

The Adapted Fresno test for speech pathologists, social workers, and dietitians/nutritionists: validation and reliability testing

Lucylynn Lizarondo
Karen Grimmer
Saravana Kumar

International Centre for Allied
Health Evidence, University of South
Australia, Adelaide, Australia

Purpose: The current versions of the Adapted Fresno test (AFT) are limited to physiotherapists and occupational therapists, and new scenarios and scoring rubrics are required for other allied health disciplines. The aim of this study was to examine the validity, reliability, and internal consistency of the AFT developed for speech pathologists (SPs), social workers (SWs), and dietitians/nutritionists (DNs).

Materials and methods: An expert panel from each discipline was formed to content-validate the AFT. A draft instrument, including clinical scenarios, questionnaire, and scoring rubric, was developed. The new versions were completed by ten SPs, 16 SWs, and 12 DNs, and scored by four raters. Interrater reliability was calculated using intraclass correlation coefficients (2,1) for the individual AFT items and the total score. The internal consistency of the AFT was examined using Cronbach's α .

Results: Two new clinical scenarios and a revised scoring rubric were developed for each discipline. The reliability among raters was excellent for questions 1, 3, and 6 across all disciplines. Question 7 showed excellent reliability for SPs, but not for SWs and DNs. All other reliability coefficients increased to moderate or excellent levels following training. Cronbach's α was 0.71 for SPs, 0.68 for SWs, and 0.74 for DNs, indicating that internal consistency was acceptable for all disciplines.

Conclusion: There is preliminary evidence to show that AFT is a valid and reliable tool for the assessment of evidence-based practice knowledge and skills of SPs, SWs, and DNs. Further research is required to establish its sensitivity to detect change in knowledge and skills following an educational program.

Keywords: Adapted Fresno test, evidence-based practice, speech pathology, social work, dietetics/nutrition

Introduction

The importance of evidence-based practice (EBP) in allied health is well documented in the literature.^{1,2} Clinical decisions that are based on patients' unique circumstances, sound clinical expertise, and the best available research evidence are known to deliver the best outcomes for patients and their families.³⁻⁵ Allied health practitioners hold positive attitudes toward EBP and believe in the value of research evidence in informing their clinical decisions. However, applying research findings to clinical decisions is not a simple process and is often difficult to achieve. One of the most commonly reported barriers to evidence uptake in allied health is the lack of knowledge of the EBP process and lack of skill in critically appraising research.⁶⁻⁸ Teaching EBP is therefore an important step in promoting evidence-based clinical decision making.

Correspondence: Lucylynn Lizarondo
International Centre for Allied Health
Evidence, University of South Australia,
GPO Box 2471, Adelaide, SA 5000,
Australia
Tel +61 8 8302 2099
Fax +61 8 8302 2766
Email lucylynn.lizarondo@unisa.edu.au

Allied health practitioners need to understand the principles of EBP before they can apply it.

Early EBP educational programs include the development of clinical questions, literature searches, and critical appraisal.⁹ To evaluate the impact of such educational programs and document competence of individual practitioners, educators need objective and psychometrically sound instruments or assessment tools. Based on a review of the literature, the Fresno test is the only available instrument that comprehensively assesses EBP competence across all relevant domains.¹⁰ The Fresno test consists of two clinical scenarios and 12 short-answer questions that require respondents to formulate a focused question, identify the most appropriate research design that will address the question, show knowledge of electronic database searching, identify issues important for determining the relevance and validity of a research paper, and discuss the magnitude and importance of research findings.¹¹ The test is scored by using a standardized grading rubric that describes explicit grading criteria. The Fresno test has content validity, good-to-excellent interrater reliability for all questions, and excellent internal consistency.¹¹ However, this tool focuses on assessing competence in medical students only, and therefore it cannot be used across different health disciplines.

In 2009, McCluskey and Bishop modified the Fresno test to measure the change in EBP skills and knowledge of occupational therapists following exposure to an EBP workshop.¹² New clinical scenarios (ie, versions 1 and 2) were developed to suit rehabilitation professionals, such as physiotherapists and occupational therapists. The 12 questions in the original Fresno test were reduced to seven (ie, questions 1–7), removing questions about diagnosis and complex statistics (ie, questions 8–12). The scoring rubric was also revised. Similar to the original Fresno test, the seven-item Adapted Fresno test (AFT) measures the following: the ability to develop a focused clinical question using the PICO (population, intervention, comparison, and outcome) format, the ability to develop a search strategy, the ability to interpret and critically appraise a research paper, and knowledge associated with understanding of the hierarchy of evidence and methodological biases in study designs, databases, and other sources of evidence and study designs. The AFT has been reported to have acceptable psychometric properties: interrater reliability ranged from good to excellent for individual items (version 1, intraclass correlation coefficient [ICC] 0.80–0.96; version 2, 0.68–0.94) and excellent for the total score (version 1, 0.96; version 2, 0.91); acceptable internal consistency (Cronbach's α 0.74); and responsive to change in novice learners.¹²

The current versions of the AFT are limited to physiotherapists and occupational therapists, and new scenarios and scoring rubrics are required for other allied health disciplines. Therefore, the aim of this study was to examine the validity, interrater reliability, and internal consistency of AFT versions developed for speech pathologists, social workers, and dietitians/nutritionists.

Materials and methods

This study was approved by the Human Research Ethics Committee of the University of South Australia and the Ethics Review Board of the University of Tasmania.

Development and content validation of AFT for speech pathology, social work, and dietetics/nutrition

An expert panel consisting of four practitioners from each discipline was formed to content-validate the AFT. Content validity refers to "... how well the combined elements used to construct the instrument truly describe the conceptual domain of interest".¹³ The panel represented practitioners with more than 10 years of clinical experience and with previous exposure to EBP training or research. The majority had graduate degrees in their respective disciplines or other clinical areas.

The panel members were presented with the original Fresno test and AFT, and were asked to examine the questionnaire and comment on which questions should be included in the new versions for speech pathologists, social workers, and dietitians/nutritionists. All members agreed that only questions adapted by AFT should be included for these disciplines. Following discussion, new clinical scenarios were developed for each discipline. The scoring rubric of the AFT was considered applicable to the new versions except for questions 1 ("Write a focused clinical question for one scenario to help you organize a search of the literature"), 2 ("Where might you find answers to these and other similar clinical questions? Name as many possible sources of information as you can, not just the ones you think are good sources"), and 4 ("If you were to search for Medline for original research to answer your question, describe the search strategy you might use"). Discipline-specific information was required to revise the scoring key for these questions.

Following consultation with the expert panel, a draft instrument including the clinical scenarios, questionnaire, and scoring rubric was prepared by the primary author. The draft instrument was emailed to the experts for feedback on the clarity of the entire instrument and completeness of the scoring rubric. The instrument and scoring rubric for each

discipline were revised based on comments from the expert panel and returned to them for a final round of feedback. No further changes were required in the instrument.

Participants

The new AFT versions were completed by ten speech pathologists, 16 social workers, and 12 dietitians/nutritionists who agreed to participate in a larger study aimed at examining the impact of a journal club on the EBP knowledge and skills of allied health professionals.¹⁴ They were asked to individually complete either a paper-and-pencil version or electronic version of the questionnaire at a time convenient for them. There were equal numbers of participants who held bachelor's degrees and postgraduate degrees. Less than half had previous training in research or EBP, and the majority had been in clinical practice for less than 10 years.

Interrater reliability of the AFT

Interrater reliability is the "... degree to which measurements of the same phenomenon by different raters will yield the same results, or the consistency of results between raters".¹⁵ Interrater reliability was calculated for individual items and the total AFT score using ICCs (2,1) and 95% confidence intervals. For interpretation of results, ICC values of ≥ 0.80 indicate excellent reliability, values between 0.60 and 0.79 denote moderate reliability, and values < 0.60 mean questionable reliability.¹⁶

Four individuals experienced in research and teaching EBP for allied health students served as raters for the study. Before the study began, the raters reviewed and discussed the AFT test, and collaboratively scored a sample test for each discipline. They were then given a practice period, where they scored another set of sample tests, then compared and discussed their differences in scoring. Following discussion, the raters were instructed to score each test independently without conferring or comparing ratings. Raters were given 2 weeks to mark all questionnaires.

Initial examination of the interrater reliability showed poor reliability between raters for questions 2, 4, and 5 of all versions (ie, AFT for speech pathology, social work, and dietetics/nutrition) and question 7 for social work and dietetics/nutrition. This prompted the first author, who has experience in using the previous AFT versions, to provide further training and discussion of the scoring procedure to the raters. The training involved an explanation of the rating system, discussion of common rater errors, advice on process for decision making, and practice on interpreting the rubric. Those questions with poor reliability were rescored 2 weeks later.

Internal consistency of the AFT

Internal consistency reflects the coherence of the components of a scale or instrument.¹⁷ The internal consistency of the AFT was examined using Cronbach's α .

Results

Content validity of the AFT for speech pathologists, social workers, and dietitians/nutritionists

The content validity of the AFT instrument was established through formal feedback from the expert panel. The comments received were consistent across disciplines, and involved issues associated with the wording of the clinical scenarios. No comments were made on the questionnaire itself; however, additional possible answers were suggested for the scoring rubric. For example, in question 1, where respondents are asked to write a focused clinical question, the expert panel provided additional PICO terms or synonyms. Some members of the panel suggested further sources of research information for question 2, such as discipline-specific electronic databases, websites, and professional organizations.

Two new clinical scenarios and a revised scoring rubric were developed for each discipline. Table 1 shows the final versions of the clinical scenarios. Table 2 lists the questions included in the new AFT versions. A copy of the scoring rubric may be obtained from the primary author upon request.

Interrater reliability of the AFT

The reliability among raters was excellent for questions 1, 3, and 6 across all disciplines, as shown in Table 3. Question 7 showed excellent reliability for speech pathology, but not for social work or dietetics/nutrition. All other reliability coefficients increased to moderate or excellent levels following further training and discussion.

Internal consistency of the AFT

Cronbach's α was 0.71 for speech pathology, 0.68 for social work, and 0.74 for dietetics/nutrition, indicating internal consistency was acceptable for all disciplines. Deletion of any of the items did not improve the internal consistency of the AFT for any discipline.

Discussion

The results provide preliminary evidence of the psychometric integrity of the AFT, and support its use in the assessment

Table 1 Discipline-specific clinical scenarios

| Clinical scenario 1 | Clinical scenario 2 |
|---|---|
| <p>Speech pathology clinical scenarios</p> <p>Laura is a 50-year-old lawyer who was recently diagnosed with squamous cell carcinoma of the nasopharynx. You have heard that providing swallowing exercises prior to her receiving chemoradiation therapy may help. You would like to find out from the literature whether there is any evidence to support the use of such exercise in improving swallowing function after chemotherapy treatment.</p> | <p>A 64-year-old lady with chronic anomic aphasia secondary to stroke has been referred to you for speech therapy. She had been participating in extensive language therapy and felt she was no longer progressing. You would like to know if using constraint-induced aphasia therapy would further improve her language skills.</p> |
| <p>Social work clinical scenarios</p> <p>You have received a referral for a 38-year-old male client with alcohol-use disorder. He began drinking 6 years ago to manage work-related stress. He indicates that he wants to reduce his alcohol consumption, but has not been successful. You want to find out whether there is any evidence to support the use of motivational interviewing for alcohol abuse over an educational intervention.</p> | <p>A 52-year-old single lady with a long history of obsessive-compulsive disorder (OCD) has been referred to you for behavior therapy. Her obsessions involve severe fear of contamination and having to urinate. Her compulsions involve excessive washing behaviors and avoiding places without an easy escape or readily accessible bathrooms. She fears being “an OCD” her entire life, having received many years of therapy with little effect on her symptoms. You would like to know if there is value in using acceptance and commitment therapy to reduce feelings of anxiety and distress.</p> |
| <p>Dietetics/nutrition clinical scenarios</p> <p>A 58-year-old housewife has been referred to a dietetics outpatient clinic for advice on dietary management of her chronic kidney disease, including high potassium levels. Her urea and creatine levels are significantly higher than normal, and have continued to gradually increase since diagnosis. This patient is not receiving dialysis; however, her renal specialist indicates that this will likely need to commence in a few years if there is ongoing deteriorating kidney function. You want to find out from the published literature the most effective dietary management to prevent progression of the kidney problem in a nondiabetic patient.</p> | <p>You have been referred to a 75-year-old male inpatient for possible enteral or parenteral feeding. The patient was admitted to hospital 3 days ago with severe abdominal pain and vomiting. He has been unable to manage an oral diet, and is currently receiving intravenous fluids. Tests indicate that the patient is suffering from pancreatitis, a condition that he has never experienced before. Doctors are managing the medical condition conservatively at present, with no indications for surgery. You want to find out the best nutrition-intervention approach for optimal outcomes.</p> |

of EBP knowledge and skills of speech pathologists, social workers, and dietitians/nutritionists. Similar to the original AFT, the new versions assess knowledge and skills of the key processes involved in EBP, including the development of

clinical questions, searching for literature, critical appraisal, and interpretation of research findings. The new AFT has content validity, moderate-to-excellent reliability and acceptable internal consistency. These results are consistent with the

Table 2 Questions in the Adapted Fresno test

Introduction: Please read the two clinical scenarios, and try to answer all of the following questions to the best of your ability. Do not worry if you are unfamiliar with the diagnoses mentioned; this should not affect your answers. You will find most of the following questions quite challenging, and will need to think carefully when answering them. If you are unsure of an answer, please say so.

Q1: Write a focused clinical question for one of the scenarios that will help you to organize a search of the clinical literature.

Q2: Where might you find answers to these and other similar clinical questions? Name as many possible sources of information as you can, not just the ones you think are “good” sources. Describe the most important advantages and disadvantages of each type of information source you have listed.

Q3: What type of study (design) would best answer your clinical question (see Q1), and why?

Q4: If you were to search Medline for original research to answer your clinical question, describe the search strategy you might use. Be as specific as you can about which topics and search categories (fields) you would use. Explain your rationale for taking this approach. Describe how you might limit your search if necessary and explain your reasoning.

Q5: When you find a report of original research on this question or any others, what characteristics of the study will you consider, to determine if it is relevant? (Q6 and Q7 will ask you how to determine if the study is valid and how important the findings are. For this question, please focus on how to determine if it is really relevant to your practice.)

Q6: When you find a report of original research related to your clinical question or any others, what characteristics of the study will you consider, to determine if its findings are valid? (You’ve already addressed relevance, and Q7 will ask how to determine the importance of the findings. For this question, please focus on the validity of the study.)

Q7: When you find a report of original research that relates to your clinical question or any others, what characteristics of the findings will you consider to determine their magnitude and significance (clinical and statistical)?

Note: Copyright © 2009 The Alliance for Continuing Medical Education, the Society for Academic Continuing Medical Education, and the Council on CME, Association for Hospital Medical Education. Reproduced with permission from John Wiley and Sons. McCluskey A, Bishop B. The Adapted Fresno Test of competence in evidence-based practice. *J Contin Educ Health Prof.* 2009;29(2): 119–126.¹²

Table 3 Interrater reliability of the Adapted Fresno test (individual items and total score)

| Adapted Fresno item | Speech pathology n=10 ICC (95% CI) | Social work n=16 ICC (95% CI) | Dietetics/nutrition n=12 ICC (95% CI) |
|---|---|---|---|
| Q1: Write a focused clinical question for one scenario to help you organize a search of the clinical literature. | 0.98 (0.94–0.99) | 0.98 (0.96–0.99) | 0.97 (0.93–0.99) |
| Q2: Where might you find answers to these questions? Name as many possible sources of information as you can. List advantages and disadvantages. | 0.95 (0.88–0.98) 0.74 (0.35–0.92)* | 0.88 (0.73–0.95) 0.24 (–0.70 to 0.71)* | 0.85 (0.66–0.95) 0.16 (–0.41 to 0.65)* |
| Q3: What type of study (design) would best answer your clinical question and why? | 0.94 (0.85–0.98) | 0.99 (0.98–0.99) | 0.97 (0.94–0.99) |
| Q4: Describe the search strategy you might use in Medline topics, fields, rationale, and limits. | 0.93 (0.82–0.98) 0.34 (–0.45 to 0.80)* | 0.81 (0.62–0.92) 0.21 (–0.68 to 0.69)* | 0.84 (0.62–0.95) 0.35 (–0.26 to 0.76)* |
| Q5: What characteristics of a study determine if it is relevant? | 0.91 (0.75–0.98) 0.20 (–0.61 to 0.74)* | 0.93 (0.86–0.97) 0.28 (–0.26 to 0.68)* | 0.84 (0.61–0.95) 0.51 (–0.06 to 0.83)* |
| Q6: What characteristics of a study determine its validity? | 0.98 (0.95–0.99) | 0.95 (0.88–0.97) | 0.89 (0.74–0.96) |
| Q7: What characteristics of the study's findings determine its magnitude and significance? | 0.98 (0.95–0.99) | 0.93 (0.87–0.97) 0.54 (0.11–0.81)* | 0.84 (0.63–0.95) 0.48 (–0.19 to 0.83)* |
| Total score | 0.93 (0.82–0.98) 0.83 (0.55–0.95)* | 0.83 (0.62–0.93) | 0.92 (0.82–0.98) 0.78 (0.42–0.92)* |

Note: *Reliability coefficients prior to further training and discussion among raters.

Abbreviations: ICC, intraclass correlation coefficient; CI, confidence interval.

previously reported validity and reliability of the original Fresno test¹¹ and AFT versions for rehabilitation professionals (ie, occupational therapist and physiotherapist).¹²

The importance of EBP training in facilitating an evidence-based approach to clinical practice has been highlighted by a number of systematic reviews.^{18–21} Many of the training programs reported in these reviews relied on self-report data, which potentially reflect inaccuracies in actual knowledge.²² Measuring the effectiveness of such training programs therefore requires objective and robust instruments to document changes in the competence of the individuals being trained. To the authors' knowledge, the AFT is the only objective measure of EBP knowledge and skills that has been tested and applied in allied health. McCluskey and Bishop, who first reported about the validity and reliability of the AFT, urged researchers to develop new clinical scenarios and modify the instrument to suit other health disciplines.¹² The current study addressed this gap and provided researchers and educators an instrument to measure EBP skills and knowledge in speech pathologists, social workers, and dieticians/nutritionists. The new versions of the AFT were content-validated, and although the internal consistency of the different versions was slightly lower than the original AFT, the Cronbach's α -values were still acceptable.

The reliability estimates for some of the items (questions 2, 4, 5, and 7) were questionable; however, after further training, the ICCs increased considerably, indicating moderate-to-excellent reliability of scores for these items. This finding highlights the importance of providing train-

ing to raters as a strategy to improve interrater reliability. Rater training has been shown to increase consistency of scoring between raters.²³ It emphasizes developing a common understanding among raters so they will apply the rating system as consistently as possible.²⁴ This common understanding, also called "frame of reference", addresses the common sources of rater disagreements, which include lack of overlap among what is observed, discrepant interpretations of descriptor meanings, and personal beliefs or biases.²⁴ However, research also suggests that even comprehensive training will not ensure rater agreement.²⁵ Studies have suggested that a rater's expertise may improve accuracy,^{23,26} which implies that rater characteristics are also an important consideration in ensuring consistency between raters. Reliability in examination scoring can be expected if the raters are highly knowledgeable in the domain in which ratings are made. Studies have found a relationship between rater expertise and rating accuracy, as well as the ability to differentiate between different domains in a rating scale.^{24,26} The raters involved in this study are experienced EBP educators and researchers, and these attributes could have contributed to the consistency in scoring. Because of their exposure to teaching, the raters may have already gained a wealth of experience in examination assessment, and could be expected to respond well to training. It is therefore not surprising to find that following training in AFT rating, the reliability estimates improved significantly for the previously questionable items. Based on the results of the current study, it appears that there are three important variables that can contribute to rater reliability: an

explicit scoring criteria (ie, scoring rubric), raters' training, and raters' professional experience.

As with any study, this research has limitations that need to be considered when interpreting the results. First, the sample size may have been too small to produce sufficiently reliable results. Second, the expert panel was limited to four practitioners, which may not represent the collective set of views in the different professions. Third, the ability of the test to detect change following educational programs has not been tested.

Despite these limitations, the results of this study provide a valuable resource for EBP educators and researchers who require an objective instrument to measure knowledge and skills among social workers, speech pathologists, and dietitians/nutritionists.

Conclusion

The authors propose the use of AFT in evaluating the EBP knowledge and skills of social workers, speech pathologists, and dietitians/nutritionists. EBP educators and researchers should identify raters with experience in EBP teaching or those with previous EBP training, who should then receive training for AFT scoring. The reliability of raters should be evaluated before they participate in the actual assessment.

While the content validity, internal consistency, and reliability of the AFT have been shown in this study, further research is required to establish its sensitivity to detect change in knowledge and skills following an educational intervention for dietitians, speech pathologists, and social workers.

Disclosure

The authors report no conflicts of interest in this work.

References

- Menon A, Korner-Bitensky N, Kastner M, McKibbin KA, Straus S. Strategies for rehabilitation professionals to move evidence-based knowledge into practice: a systematic review. *J Rehabil Med*. 2009;41(13):1024–1032.
- Heiwe S, Kajermo KN, Tyni-Lenné R, et al. Evidence-based practice: attitudes, knowledge and behaviour among allied health care professionals. *Int J Qual Health Care*. 2011;23(2):198–209.
- Bahtsevani C, Udén G, Willman A. Outcomes of evidence-based clinical practice guidelines: a systematic review. *Int J Technol Assess Health Car*. 2004;20(4):427–433.
- Alberts M, Easton D. Stroke best practices: a team approach to evidence-based care. *J Natl Med Assoc*. 2004;96 Suppl 4:5S–20S.
- Leufer T, Cleary-Holdforth J. Evidence-based practice: improving patient outcomes. *Nurs Stand*. 2009;23(32):35–39.
- Metcalfe C, Lewin R, Wisher S, Perry S, Bannigan K, Moffett JK. Barriers to implementing the evidence base in four NHS therapies: dietitians, occupational therapists, physiotherapists, speech and language therapists. *Physiotherapy*. 2011;87(8):433–441.
- Jette DU, Bacon K, Batty C, et al. Evidence-based practice: beliefs, attitudes, knowledge, and behaviors of physical therapists. *Phys Ther*. 2003;83(9):786–805.
- Iles R, Davidson M. Evidence based practice: a survey of physiotherapists' current practice. *Physiother Res Int*. 2006;11(2):93–103.
- Hockenberry M, Brown T, Walden M, Barrera P. Teaching evidence-based practice skills in a hospital. *J Contin Educ Nurs*. 2009;40(1):28–32.
- Ilic D. Assessing competency in evidence based practice: strengths and limitations of current tools in practice. *BMC Med Educ*. 2009;9:53.
- Ramos KD, Schafer S, Tracz SM. Validation of Fresno test of competence in evidence based medicine. *BMJ*. 2003;326(7384):319–321.
- McCluskey A, Bishop B. The Adapted Fresno Test of competence in evidence-based practice. *J Contin Educ Health Prof*. 2009;29(2):119–126.
- Mastaglia B, Toye C, Kristjanson L. Ensuring content validity in instrument development: challenges and innovative approaches. *Contemp Nurse*. 2003;14(3):281–291.
- Lizarondo LM, Grimmer-Somers K, Kumar S, Crockett A. Does journal club membership improve evidence uptake in different allied health disciplines: a pre-post study. *BMC Res Notes*. 2012;5:588.
- Amelang A. *Inter-rater Reliability of the Clinical Practice Assessment System used to Evaluate Pre-service Teachers at Brigham Young University* [master's thesis]. Provo (UT): Brigham Young University; 2009.
- Richman J, Makrides L, Prince B. Research methodology and applied statistics. *Physiother Can*. 1980;32(4):253–257.
- McCrae RR, Kurtz JE, Yamagata S, Terracciano A. Internal consistency, retest reliability, and their implications for personality scale validity. *Pers Soc Psychol Rev*. 2011;15(1):28–50.
- Green ML. Graduate medical education training in clinical epidemiology, critical appraisal, and evidence-based medicine: a critical review of curricula. *Acad Med*. 1999;74(6):686–694.
- Parkes J, Hyde C, Deeks J, Milne R. Teaching critical appraisal skills in health care settings. *Cochrane Database Syst Rev*. 2001;(3):CD001270.
- Taylor RS, Reeves BC, Ewings PE, Taylor RJ. Critical appraisal skills training for health care professionals: a randomized controlled trial [ISRCTN46272378]. *BMC Med Educ*. 2004;4(1):30.
- Coomarasamy A, Khan K. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. *BMJ*. 2004;329(7473):1017.
- Shaneyfelt T, Baum K, Bell D, et al. Instruments for evaluating education in evidence-based practice: a systematic review. *JAMA*. 2006;296(9):1116–1127.
- Barrett S. The impact of training on rater variability. *Int Educ J*. 2001;2(1):49–58.
- Graham M, Milanowski A, Miller J. *Measuring and Promoting Inter-rater Agreement of Teacher and Principal Performance Ratings*. Madison (WI): Center for Educator Compensation Reform; 2012. Available from: http://cecr.ed.gov/pdfs/Inter_Rater.pdf. Accessed June 13, 2013.
- Hoyt WT, Kerns M. Magnitude and moderators of bias in observer ratings: a meta-analysis. *Psychol Methods*. 1999;4(4):403–424.
- Lizarondo L, Grimmer K, Kumar S. Inter-rater reliability of Adapted Fresno Test across multiple raters. *Physiother Can*. 2012;65(2):135–140.

Journal of Multidisciplinary Healthcare

Dovepress

Publish your work in this journal

The Journal of Multidisciplinary Healthcare is an international, peer-reviewed open-access journal that aims to represent and publish research in healthcare areas delivered by practitioners of different disciplines. This includes studies and reviews conducted by multidisciplinary teams as well as research which evaluates the results or conduct of such teams or

healthcare processes in general. The journal covers a wide range of areas and welcomes submission from practitioners at all levels, from all over the world. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/journal-of-multidisciplinary-healthcare-journal>