the reviews within a 2-week time frame.

Reimbursement. The experienced clinicians were given a gift card for \$10 (US) or Australian equivalent, respectively, for their time involved in prepara-

tion and in performing the interview. The assessors were paid \$25 per hour for their work in assessing the tapes. It was expected that it would take approximately 30 minutes to review and score each tape, with total time worked per assessor estimated at approximately 30 hours. A grant from the Department of Orthopedics and Rehabilitation at the School of Medicine and Public Health at UW-Madison covered these costs.

Data Analysis

Reliability was assessed by calculating intraclass correlation coefficients (ICCs) as defined by Shrout and Fleiss.¹⁵ Intrarater reliability was determined by calculation of an individual ICC (3,1) for each assessor, and interrater reliability was calculated using ICC (2,1). Ninety-five percent confidence intervals (CIs) also were calculated for the ICC types. Reliability was assessed separately for overall ECHOWS score, total ECHOW score, and total S score. Reliability was categorized according to Goldstein et al,¹⁶ with ICCs <.40 classified as poor, .40 to .60 as moderate; .61 to .80 as good, and >.80 classified as excellent. Ninety-five percent limits of agreement¹⁷ also were calculated for intrarater and interrater assessments for the ECHOWS overall and for the ECHOW and S components separately.

Additional analyses were conducted to assess whether there was any bias on the part of assessors in relation to students and clinicians from their home country or the foreign country. For example, Australian assessor scoring on Australian participants' tapes was compared with US assessor scoring on US participants' tapes, and Australian assessor scoring on US participants' tapes was compared with US assessor scoring on Australian participants' tapes. We also assessed whether there were significant differences between the Australian and US assessors on their scoring of the tapes as a whole.

The student cohort was compared with the experienced clinician cohort to see if there was a significant difference in performance on the ECHOWS tool. Comparisons were assessed using Student *t* tests

Reliability of the ECHOWS Tool

Table 1.
Intrarater Reliability^a

Assessor	Country	ICC (95% CI)	First Rating X̄ (SD)	Second Rating X̄ (SD)	Difference X̄ (SD)
ECHOWS overall					
1	United States	.80 (.57, .92)	26.5 (4)	26.9 (3.5)	-0.4 (2.3)
2	United States	.87 (.70, .95)	22.6 (4.9)	22.9 (4.5)	-0.4 (2.4)
3	United States	.83 (.61, .93)	21.6 (6.4)	20.9 (6.5)	0.8 (3.8)
4	Australia	.74 (.45, .89)	25.6 (5.1)	25.3 (4.8)	0.4 (3.6)
5	Australia	.89 (.74, .95)	28.4 (6.9)	28.5 (6.5)	-0.2 (3.2)
6	Australia	.82 (.59, .92)	23.7 (5.6)	21.8 (5.6)	2 (3.4)
Total (or overall)		.82	24.7 (5.9)	24.4 (5.9)	0.4 (3.2)
ECHOW sections					
1	United States	.93 (.83, .97)	14.3 (3.3)	14.2 (3.3)	0 (1.2)
2	United States	.95 (.87, .98)	13.8 (3.6)	13.6 (3.3)	0.2 (1.1)
3	United States	.92 (.81, .97)	13.2 (3.7)	12.3 (3.6)	0.8 (1.5)
4	Australia	.83 (.62, .93)	14.7 (2.9)	14.1 (3.3)	0.6 (1.8)
5	Australia	.91 (.78, .96)	13.2 (3.5)	13.1 (3.6)	0.2 (1.5)
6	Australia	.95 (.87, .98)	13.1 (3.1)	12.6 (3.1)	0.4 (1)
Total (or overall)		.93	13.7 (3.4)	13.3 (3.4)	0.4 (1.4)
S Section					
1	United States	.69 (.36, .86)	12.2 (2.5)	12.7 (2)	-0.4 (1.8)
2	United States	.68 (.35, .86)	8.8 (2.3)	9.3 (2.6)	-0.5 (2)
3	United States	.64 (.28, .84)	8.5 (3.2)	8.6 (3.4)	0 (2.8)
4	Australia	.67 (.34, .86)	10.9 (3.6)	11.2 (2.7)	-0.3 (2.6)
5	Australia	.85 (.65, .94)	15.1 (4.4)	15.4 (4.2)	-0.3 (2.4)
6	Australia	.74 (.45, .89)	10.7 (3.7)	9.2 (3.7)	1.5 (2.7)
Total (or overall)		.71	11.0 (4.0)	11.1 (3.9)	0 (2.4)

^a ICC=intraclass correlation coefficient, CI=confidence interval.

and repeated-measures analysis of variance models.

Role of the Funding Source

Research funding was provided by a grant from the Department of Orthopedics and Rehabilitation at the University of Wisconsin–Madison School of Medicine and Public Health.

Results

The average intrarater reliability values for the ECHOWS overall and the ECHOW sections (E, C, H, O, and W sections combined) were excellent, with ICC (3,1) values for all individual raters ranging from .74 to .89 for the ECHOWS overall and from .83 to .95 for the ECHOW sec-

tions. The overall intrarater reliability for the S section was good, with ICC (3,1) values ranging from .64 to .85 (Tab. 1). The 95% CI values overlapped for all assessors and subscales, indicating no significant differences in reliability between subscales or raters. The intrarater limits of agreement are: ECHOWS overall, ± 7 points; ECHOW sections, ± 3 points; and S section, ± 5 points. The interrater reliability values with ICC (2,1) for the ECHOWS overall and the ECHOW sections were moderate (ICC=.55) and excellent (ICC=.82), respectively. The S section, however, had poor interrater reliability (ICC=.27). For interrater agreement, the limits of agreement were: total, ± 8 points; ECHOW section, ± 4

points; and S section, ± 7 points. Further analysis of the intrarater data shows there was a significant difference between first and second scorings only for reviewer 3 in the ECHOW sections ($P=.018$), with an average difference of 0.85 points (SD=1.46). Looked at as a whole group across reviewers, there was a significant difference in ECHOW sections between the 2 readings (-0.39 ; 95% CI= -0.64 to -0.13 ; $P=.003$) but not for ECHOWS overall or for total S score. However, the difference was of less than 1 point and is not likely clinically relevant.

Additional analyses were conducted to assess whether there was any bias on the

Table 2.
Interrater Reliability^a

Measurement	All Tapes ICC (95% CI)	Student Tapes ICC (95% CI)	Clinician Tapes ICC (95% CI)
ECHOWS total	.55 (.36, .71)	.56 (.32, .77)	.55 (.33, .72)
ECHOW sections	.82 (.73, .89)	.85 (.73, .93)	.80 (.67, .90)
S section	.27 (.12, .44)	.23 (.08, .46)	.29 (.12, .53)

^a ICC=intraclass correlation coefficient, CI=confidence interval.

part of assessors in relation to students or clinicians from their home or foreign country. No such biases were found (Tab. 2).

With regard to construct validity, there was no statistical difference in performance on the tool between novice and experienced clinicians ($P=.59$). The results of the construct validity analysis are presented in Table 3.

Discussion

This study builds upon previous pilot work¹⁴ and provides further validation of the ECHOWS tool for evaluating history taking by entry-level physical therapist students. Limitations of the pilot study included the small numbers of reviewed videotapes and tape reviewers, making assessment of interrater reliability difficult. In addition to increasing the numbers of tapes and reviewers in the current study, the student and reviewer pools also were expanded to include participants from both the United States and Australia. These changes provided a more powerful design for assessment of intrarater and interrater reliability and construct validity and speak to the validity of use of the tool in Australia in addition to its use in the United States.

Compared with findings in the pilot study, intrarater reliability values were higher in the current study. In the pilot study, intrarater reliability was moderate for the ECHOW sections and good for the S section and the ECHOWS tool overall. In the current study, ratings were excellent for the ECHOWS overall and ECHOW sections and good for the S section. Interrater reliability results for the ECHOWS overall improved from poor in the pilot study to moderate in the current study and demonstrated excellent interrater reliability for the ECHOW sections. The S section interrater reliability remained low. There was no indication of cultural bias influencing the repeatability of the instrument, as no differences were detected between the assessors from the 2 nations or when assessors rated interviewers from their own country or the other country.

The following discussion will focus on the properties of the ECHOW sections and the S section before concluding how the 2 sections together might be used to provide both summative and formative feedback for students. It is perhaps not surprising that the items in the ECHOW sections had greater reliability than the S

section, as this section requires only a dichotomous decision of whether a particular behavior is observed or not. The S section, on the other hand, requires a more subjective, value judgment (ie, “needs improvement,” “satisfactory,” or “superior”) of more global items such as logical sequencing or communication strategies.

The intrarater and interrater limits of agreement for the ECHOW sections were 3 and 4 points, respectively, out of a possible 22 points, with 95% of the values in the ECHOW sections (mean±2 standard errors) between 7 and 20 points. These results suggest that, for any given individual, it would be possible to detect 3 or 4 significant changes in the score. For people with scores below 10 or 11 points, it would be difficult to detect a real deterioration due to a floor effect, and for those with scores over 16 or 17 points, it would be difficult to detect improvement due to a ceiling effect. The intrarater and interrater limits of agreement for the S section were 5 and 7 points, respectively, out of a possible 20 points, with 95% of the values between 3 and 19 points. These results suggest that, for a given individual, it would be possible to detect 1 to 3 significant changes in the score. For people with scores below 8 or 10 points, it would be difficult to detect a real deterioration due to a floor effect, and for those with scores over 12 or 14 points, it would be difficult to detect improvement due to a ceiling effect.

For the ECHOWS overall, the intrarater and interrater limits of agreement were 7 and 8 points, respectively, out of a pos-

Table 3.
Construct Validity Analysis^a

Measurement	Time	Student Tapes X̄ (95% CI)	Clinician Tapes X̄ (95% CI)	P
ECHOWS total	First viewing	25.1 (23.0, 27.2)	25.9 (23.8, 28.1)	.591
	Second viewing	25.0 (22.0, 27.9)	23.8 (20.8, 26.7)	.576
ECHOW sections	First viewing	14.0 (12.6, 15.4)	14.0 (12.6, 15.4)	.974
	Second viewing	13.7 (11.7, 15.8)	12.9 (10.9, 14.9)	.575
S section	First viewing	11.1 (10.1, 12.1)	12.0 (11.0, 13.0)	.245
	Second viewing	11.2 (9.8, 12.7)	10.9 (9.4, 12.3)	.724

^a CI=confidence interval.

sible 42 points, indicating the magnitude of a difference between repeated measures that is necessary to be confident of a real difference. With 95% of the scores between 13 and 36 points, the results suggest that, for any given score, it would be possible to detect 2 or 3 significant changes in the score. For people with scores below 20 or 21 points, it would be difficult to detect a real deterioration due to a floor effect, and for those with scores over 28 or 29 points, it would be difficult to detect improvement due to a ceiling effect.

For an evaluation or assessment tool to be useful, it is necessary for it to be able to discriminate between different levels of performance. The ECHOW sections fulfill this criterion, with the ability to discriminate 3 significant differences in a student's level of performance if the same educator is repeating the measure or 2 significant differences if a different educator is doing the evaluation. The S section, on the other hand, is less sensitive to change but is still able to distinguish at least between 2 extremes of performance. We suggest that a reasonable way of using the ECHOWS tool is to use the ECHOW sections for both summative and formative feedback and to use the S section primarily for formative evaluation. In other words, both sections can be used as a framework to provide feedback to assist student learning, but only the ECHOW sections are appropriate to be included in student marks.

Because the S section is more subjectively scored than the ECHOW sections and is not as sensitive to change, looking at the developed training module may be warranted to limit variability. The interrater reliability results support this need. Assessors were given training online, but the training may have been insufficient to ensure reliability in scoring this section of the tool, or there may have been variation in use of the online module. Additional practice scoring taped interviews prior to use in grading students or in formative assessment, combined with emphasizing the relative complexity of the S section to assessors and the importance of understanding the criteria for scoring, may enhance an assessor's reliability with this section of the tool. Addi-

tionally, there may be a need for refinement and additions to the training to promote consistency for scoring the S section. No significant differences were found between the novice and experienced practitioners. Although unexpected, there are a number of possible explanations for this finding. The ECHOWS tool was developed to evaluate students, and the characteristics of a history taken by a student may not correspond to that of an experienced practitioner. For example, Jensen et al¹⁸ noted that although master clinicians were able to deviate from the patient examination framework when it was deemed necessary to gather more detailed patient-focused information, novice clinicians organized their examination scheme based on following standard examination routines. Students implement a more scripted flow to ensure all areas are discussed, whereas experienced practitioners are flexible in their approach, leading to greater or lesser prioritization for some elements of the patient history. As for the S section of the tool, designed to capture interviewing skill level,¹⁴ one would expect more experienced clinicians to perform at higher levels than students; however, as noted above, the relatively small point total for the S section makes detecting a difference difficult.

Limitations and Future Studies

Further research should be conducted for assessment of entry-level students to establish instrument validity and applicability of use during clinical internships. The live clinical setting carries more uncertainty regarding patient responses, clinical interruptions, and time constraints—items that can be controlled for in an artificial setting. Academic programs would encourage clinical instructors to incorporate the ECHOWS tool and Guide to ECHOWS into the clinical education experience to begin determination of the tool's applicability in this setting. Such use would provide a resource for clinical instructors to reinforce and further develop previously learned interviewing skills and could be implemented at the beginning of the clinical education experience to identify areas in need of improvement. The stu-

dents could then use the familiar tool for continued self-assessment.

The results of this study suggest the need for the S section to evolve in order to provide a more discerning measure of interviewing skill level. Utilization of the S section to provide feedback to assist student learning provides value in an academic setting, but the difficulty in discerning different levels of skill is a limitation of the tool as it is currently formatted. Besides providing a mechanism to monitor student progress, rectifying this issue may allow the tool to be of greater value in self-assessment and peer assessment and for enhancement of these skills in physical therapist practitioners at various stages of their careers. Such tool refinement also would provide researchers with another instrument to measure interviewing skill levels.

Regarding the inability of ECHOWS, in its current format, to discern novice from advanced practitioners, future research could utilize patient scenarios with greater complexity and range of responses, along with concomitant training in more complex and obtuse responses for the standardized patient, allowing for enhanced clinical decision-making opportunities for the student in a videotaped situation. Additionally, this study did not assess whether student or clinician race or ethnicity influenced scoring on the ECHOWS tool other than the comparison of scores between US and Australian participants. This is an area that could be assessed in future studies. Lastly, the relatively small sample, particularly of assessors, limited the information that could be gained from the current study. Despite these limitations, this study provides further support for the reliability of a tool consistent with the recently proposed Educational Research Agenda.

In conclusion, intrarater and interrater reliability have been established for the ECHOWS tool in assessing physical therapist student history-taking skills. This is the first study of this type of educational assessment tool that included physical therapist students from 2 different countries. As there were no differences in ratings completed by US and Australian

assessors, nationality-based bias in rating does not appear to be a factor in scoring of students and clinicians with the ECHOWS tool. The ECHOWS tool may be valid in assessing Australian physical therapist students' patient interviewing skills in addition to assessing skills in US DPT students. Assessor training with the ECHOWS tool appears to be important in maintaining reliability in student assessment.

All authors provided concept/idea/research design and contributed to writing. Dr J. Boissonnault, Dr W. Boissonnault, Dr Evans, and Dr Tuttle provided data collection. Dr Tuttle, Mr Hetzel, Dr J. Boissonnault, and Dr W. Boissonnault provided data analysis. Dr J. Boissonnault provided project management and fund procurement. Dr J. Boissonnault and Dr Tuttle provided participants. Dr J. Boissonnault, Dr W. Boissonnault, Dr Evans and Dr Tuttle provided facilities/equipment. All authors provided institutional liaisons and consultation (including review of manuscript before submission).

The authors acknowledge the assistance of Nicholas Conte, DPT, LAT, for his help in development of the ECHOWS training module, and Sarah Stream, MS, DPT, for her assistance in data input. Additionally, they recognize and thank the University of Wisconsin–Madison and Griffith University students, and the US and Australian faculty and clinicians who assisted the project as participants and assessors.

The study received UW–Madison Social Science Institutional Review Board approval and Griffith University Institutional Review Board approval.

Research funding was provided by a grant from the Department of Orthopedics and Rehabilitation of the University of Wisconsin–Madison School of Medicine and Public Health.

DOI: 10.2522/ptj.20150172

References

- 1 *Guide to Physical Therapist Practice 3.0*. Alexandria, VA: American Physical Therapy Association; 2014. Available at: <http://guidetoptpractice.apta.org>. Accessed March 13, 2015.
- 2 American Physical Therapy Association. Minimum Required Skills of Physical Therapist Graduates at Entry Level [Guideline] (BOD G11-05-20-49). Available at: http://www.apta.org/uploadedFiles/APTAorg/About_Us/Policies/Education/Minimum_RequiredSkillsPTGrads.pdf. Accessed March 13, 2015.
- 3 Gwyer J, Hack LM. A challenge [editorial]. *JOPT*. 2015;29:4–5.
- 4 *Australian Standards for Physiotherapy: Safe and Effective Physiotherapy*. Canberra, Australia: Australia Physiotherapy Council; 2006. Available at: <https://physiocouncil.com.au/media/1021/the-australian-standards-for-physiotherapy-2006.pdf>. Accessed March 13, 2015.
- 5 Refshauge K, Gass E. *Musculoskeletal Physiotherapy: Clinical Science and Evidence-Based Practice*. Oxford, United Kingdom: Butterworth Heinemann; 2004.
- 6 Davis CM. *Patient Practitioner Interaction. An Experiential Manual for Developing the Art of Health Care*. 5th ed. Thorofare, NJ: Slack Inc; 2011:159–177.
- 7 Garber M, Boissonnault WG. The patient interview: the science behind the art. In: Boissonnault WG, ed. *Primary Care for the Physical Therapist: Examination and Triage*. 2nd ed. St Louis, MO: Elsevier Saunders; 2010:53–63.
- 8 Fruth SJ. *Fundamentals of the Physical Therapy Examination: Patient Interview and Tests and Measures*. Burlington, MA: Jones & Bartlett Learning; 2014.
- 9 Ambady N, Koo J, Rosenthal R, Winograd CH. Physical therapists' nonverbal communication predicts geriatric patients' health outcomes. *Psychol Aging*. 2002;17:443–452.
- 10 Hall AM, Ferreira PH, Maher CG, et al. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. *Phys Ther*. 2010;90:1099–1110.
- 11 Hush JM, Cameron K, Mackey M. Patient satisfaction with musculoskeletal physical therapy care: a systematic review. *Phys Ther*. 2011;91:25–36.
- 12 Drench ME, Noonan AC, Sharby N, Ventura H. *Psychosocial Aspects of Health Care*. 3rd ed. Upper Saddle River, NJ: Pearson; 2012.
- 13 Jensen GM, Shepard KF. Techniques for teaching and evaluating students in academic settings. In: Shepard KF, Jensen GM, eds. *Handbook of Teaching for Physical Therapists*. 2nd ed. Boston, MA: Butterworth Heinemann; 2002:71–131.
- 14 Boissonnault JB, Boissonnault WG, Hetzel SJ. Development of a physical therapy patient-interview student assessment tool: a pilot study. *JOPT*. 2013;27:35–47.
- 15 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–428.
- 16 Goldstein G, Beers SR, Hersen M, eds. *Comprehensive Handbook of Psychological Assessment: Intellectual and Neuropsychological Assessment*. Vol 1. Hoboken, NJ: John Wiley & Sons Inc; 2003:317.
- 17 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 2011;1:307–310.
- 18 Jensen GM, Shepard KF, Gwyer J, Hack LM. Attribute dimensions that distinguish master and novice physical therapy clinicians in orthopedic settings. *Phys Ther*. 1992;72:711–722.

Reliability of the ECHOWS Tool

Appendix 1. ECHOWS Tool^a

<p>ECHOWS</p> <p>A Physical Therapy Patient Interview Assessment Tool</p> <p>Name/Number: _____</p> <p>Date: __/__/__</p>	<p>Comments: Available for additional information</p> <p>Scoring for sections E, C, H, O, and W: Observed or Not Observed: Item is pertinent to the interview, and the interviewer either includes it or does not. Number of observed scores given is tallied as the numerator for each section.</p> <p>Scoring for section S: 0=Needs Improvement: Check this if the interviewer does not perform the item well and elaborate in "Comments" section</p>		
E: Establish Rapport	Observed	Not Observed	Comments
1. Introduction/greeting			
2. Orients patient to the flow of the visit			
Number of Observed items for E_/2			
C: Chief Complaint	Observed	Not Observed	Comments
1. Reason for visit: (chief complaint, including location of symptoms)			
2. Functional status in various roles and realms (eg, home, work, school, social)			
3. Patient's goals and expectations for treatment and prevention			
4. History of chief complaint			
5. Location/behavior of symptoms: aggravation, alleviation, nature (includes intermittent or constant, description, how long symptoms last, quantification of symptoms, 24-h presentation)			
6. Previous examination/tests/interventions for chief complaint			
Number of Observed items for C_/6			
H: Health History	Observed	Not Observed	Comments
1. Review of constitutional symptoms (fatigue, weakness, sweats, night pain, weight loss, confusion)			
2. Review of body systems			
3. Surgeries (including type and date)			
4. Allergies (including latex and to drugs)			
5. Other illnesses/health conditions			
6. Medication: prescription and OTC/herbals			
7. Health habits: substance use (caffeine, tobacco, alcohol) and exercise			
8. Abuse history (family violence, sexual, physical, and/or emotional abuse)			
9. Pertinent family medical history			
Number of Observed items for H_/9			
O: Obtain Psychosocial Perspective	Observed	Not Observed	Comments
1. Patient perception of chief complaint			
2. Family, social, and personal circumstances			
3. Environmental barriers/accommodations			
Number of Observed for O_/3			

(Continued)

Appendix 1.

Continued

W: Wrap-up		Observed	Not Observed	Comments	
1. Asks patient about additional questions/concerns					
2. Transition into physical exam (gives clear information about next steps of the process)					
Number of Observed items for W_/2					
Total Number of Observed items for E, C, H, O, and W: ____/22					
S: Summary of Performance		Needs Improvement 0	Satisfactory 1	Superior 2	Comments
1. Attends to patient comfort and privacy					
2. Logical sequencing/follows an organized format					
3. Time management/keeps the interview on task					
4. Questioning strategies (eg, avoids leading questions, avoids duplication, explains rationale for questions, all questions are relevant)					
5. Verbal communication strategies (avoid jargon and repetitive verbal habits ["umm"], checks for patient understanding, rephrases/summarizes, uses transition statements)					
6. Documents without interfering with the flow of the interview/not distracting					
7. Attentive listening (interrupts the patient only when redirecting is needed)					
8. Respect and interest toward patient; makes a personal connection					
9. Nonverbal behavior (eg, distance and eye contact are comfortable for patient; responds to nonverbal patient cues)					
10. Social skills (eg, empathy, poise, handles embarrassing/sensitive topics, accepts and legitimizes patient's feelings and beliefs)					
TOTAL SCORE for S_/20		0	Subtotal × 1	Subtotal × 2	
TOTAL SCORE FOR ECHOWS ____/42					

Reviewed by: _____

^a Property of Jill S. Boissonnault and William G. Boissonnault. Developed at the University of Wisconsin–Madison. Version dated May 11, 2011. OTC=over the counter.

Reliability of the ECHOWS Tool

Appendix 2.

Guide to ECHOWS^a

Guide to ECHOWS (A Physical Therapy Patient Interview Assessment Tool)

This “Guide” is offered to assist faculty, clinical supervisors, clinical instructors, or anyone assessing a therapist’s interviewing skill in appreciating the details of the subcategories under each section of the ECHOWS tool. The gray-bar subheadings correspond to the items on the tool itself. The Guide can be used as an educational tool by the trainer. Identifying the subcategory details should help the learner ascertain content criteria that, when included in an interview, improves it.

A note on the ECHOWS tool: ECHOWS is designed to assess how effectively and efficiently the student collects patient data in order to move to the next steps in an examination session: generation of a problem list, development of hypotheses, planning of the physical examination, and setting goals for therapy. It is not intended to assess skill in hypothesis formation or critical thinking.

ECHOWS sections can be used independently of one another for teaching a particular set of items from the patient history (ie, establishing rapport or history of the chief complaint). The S section (summary of performance) may be used without any of the other sections to assess skill vs the checklist of content found in E, C, H, O, and W. History taking is not always sequenced as we have categorized it (eg, items listed in the O section are likely to have been asked in conjunction with pieces of the chief complaint listed in the C section). Each interview may flow differently, and practitioners find ways to connect questions together that make sense for that patient, that situation, and their own style. Therefore, the location of the categories in the document is meant as a guide and is not intended to infer that items must be asked in the order found in the instrument.

Finally, ECHOWS assessment does not preclude the use other patient history forms or databases such as a body diagram, medical record, or health history questionnaire. Standardized outcome measurement tools can be completed prior to the initial visit and may serve as benchmarks and assist in quantification of symptoms.

Suggested Pedagogy

A suggested teaching strategy would be to provide students with this Guide as part of an instructional unit on patient history taking, accompanied by any readings. Once the students grasp details of the patient history and have an opportunity to practice with case scenarios, they might then be assessed interviewing patients or simulated patients. The ECHOWS tool can be used: (1) to self-assess taped interviews, (2) to assess peers in real or taped interviews, (3) by faculty or supervisors to assess students or clinicians in real time or with taped interviews, and (4) by simulated patients to score their interviewer. This Guide can also be used to develop interview intake forms.

Scoring

Please read the legend at the top of the ECHOWS tool. Note that for each section of E, C, H, O, and W, you score simply as to whether the interviewer has completed the task: observed or not observed. Notes can be made in the comments section to the right for any construct. For the S section, you are asked to score the interviewer based on skill or minimum competence for a new graduate. A “needs improvement” score would indicate the student or therapist is not yet competent at this skill or activity. A score of “satisfactory” means the interviewer has performed at a competent level for that particular construct, akin to a new graduate’s performance. A score of “superior” should be given when an interviewer exceeds basic competency by demonstrating a higher, more advanced skill level; beyond expectations for a new graduate. PLEASE SCORE EVERY ITEM. The Final Score is a summation of the ECHOW score and the S score.

E: Establish Rapport
Introduction/greeting
(a) Introduce oneself, including name and profession (physical therapy)
(b) Welcome patient and family (determine patient’s preferred way to be addressed)
(c) Uses patient’s name/knows patient’s name
Orients patient to the flow of the visit
(a) Explains how the history taking will unfold
(b) Explain what happens after the history is completed

(Continued)

Appendix 2.
Continued

C: Chief Complaint
Reason for visit: What brought the patient into physical therapy (chief complaint)?
Functional status: How does chief complaint affect patient's function and quality of life (various roles and realms: home, work, school, social)?
Patient's goals and expectations for treatment and prevention
History of chief complaint
(a) Facilitates patient story: begins with current/recent history, then explores the past chronologically, including dates and sequences
(b) Allows patient to tell story in his or her own words
Signs and symptoms; include intensity scale 0–10 (VAS or other)
(a) Location and symptom descriptors; may refer to a body diagram
(b) Behavior of symptoms: aggravation, alleviation, nature: intermittent or constant. Description; how long symptoms last. Quantification of symptoms; VAS/other outcome measure.
Previous examination/intervention for chief complaint
(a) Self-treatment
(b) Previous physical therapy. If yes: was it beneficial? Which treatment helped or not?
(c) Other interventions from non-PT providers
(d) Past tests for chief complaint
H: Health History
Review of systems—to identify secondary complaints
(a) Constitutional symptoms (fatigue, weakness, fever, sweats, malaise, night pain, unexplained weight loss/gain, confusion)
(b) Cardiovascular and pulmonary
(c) Neuromuscular
(d) Musculoskeletal
(e) Integumentary
(f) Genitourinary/reproductive system (including number of pregnancies)
(g) Endocrine
(h) Gastrointestinal
(i) Psychological (including depression)
Surgeries (including type and date)
Allergies
(a) Do you have any?
(b) Nature of reactions
(c) Latex allergies
(d) Allergy treatment, including emergency situations, and does the patient carry the medications?
Other current/recent illnesses/health conditions
Medications
(a) Reason for taking, effectiveness, dose, dosing schedule
(b) How long patient has been taking the medication, and is it effective?
(c) Over-the-counter medications
(d) Prescription medications
(e) Herbal/supplements

(Continued)

Reliability of the ECHOWS Tool

Appendix 2.

Continued

Health habits
(a) Substance abuse
(b) Alcohol consumption
(c) Tobacco use
(d) Caffeine consumption
(e) Exercise (including mode, frequency, intensity)
Abuse history (family violence, sexual abuse, physical abuse, emotional abuse)
Pertinent family medical history (eg, diabetes, CAD, CA, HTN)
O: Obtain Psychosocial Perspective
How does patient perceive his or her chief complaint (ie, what patient believes is going on; includes patient's world view and beliefs about illness/health and may include family beliefs)?
(a) The etiology
(b) Its progression
Family, social, and personal circumstances: contextual factors that may affect health
(a) Family issues
(b) Cultural issues (ethnicity, religion, community)
(c) Socioeconomic issues
(d) Educational level
(e) Personal circumstances (including stressors)
(f) Available support network
Environmental barriers/accommodations; physical spaces, structures, and obstacles they interact with
(a) Home and surroundings
(b) Work
(c) Social/public spaces
W: Wrap-up
Asks patient if there are any additional questions or concerns
Transition into physical exam (gives clear information about next steps of the process)
S: Summary of Performance
Utilizes patient information from additional sources (eg, health history questionnaire, body diagrams, medical record); reviews orally with patient pertinent items of the available patient data
Attends to patient comfort and privacy/asks patient if he or she is comfortable
(a) Sitting comfort (room temperature, stool, pillow, type of seat)
(b) Provides/maintains patient privacy
Logical sequencing: follows an organized format and sequence makes sense
Time management: keeps the interview on task
Questioning strategies
(a) Moves from open-ended questions to more focused questions when appropriate
(b) Avoids leading questions
(c) Explains rationale for questions when needed
(d) Avoids duplication of questions
(e) All questions are relevant to the interview

(Continued)

Appendix 2.
Continued

Verbal communication strategies
(a) Allows patient to express full set of concerns or questions
(b) Avoids jargon/uses understandable language
(c) Avoids repetitive verbal habits (eg, "umm," "Okay," "That's good")
(d) Periodically, checks for patient understanding of flow of the process, the questions themselves
(e) Rephrases and summarizes to ensure mutual comprehension
(f) Utilizes transition statements as the interview moves from one category to another
Documents without interfering with the flow of the interview/not distracting
Attentive listening (interrupts the patient only when redirecting is needed)
Demonstrates respect and interest toward patient: makes a personal connection
(a) Tone of voice demonstrates respect and interest toward patient
(b) Questions preferred learning style for patient if patient seems lost and is not following questions well
(c) Determines whether the patient has any learning issues (eg, learning disabilities, literacy issues, ESL issues) if patient seems lost and is not following questions well
Displays effective nonverbal behavior/responds appropriately to nonverbal behavior by patient
(a) Body language/posture/proxemics (spacing between patient and practitioner, attempts to be at approximately eye level with patient); demonstrates care, concern, interest
(b) Eye contact parallels patient's cultural practice and comfort
(c) Head nodding, facial expression, and gestures encourage communication
(d) Recognizes and appropriately responds to nonverbal patient cues
Demonstrates appropriate social skills
Demonstrates empathy
Is calm, poised, and relaxed
Deals sensitively with embarrassing or disturbing topics
Accepts and legitimizes patient's feelings and beliefs

^a Property of Jill S. Boissonnault and William G. Boissonnault. Developed at the University of Wisconsin–Madison. Version dated May 10, 2011. VAS=visual analog scale, PT=physical therapy, CAD=coronary artery disease, CA=cancer, HTN=hypertension, ESL=English as a second language.