

RESEARCH

Open Access



Applying modern measurement approaches to constructs relevant to evidence-based practice among Canadian physical and occupational therapists

Fadi Al Zoubi^{1,2*} , Nancy Mayo^{1,3}, Annie Rochette^{2,4} and Aliko Thomas^{1,2,5}

Abstract

Background: Evidence-based practice (EBP) is a complex process. To quantify it, one has to also consider individual and contextual factors using multiple measures. Modern measurement approaches are available to optimize the measurement of complex constructs. This study aimed to develop a robust measurement approach for constructs around EBP including practice, individual (e.g. knowledge, attitudes, confidence, behaviours), and contextual factors (e.g. resources).

Methods: One hundred eighty-one items arising from 5 validated EBP measures were subjected to an item analysis. Nominal group technique was used to arrive at a consensus about the content relevance of each item. Baseline questionnaire responses from a longitudinal study of the evolution of EBP in 128 new graduates of Canadian physical and occupational therapy programmes were analysed. Principles of Rasch Measurement Theory were applied to identify challenges with threshold ordering, item and person fit to the Rasch model, unidimensionality, local independence, and differential item functioning (DIF).

Results: The nominal group technique identified 70/181 items, and modified Delphi approach identified 68 items that fit a formative model (2 related EBP domains: self-use of EBP (9 items) and EBP activities (7 items)) or a reflective model (4 related EBP domains: attitudes towards EBP (17 items), self-efficacy (9 items), knowledge (11 items) and resources (15 items)). Rasch analysis provided a single score for reflective construct. Among attitudes items, 65% (11/17) fit the Rasch model, item difficulties ranged from -7.51 to logits (least difficult) to $+5.04$ logits (most difficult), and person separation index (PSI) = 0.63. Among self-efficacy items, 89% (8/9) fit the Rasch model, item difficulties ranged from -3.70 to $+4.91$, and PSI = 0.80. Among knowledge items, 82% (9/11) fit the Rasch model, item difficulties ranged from -7.85 to 4.50 , and PSI = 0.81. Among resources items, 87% (13/15) fit the Rasch model, item difficulties ranged from -3.38 to 2.86 , and PSI = 0.86. DIF occurred in 2 constructs: attitudes (1 by profession and 2 by language) and knowledge (1 by language and 2 by profession) arising from poor wording in the original version leading to poor translation.

Conclusions: Rasch Measurement Theory was applied to develop a valid and reliable measure of EBP. Further modifications to the items can be done for subsequent waves of the survey.

Keywords: Rasch Measurement Theory, Evidence-based practice, Self-efficacy, Attitudes, Knowledge, Resources, Measure, Measurement, Rehabilitation, Physical therapy, Occupational therapy

* Correspondence: fadi.alzoubi@mail.mcgill.ca

¹McGill University, Montreal, QC, Canada

²Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Montreal, QC, Canada

Full list of author information is available at the end of the article



Background

Health care professionals are expected to integrate best available research evidence, patients' preferences, and their clinical expertise to support clinical decision-making, a process known as evidence-based practice (EBP) [1]. Production of high-quality research in fields related to rehabilitation over the past 15 years [2] has provided evidence for occupational therapists (OTs) and physical therapists (PTs) to guide practice [3–5]. As a result of this exponential rise in knowledge, there is an urgent need to mobilize evidence into clinical practice [6]. According to the World Health Organization (WHO), rehabilitation is defined as “a set of interventions designed to optimize functioning and reduce disability in individuals with health conditions in interaction with their environment” [7]. Without implementing effective interventions, rehabilitation will not be successful [8–11]. As a response to these expectations [12, 13], a priority for all professional OT and PT programmes was to emphasize EBP. Despite the emphasis on teaching and

promoting the competencies associated with EBP in professional programmes, measuring EBP remains a daunting challenge.

EBP is a complex area of enquiry necessitating multiple measurement approaches to identify the practice itself and the individual and contextual factors influencing it. Several socio-cognitive theories have been applied in research to identify and tackle these factors [14–16]. To determine if EBP is changing professional practices, improving the quality of care, and informing organizations [17], there is a need to administer several multi-item questionnaires covering the relevant domains [18]. The use of validated and reliable measures of EBP outcomes is essential to improve EBP across studies and to inform areas for contextual change [17].

Measurement challenges

There are over 100 tools to measure domains related to EBP [19–22]. As a result, many items cover the

Table 1 Explanation of steps taken to fit the data to the Rasch model

Threshold order	There should be a logical ordering to the response options such that endorsing a more optimal response option should situate the person at a higher level of the latent trait. That means a person with higher ability (for example a knowledgeable clinician in EBP) is expected to select higher response options on an ordinal scale. At lower ability, more clinicians should endorse a lower response level, and fewer should endorse a higher response level. If the thresholds are disordered, the response options need to be rescored, sometimes reducing the responses to binary. The number of thresholds is equal to the number of response options - 1 and reflects the number of “jumps” the person has to make for each item.
Fit to the Rasch model	The items should line up hierarchically such that those items that need little ability to endorse at the most optimal response level are at the low end and those items requiring more ability to endorse are higher. Overall goodness of model fit is indicated by a non-significant chi-square test ($p > 0.05$) after a Bonferroni adjustment for the number of items. The fit of each item and each person is as important, or even more important, than overall fit. Item and person fit is indicated when fit residual (deviance from pure linearity) values are within ± 2.5 and the chi-square test for fit is non-significant (> 0.05). Those items that fail this criterion need to be looked at carefully to ensure their importance in scoring the latent trait. A fit residual of greater than $+2.5$ indicates the item does not fit the latent trait; a fit residual of less than -2.5 indicates the item overfits and may be redundant.
Unidimensionality	A requirement of the Rasch model is that a single latent trait is being measured. This is assessed using a principal component analysis (PCA) of the fit residuals. The person-ability estimates derived from all pair-wise comparisons of the two most disparate set of items (those with the highest positive and negative loadings on the first factor) are compared using independent t tests. For a set of items to be considered unidimensional, less than 5% of t values should be outside ± 1.96 . When this value is greater than 5%, a binomial test of proportions is used to calculate the 95% confidence interval (CI) around the t test estimate. Evidence of unidimensionality is still supported if the 5% value falls within the 95%CI.
Response dependency	The uniqueness of the information provided by the items is a requirement of the Rasch model. Items with pair-wise residual (after controlling for the latent trait) correlations greater than 0.3 could indicate lack of independence of the responses which inflates the reliability. Solutions include creating a super-item which combines the response options across items or choosing the one item that best suits the testing context.
Differential item functioning (DIF)	The items should have the same ordering of difficulty across all people being measured defined by personal factors such as in this study, PT or OT, gender, and language. DIF is an indicator of item bias. Typically, DIF is indicated with a significant F test from a two-way analysis of variance. A caution is that with large and sample sizes, anything may be significant; with small sample sizes, nothing may be significant. A close visual inspection of the item characteristic curve plotted by the level of each factor will support or not the information from the statistical approach. Two options are available for items with DIF, deletion or split scoring.
Targeting	An ideally targeted measure should include a set of items that spans the full range of the theoretical latent construct (-4 to $+4$ logits) and have a mean location of 0 with a standard deviation (SD) of 1. Ideally, the person estimates from this measure should be centred on location 0 with a SD of 1.
Discrimination or person separation	This indicates how well people are differentiated by the spread of the item difficulty. The person separation index (PSI) is interpreted like a Cronbach's alpha. The larger the index, the better is the discrimination which facilitates the measurement of change. Values of > 0.9 are suitable for measuring within-person change; values > 0.7 are suitable for detecting group differences.

same construct but with different phrasing and response options. There is always a need for parsimony in measurement, as redundancy can be a reason for non-completion and can produce false reliability [23]. The removal of redundant items from the questionnaire can raise concerns for fear of invalidating the interpretation of the total score derived from the original set of items. However, there is a vast literature on the validity of total scores derived from ordinal measures [24]. Typically, summing the numerical labels assigned to each ordinal response option is done to produce a total score. This score does not necessarily have mathematical properties nor does every item contribute equally to the total. Application of Rasch Measurement Theory (RMT) can shed light on how individual items contribute to a theoretically defined construct used to form a measure [25]. The basis of RMT is the Rasch model, named for Georg Rasch the Danish statistician, which is a probabilistic model used to situate a person and an item on a

linear continuum from least able (easiest item) to most able (hardest item) [25].

RMT is one of the several measurement theories that have been applied in the context of health care. The best-known theory is Classical Test Theory which assumes that the true score is a function of the total score plus error; the error is assumed to be the same for each person and for each item [26]. Theories that are based on how each item behaves with respect to other items in the theoretical construct and how the items align along an expected linear hierarchy, from easiest to hardest, do not assume that these errors are the same. Therefore, the location and error for each item and each person are estimated. The location of each item along this “ability” continuum is estimated by a logit transformation of participants’ responses to each level of each item. An item response category that 50% of participants endorse, the middle item, has a logit of 0. The optimal linear scale for the items is required to have a mean of 0 and a standard deviation of 1, with locations all along the continuum

Table 2 Measures—description and psychometric properties

Original measure	Description	New measures	Items
EBPQ ² [39] Practice subscale	74 items; 5-point Likert Scale Time frame: past 6 months Psychometric data: acceptable internal consistency (Cronbach’s alpha (α) 0.85), test-retest reliability (intraclass correlation coefficient (ICC) 0.83 and convergent validity 0.66 [40].	Self-use of EBP: It was a term we chose to reflect actual application of EBP concepts, tools, and procedures into specific actions such as identifying knowledge related to a patient situation or the ability to formulate a research question to guide a literature search based on this gap. EBP activities: It can be defined as the implementation of research evidence to the surrounding environment such as in/formally shared and discussed literature/research findings with colleagues at work or patients.	9 items on a 5-point scale ranging from “never” to “more than 10 times a month” 7 items on a 5-point scale ranging from “never” to “daily”
EBPAS [41]	50 items; 5-point ordinal scale ⁵ Psychometric data: acceptable good internal consistency (Cronbach’s α 0.90–0.93), good reliability (ICC 0.83), and convergent validity 0.66 [41–43].	Attitudes towards EBP	17 items on a 5-point scale ranging from “strongly disagree” to “strongly agree”
EPIC [44]	11 items; confidence from 0 to 100% Psychometric data: validated for both PTs and OTs. It has almost excellent internal consistency (Cronbach’s α 0.89) [44]. The ICC for test-retest reliability was 0.89 for PTs [45] and 0.92 for OTs [46].	EBP self-efficacy	9 items on a scale from 0 to 10 representing 0–100%
EBPQ [40]	24 items; 7-point Likert Scale Psychometric data: acceptable internal consistency (Cronbach’s α 0.91) and excellent test-retest reliability (ICC 0.94) [47].	Knowledge of EBP	11 items on a 5-point scale ranging from “never heard the term” to “understand and could explain to others”
ACT [36]	57 items; 5-point Likert Scale Psychometric data: acceptable internal consistency (Cronbach’s α > 0.80) [48].	EBP resources: It can be defined as the available resources at the workplace that allow clinicians to access and use EBP or encourage the clinicians to use EBP such as receiving recognition from manager/supervisor and workplace/college support the best practice.	17 items on a 5-point scale ranging from “strongly disagree” to “strongly agree”

EBPQ Evidence-Based Practice Profile Questionnaire, EBPQ² Evidence-Based Practice Profile Questionnaire-2, EBPAS Evidence-Based Practice Attitude Scale, EPIC Evidence-Based Practice Confidence Scale, ACT Alberta Context Tool

⁵We used only 15 items since some of these items are repetitive in other measures

from -4 to $+4$ logits. This represents the theoretical range of a standard normal distribution with a mean of 0 and a standard deviation of 1. The person's values along this continuum are also optimal when they follow this standard normal distribution.

RMT has a number of requirements including measuring only one construct (i.e. unidimensionality). The analytical details of how to apply and interpret Rasch analysis to a set of person responses to items are presented in Table 1.

Items that do not fit the Rasch model should not be used in the total score until improved. Only few EBP measures [27–32] have used Rasch analysis for developing the scoring system, and none of these items were designed for rehabilitation EBP. Moreover, none of these studies were designed for Canadian professionals where the items have to be both in English and French.

Study context

In the context of a study on the evolution of individual characteristics (including knowledge, attitudes, confidence), contextual factors (including support for work setting), and actual use of EBP in graduates of the professional M.Sc. programme in PT and OT (grant number 148544) [33], 5 existing measures were assembled to tap into these important EBP constructs. This resulted in 181 items which clearly would be a barrier to study recruitment and completion. In addition, many of the items were redundant leading to low efficiency [34] or had multiple concepts in one item leading to high content density [35]. An example of item redundancy is the “feedback process” subcategory of the Alberta Context Tool [36] that includes 5 items that can be reflected in one. These items (*I routinely receive information on my team's performance on data like the examples provided above, our team routinely discuss this data informally, our team has a scheduled formal process for discussing this data, our team routinely formulates action plans based on the data, our team routinely compares our performance with others*) can be captured in one item (*our team routinely monitors our performance with respect to the action plans*). Another example of multiple concepts in one item exists in the subcategory “sympathy” in the Evidence-Based Practice Profile Questionnaire-2 which is *Critical appraisal of the literature and its relevance to the client is not very practical in the real world of my profession*. This item fits in both the “sympathy” and “relevance” subcategories. These shortfalls in item development can lead to inconsistency in responses and biased estimates of change. From a measurement perspective, there is value in the items from multiple questionnaires as they could be considered to form a pool of items from which new combinations could be constructed with legitimate total scores.

Therefore, the global aim of this study was to apply a robust measurement approach on constructs around EBP including practice, individual characteristics variants (e.g. knowledge, attitudes, confidence, behaviours), and contextual factors (e.g. resources). The specific study objectives were to (i) identify the extent to which items from existing EBP measures reflect constructs suitable for use in a survey of EBP in PTs and OTs and (ii) the extent to which the items reflective of EBP constructs fit the expected unidimensional hierarchy sufficient to create a total score with interval-like properties optimizing the estimation of change or differences across groups.

Methods

Study design

This study was based on three steps: a nominal group process [37], a modified Delphi approach [38], and a cross-sectional electronic survey. An item analysis was conducted on all items arising from the five EBP measures chosen for inclusion in the longitudinal study of the evolution of EBP in new graduates of Canadian PT and OT programmes mentioned above. Data from the newest PT/OT cohort to enter clinical practice were analysed to refine the measurement strategy for future phases.

Population

The target population for the cross-sectional survey was all graduates ($n = 1703$) of the 28 Canadian OT and PT programmes that completed their professional education during the 2016–2017 academic year. To ensure complete ascertainment of graduates, participants were identified from university academic programmes. For the item analysis, the newest entry cohort was queried within 1-month post-graduation. The new graduates were prioritized for recruitment in order to assess EBP-related constructs at entry to practice and then subsequently over time. At the time of analysis, data from new graduates from 12 of the planned 28 university programmes were available.

Measurement

To identify potential measures for the broader study, all relevant EBP measures reported in the literature were identified. Table 2 lists the 5 measures used in this study from which the 181 items were chosen by the research team as targeting the constructs of self-use of EBP and EBP activities, individual factors (attitudes towards EBP, confidence in applying EBP, knowledge), and contextual factors (specifically resources).

Procedures

The procedures for appraising the 181 items and identifying potential constructs for further analyses are shown

in Fig. 1. We used a three-phase process with two groups of experts (potential future respondents (user panel) and methodological (expert panel)) for the item analysis and then translated the results to French. The user panel comprised 12 PTs and OTs all with experience in clinical practice and training in EBP. The participants had a range of experience that covered the scope of practice of PT and OT including recent graduates and experienced clinician-researchers. The expert panel comprised the core research team members (AT, NEM, FAZ, AR).

Phase 1: Nominal group process

A nominal group process [36], involving both the user and expert panels, was used to screen each item by applying three criteria. Contextual relevance was defined based on whether the item was judged (by 80% or more of the user panel) to be relevant to the EBP and research/clinical scenarios facing PTs and/or OTs. Redundancy was to be avoided, and so, items were excluded if an item had a similar meaning to another item. In case of repetition, the item judged by > 50% of the user panel to have the clearest wording was retained. Re-wording was needed for items created for other health care providers such as doctors or nurses but relevant for EBP in our context. The expert panel suggested alternatives for re-wording the item, and the final decision was made by consensus (see the Additional file 1). This process reduced the initial 181 items to a pool of 81 items.

A modified Delphi process involving all the investigators on the research team ($n = 8$) was used to refine the items.

First round An invitation e-mail was sent enclosing an explanation of the aim of the work and a description of the previous steps that had led to the generation of the item pool. Once the invitee accepted to participate, the item pool was sent by e-mail with specific instructions. Each investigator was asked to vote by marking on an Excel sheet an X under the category Item Clarity (completely clear or completely unclear) and under the category Informative (highly informative, moderately informative, not informative, not sure). Another category was created for additional comments. In this round, the goal was to clarify any redundancy or comprehension problems regarding each item [49]. Response frequencies for each item were calculated, and the experts' identity was anonymized by a research assistant. Each item required 80% agreement from the panel in order to keep or remove the item as suggested by Lynn [49]. If the item showed < 80% agreement and there were comments for re-wording, then the item was kept to be re-worded and re-administered in the next round. All the responses were assigned to one of the three categories: keep without changes, keep with re-wording, and remove because of redundancy. At the end of this round, the expert panel met and re-worded the items that fell under the category of "re-word". In case of any ambiguous

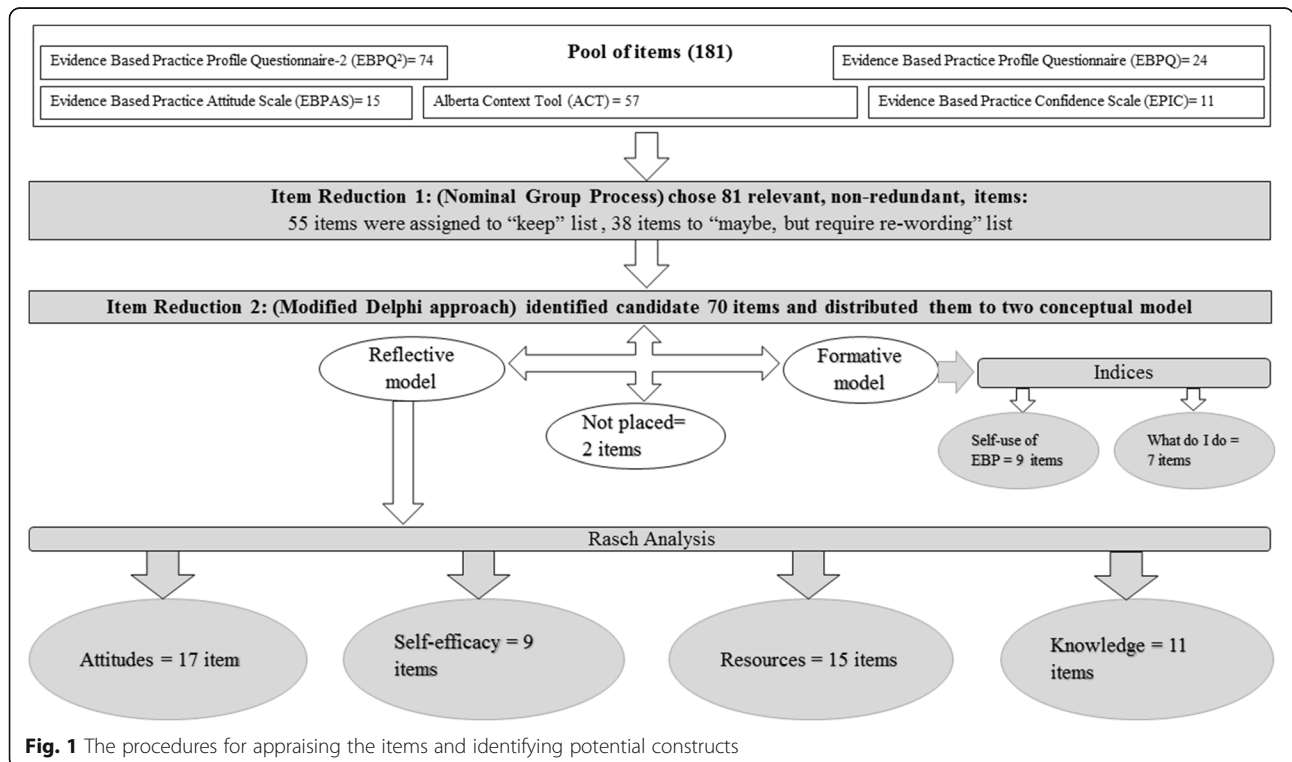


Fig. 1 The procedures for appraising the items and identifying potential constructs

comments, the investigator was contacted to seek clarifications for the comment. After reviewing all items, a final list was prepared for the second round.

Second round An e-mail was sent to the expert panel for a final review and feedback on the changes applied to the items using similar procedures for rating as described in the first round.

This process resulted in 70 endorsed items from the pool of 81 items.

Phase 2: Meeting with the core team

Once all comments were in, two meetings with the core research team were held to identify the ideal set of items that would be carried forward in subsequent phases. Four activities were carried out. The first was to confirm the constructs and their labels. Second, items were assigned to constructs or left unassigned. Third, each item was prioritized for inclusion according to its relevance to the construct and whether the item indicated a high or low degree of EBP. The fourth activity was to consider whether the constructs fit best with a formative or reflective conceptual model. Formative models are where the items form the construct rather than reflecting it. For example, a priori, self-use of EBP practices was considered formative as there was a list of recommended practices. The more practices, the higher the EBP use. The most valid method for creating a legitimate total score for formative constructs is to count the number of items at a particular level of expertise. The construct self-efficacy is a good example of a reflective model, as having more self-efficacy or confidence in applying EBP in practice is reflected in people endorsing confidence in certain behaviours chosen to reflect the construct. A necessary but not sufficient criterion for a reflective model is that the items fit the underlying measurement model, here the Rasch model.

Phase 3: Translation

The quality of translation followed the *Guidelines for the Process of Cross-Cultural Properties of Self-Report Measures* [50]. This guideline helps the translation process into a new language through the following six steps: (1) initial (forward) translation—two professional translators with French-first language as their mother tongue (one was informed about the project, and the other translator was novice to the area) independently translated the final list of items from English to Canadian French; (2) synthesis of the translation—the same two translators synthesized the results of the two translations by preparing a consensus translation in the Canadian French language with the help of a research assistant; (3) back translation—two professional translators with English-

first language as their mother tongue that are naive to measurement made back-translations; (4) use of an expert committee—in this step, a methodologist, PT, and OT professionals and a language professional met to come up with a clean version of the translation by solving any discrepancies through discussions until all the final items were judged to be linguistically equivalent; (5) testing the pre-final version—the clean version of the Canadian French set of items was completed by five clinicians and graduate students who were recent graduates (less than 5 years of clinical experience). This group identified some items to be hard to answer for recent graduates who have not worked in clinical practice. That led to suggest adding “My work does not involve clinical care” as a response option to all questions except those in knowledge or confidence domain; and (6) appraisal of the adaptation process—after adding “My work does not involve clinical care” as a response option, the final set of items was approved by the core team to be administered online.

Rasch analysis

Rasch measurement analysis was carried out to test the fit of the items related to EBP to the Rasch model. Rasch model was considered fit by item/person if the observed item/person perform consistently like the expected item/person performance. This can be quantified using chi-square (χ^2) probability value if the value is > 0.05 with a Bonferroni adjustment and item/person fit residuals (sum of person and item deviations) that were close to 0 with a standard deviation of 1. Item residual correlation matrix was examined for possible local item independence, and unidimensionality was explored using principal component analysis of the residuals. Individual item fit was evaluated using the χ^2 probability value and the fit residual values. If χ^2 probability value of < 0.05 (Bonferroni adjustment) and fit residual values ≥ 2.5 , then the individual item was considered misfit [51, 52]. The person separation index (PSI) was used to assess the internal consistency

Table 3 Characteristics of graduates (n = 128)

Variable	Mean \pm SD or N (%)
Age (years)	27.3 (6.9)
Gender	
Men/women/not answered	19/105/14 (14.9/82.9/3.1)
Language	
English/French	70/65 (51.9/48.1)
Degree	
PT/OT	53/75 (41.4/58.6)
Currently working	
Yes/no	73/55 (57.0/43.0)

reliability of the scale which is equivalent to Cronbach's α [47]; however, it only uses the logit values instead of the raw scores [52], where a value of ≥ 0.7 was considered acceptable representing the minimum required to divide the participants into two distinct groups (low/high ability) [53, 54]. Differential item functioning (DIF) was assessed to identify items that work differently for some groups who have the same level of ability in the sample. DIF was tested by profession, gender, language, and the type of clinical setting. The items which did not fit theoretically or mathematically

the Rasch model were removed. Items that showed DIF were either deleted or split. This step was repeated multiple times until all items fit the model and the measure was formed.

All Rasch measurement analyses were performed using the Rasch Unidimensional Measurement Model Software (RUMM) version 2030 [55]. All descriptive analyses were performed using the SAS statistical software (version 9.4) [56]. Data were reported as means \pm standard deviation (SD) or as frequencies (percentages).

Table 4 Results of analysis for self-use of EBP

Instructions: For each of the following activities, how often have you done the following in the past 6 months?: 5-point scale
(Directives: Depuis 6 mois, à quelle fréquence avez-vous...)

Item #	Description of item	Never	1 to 2 times	Almost every month	2 to 10 times a month	More than 10 times a month
1	E Identify a gap in your knowledge related to a patient or client situation (e.g. history, assessment, treatment)?	0	1	1	1	1
	F Cerner une lacune dans vos connaissances sur la situation d'un patient ou client (ex. antécédents, évaluation, traitement)?					
2	E Formulate a question to guide a literature search based on a gap in your knowledge?	0	1	1	1	1
	F Formuler une question pour orienter une recherche de la littérature fondée sur cette lacune dans vos connaissances?					
3	E Effectively conduct an online literature search to address the question?	0	1	1	1	1
	F Mener efficacement une recherche en ligne de la littérature pour tenter de répondre à mes questions?					
4	E Critically appraise the strengths and weaknesses of study methods (e.g. appropriateness of study design, recruitment, data collection, and analysis)?	0	1	1	1	1
	F Évaluer de manière critique les forces et faiblesses de certaines méthodes de recherche (ex. pertinence de la conception d'une étude, recrutement, collecte et analyse de données)?					
5	E Critically appraise the measurement properties (e.g. reliability and validity, sensitivity and specificity) of standardized tests or assessment tools you are considering using in your practice?	0	1	1	1	1
	F Évaluer de manière critique les caractéristiques de mesure (ex. fidélité et validité, sensibilité et spécificité) des tests normalisés ou des outils d'évaluation que vous pensez utiliser dans votre pratique?					
6	E Interpret study results obtained using statistical tests and procedures (e.g. <i>t</i> tests, logistic regression)?	0	1	1	1	1
	F Interpréter les résultats d'étude à l'aide d'outils et de procédures statistiques (ex. tests <i>t</i> , régression logistique)?					
7	E Determine if evidence from the research literature applies to your patient's/client's situation?	0	1	1	1	1
	F Déterminer si des preuves découlant d'une recherche de la littérature s'appliquent à la situation de votre patient ou client?					
8	E Decide on an appropriate course of action based on integrating the research evidence, clinical judgment, and patient or client preferences?	0	1	1	1	1
	F Décider d'un plan d'action approprié intégrant des données probantes, le jugement clinique et les préférences du client ou patient?					
9	E Continually evaluate the effect of your course of action on your patient's/client's outcomes?	0	1	1	1	1
	F Évaluer régulièrement les conséquences de votre plan d'action sur les résultats chez le patient ou client?					

0 = never; 1 = one time or more

E English, F French translation

Ethics

The research project from which the data were taken had ethical approval from all relevant university ethics committees to carry out the study.

Results

The characteristics of the participants are shown in Table 3. Figure 1 shows the process and results of the item analysis. The nominal group process resulted in a streamlined group of 81 items, from the original pool of 181 items: 55 which could be used without modification (“keep” list) and 38 that would need some re-wording. For the modified Delphi approach, all 8 investigators participated in all rounds. For the first round, experts provided their ratings for all the items. For the additional comments, redundant ones having similar suggestion were grouped and reduced to produce 57 comments which included suggestions to rephrase the item or make it clearer. There was an agreement among the experts to keep without changes (51), keep with re-wording (29), and remove because of redundancy (11). For both the item clarity and informative agreement, initially, 41 items were rated to be completely clear and highly informative (> 80% agreement) while 29 items were rated to be completely unclear or moderately/not informative (< 80% agreement). The items that

were rated to be completely clear were not necessarily rated as highly informative and the opposite way.

For the second round, items were either kept as they were, removed, or re-worded in light of the experts’ comments. In total, 70 items were examined by the expert panel where no more major changes were required.

In total, 2 items were found to not be related to any domain resulting in a total of 68 items distributed on 6 EBP domains that were covered by the items. Two sets were considered to fit a formative model: self-use of EBP (9 items) and EBP activities (7 items). For these measures, the best total score is derived by counting the number of uses or activities carried out per day over the specified time frame.

Self-use of EBP

For this construct, more frequent use of EBP does not necessarily translate to better EBP if every contact with a patient initiates an EBP activity. For example, identifying a gap in knowledge more than 10 times a month would seem problematic rather than desirable. For this measure, counting the number of practices used in the past 6 months would eliminate giving problematic behaviours more weight. A total score from 0 to 9 would now be the indicator for self-use of EBP (see Table 4).

Table 5 Results of analysis for the original items of the construct EBP activities

Instructions: In the past month, how often have you?: 5-point scale
(Directives: Depuis un mois, à quelle fréquence avez-vous...)

Item #	Description of item	Never	Monthly or less	Bi-weekly	Weekly	Daily
1	E Integrated research evidence with your expertise F Intégré des preuves découlant de recherches à votre expertise?	0	1	2	4	20
2	E Informally (e.g. outside of formal team or family meetings) shared and discussed literature/research findings with colleagues at work F Partagé et discuté de manière informelle (ex. hors du cadre de réunions d'équipe ou de famille structurées) de résultats publiés ou de recherches avec des collègues au travail?	0	1	2	4	20
3	E Formally (e.g. during team or family meetings) shared and discussed literature/research findings with colleagues at work F Partagé et discuté dans un cadre structuré de résultats publiés ou de recherches avec des collègues au travail?	0	1	2	4	20
4	E Shared and discussed literature/research findings with patients/clients F Partagé et discuté de résultats publiés ou de recherches avec des patients ou clients?	0	1	2	4	20
5	E Read published research reports F Lu des rapports de recherche publiés?	0	1	2	4	20
6	E Made time to read research F Réserve du temps à la lecture de travaux de recherche?	0	1	2	4	20
7	E Attended in-services/workshops/courses in your organization? F Assisté à des ateliers, séances de formation ou cours dans votre organisation?	0	1	2	4	20

E English, F French translation

Table 6 Results of Rasch Analysis for the Original Items of the Construct Attitudes towards EBP

Instructions: Please indicate your level of agreement with the following statements: 5-point Likert Scale
(Directives: Veuillez indiquer à quel point vous êtes en accord avec les énoncés suivants)

Item #	Description of Item	Response option rescored	Result		
			Item misfit	Local item dependency	DIF
1	E New evidence is so important that I make the time in my work schedule F (Les nouvelles données probantes sont tellement importantes que j'y consacre du temps dans mon horaire de travail	√	No	Yes	No
2	E My practice has changed because of evidence I have found F Ma pratique a changé en raison de données probantes que j'ai découvertes	√	No	Yes	Profession (Item 2 split)
3	E Evidence based practice is fundamental to my professional practice F La pratique fondée sur des données probantes est essentielle à l'exercice de ma profession	√	No	Yes with Item 5 (Item 3 deleted)	No
4	E I need to increase the use of evidence in my daily practice F Je dois augmenter le recours aux données probantes dans ma pratique		No	No	No
5	E An evidence based practice approach improves the quality of my practice F Une pratique fondée sur des données probantes améliore la qualité de mon travail professionnel	√	No	No	Language (Item 5 deleted)
6	E Literature and research findings are useful in my daily practice F Les données tirées de la littérature et de la recherche sont utiles dans ma pratique de tous les jours	√	No	Yes	No
7	E Evidence based practice helps me to make decisions about patients/clients in my practice F Le recours à des données probantes m'aide à prendre des décisions au sujet des patients ou clients de ma pratique	√	No	Yes with Item 2 and 6 (Item 7 deleted)	No
8	E I am willing to use new and different types of clinical interventions (e.g. assessment, treatment) developed by researchers to help my patients/clients F J'accepterais de bon gré d'utiliser divers types d'interventions cliniques inédites (ex. évaluation, traitement) mises au point par des chercheurs pour aider mes patients ou clients	√	No	Yes	No
9	E I would try a new therapy/intervention even if it were very different from what I am used to doing F J'essaierais un nouveau traitement ou une nouvelle approche même si elle était très différente de ce que je fais d'habitude	√	No	Yes with Item 8 (Item 9 deleted)	No
10	E I resent having my clinical practice questioned F Je n'aime pas qu'on mette en doute ma pratique clinique	√	No	No	No
11	E I stick to tried and trusted methods in my practice rather than changing to anything new F Dans ma pratique, j'adopte des méthodes fiables et éprouvées plutôt que de changer et essayer une nouvelle approche	√	No	No	Profession and language (Item 11 deleted)
12	E Clinical experience is the most reliable way to know what really works F L'expérience clinique est la meilleure façon de voir ce qui fonctionne vraiment	√	No	Yes with Item 13 (Item 12 deleted)	No
13	E Clinical experience is more useful than scientific studies when I make decisions about my patients/clients F L'expérience clinique est plus utile que les études scientifiques au moment de prendre des décisions au sujet de mes patients ou clients	√	No	Yes	No
14	E Critical appraisal of the literature is not very practical to do in my day-to-day practice F Faire une évaluation critique de la littérature n'est pas très pratique dans ma pratique au jour le jour	√	No	Yes with Item 15 (Item 14 deleted)	No

Table 6 Results of Rasch Analysis for the Original Items of the Construct Attitudes towards EBP (*Continued*)

Instructions: Please indicate your level of agreement with the following statements: 5-point Likert Scale
 (Directives: Veuillez indiquer à quel point vous êtes en accord avec les énoncés suivants)

Item #	Description of Item	Response option rescored	Result		
			Item misfit	Local item dependency	DIF
15	E Seeking relevant evidence from scientific studies is not very practical in the real world	√	No	Yes	No
	F Chercher des données probantes dans des études scientifiques n'est pas très pratique dans la réalité				
16	E I know better than academic researchers how to care for my patients/clients	√	No	No	No
	F Je sais mieux que des chercheurs théoriciens comment m'occuper de mes patients ou clients				
17	E Research based treatments/interventions are not clinically useful	√	No	No	No
	F Les traitements ou interventions découlant de la recherche ne sont pas utiles en clinique				

DIF differential item functioning, E English, F French translation

EBP activities

For EBP activities, more frequent use does indicate more EBP, and hence, a total number of activity days would be a reasonable metric. This requires assigning a frequency for each of the categories of never, monthly or less, bi-weekly, weekly, and daily over the past working month and giving the estimates of 0, 1, 2, 4, and 20 for these categories respectively. The total score is the cross-product of item days yielding activity days on a continuous scale (see Table 5).

Conceptual model

Four sets of items were considered to potentially fit a reflective conceptual model: attitudes towards EBP (*n* = 17 items), self-efficacy (*n* = 9 items), knowledge (*n* = 11 items), and resources (*n* = 15 items). Items that were compatible with a reflective model were tested to estimate the extent to which they fit the Rasch model, a necessary condition for a reflective model. Tables 4, 5, 6, and 7 and Figs. 2, 3, 4, and 5 present the results of the Rasch analysis for the constructs originally considered to be reflective.

Attitudes towards EBP items

The attitudes towards EBP items are shown in Table 6. Seventeen items were tested, all measured on a 5-point Likert Scale for the degree to which the respondent agrees or disagrees with the statement with an additional response options for “My work does not involve clinical care” in the event that the respondent does not provide hands-on direct care to patients but is involved in another aspects or rehabilitation practice (e.g. case management, clinical research). This latter option was merged with the “neutral” option. Nine items had disordered thresholds and needed rescoring. For 7 items, the categories “strongly disagree” and “disagree” were

collapsed (items 1, 2, 3, 5, 7, 10, 11); for 3 items, the final rescoring resulted in only 2 (binary) categories (items 10, 14, 15). All 17 items fit the Rasch model after rescoring. Five items showed dependency with at least 1 other item, and the best-worded item was kept resulting in the deletion of items 3, 7, 9, 12, and 14. Three items showed DIF: item 2 showed DIF by profession, and the item was split to allow 2 different scoring; item 5 showed DIF by language, and as it was very close in content to item 3, it was deleted; and item 11 (*I stick to tried and trusted methods in my practice rather than changing to anything new*) showed DIF by language and profession. On close inspection, the English wording was idiomatic, and the French translation did not reflect the right meaning of the idiom. Item 13 is close in content, and so, item 11 was deleted. The final set of 10 items best reflecting the attitudes the EBP all fit the Rasch model and formed a measure ($\chi^2 = 25.14$, *df* = 22, *p* = 0.29). Figure 2 shows the threshold map (part a) and the targeting map (part b). The threshold map lists the items according to their average location on the latent trait (*x*-axis). The thresholds, transitions from 1 category of response to the next, are illustrated in different colours. There is a gradient across the items in the threshold location such that the easiest item, item 17 (see Table 6), is situated at the lowest end of the latent trait (-7.51). The hardest item (item 2, split by profession) has the most difficult threshold at +5.04. Figure 2b shows the location of the people and the items along the latent trait (*x*-axis) with the people above the line in pink and the items threshold below the line in blue. The items are poorly targeted to the sample primarily because the sample scored above 0, the expected mean location (mean 1.38; SD 0.83), whereas the range of items is from ≈ -8.0 to +5.0.

Table 7 Results of Rasch analysis for the original items of the construct self-efficacy towards EBP

Instructions: Please indicate how confident you are in your current level of ability by choosing the corresponding number on the following rating scale: 11-point continuous scale
(Directives: Veuillez indiquer à quel point vous avez confiance en vos capacités actuelles en choisissant le nombre correspondant sur l'échelle d'appréciation suivante)

Item #	Description of item	Response option rescored	Result		
			Item misfit	Local item dependency	DIF
1	E Identify a gap in your knowledge related to a patient or client situation (e.g. history, assessment, treatment)?	✓	No	No	No
	F Cerner une lacune dans vos connaissances sur la situation d'un patient ou client (ex. antécédents, évaluation, traitement)?				
2	E Formulate a question to guide a literature search based on a gap in your knowledge?	✓	No	No	No
	F Formuler une question pour orienter une recherche de la littérature fondée sur cette lacune dans vos connaissances?				
3	E Effectively conduct an online literature search to address the question?	✓	No	No	No
	F Mener efficacement une recherche en ligne de la littérature pour tenter de répondre à mes questions?				
4	E Critically appraise the strengths and weaknesses of study methods (e.g. appropriateness of study design, recruitment, data collection, and analysis)?	✓	No	No	No
	F Évaluer de manière critique les forces et faiblesses de certaines méthodes de recherche (ex. pertinence de la conception d'une étude, recrutement, collecte et analyse de données)?				
5	E Critically appraise the measurement properties (e.g. reliability and validity, sensitivity and specificity) of standardized tests or assessment tools you are considering using in your practice?	✓	No	No	No
	F Évaluer de manière critique les caractéristiques de mesure (ex. fidélité et validité, sensibilité et spécificité) des tests normalisés ou des outils d'évaluation que vous pensez utiliser dans votre pratique?				
6	E Interpret study results obtained using statistical tests and procedures (e.g. <i>t</i> tests, logistic regression?)	✓	No	No	No
	F Interpréter les résultats d'étude à l'aide d'outils et de procédures statistiques (ex. tests <i>t</i> , régression logistique)?				
7	E Determine if evidence from the research literature applies to your patient's/client's situation?	✓	No	Yes	No
	F Déterminer si des preuves découlant d'une recherche de la littérature s'appliquent à la situation de votre patient ou client?				
8	E Decide on an appropriate course of action based on integrating the research evidence, clinical judgment, and patient or client preferences?	✓	No	Yes with items 7 and 9 (item 8 deleted)	No
	F Décider d'un plan d'action approprié intégrant des données probantes, le jugement clinique et les préférences du client ou patient?				
9	E Continually evaluate the effect of your course of action on your patient's/client's outcomes?	✓	No	Yes	No
	F Évaluer régulièrement les conséquences de votre plan d'action sur les résultats chez le patient ou client?				

DIF differential item functioning, E English, F French translation

Self-efficacy towards EBP

The self-efficacy towards EBP items are shown in Table 7 and Fig. 3. Nine items were tested; all measured on an 11-point scale from 0% (no confidence) to 100% (completely confident). All the items had disordered thresholds and needed rescoring. For eight items, the categories "90%" and "100%" were collapsed (items 1, 2, 4, 5, 6, 7, 8, 9); for one item, the final rescoring resulted in only two (binary) categories (item 3). For all items,

the first four categories (0%, 10%, 20%, 30%) were merged; for all but item 3, the middle categories (40%, 50%, 60%) were collapsed. All nine items fit the Rasch model after rescoring. Item 8 showed dependency with two other items (items 7 and 9), and the best-worded items were kept resulting in the deletion of item 8. None of the items showed DIF. The final set of eight items best reflecting the self-efficacy towards EBP all fit the Rasch model and formed a measure ($\chi^2 = 10.89$, $df = 16$,

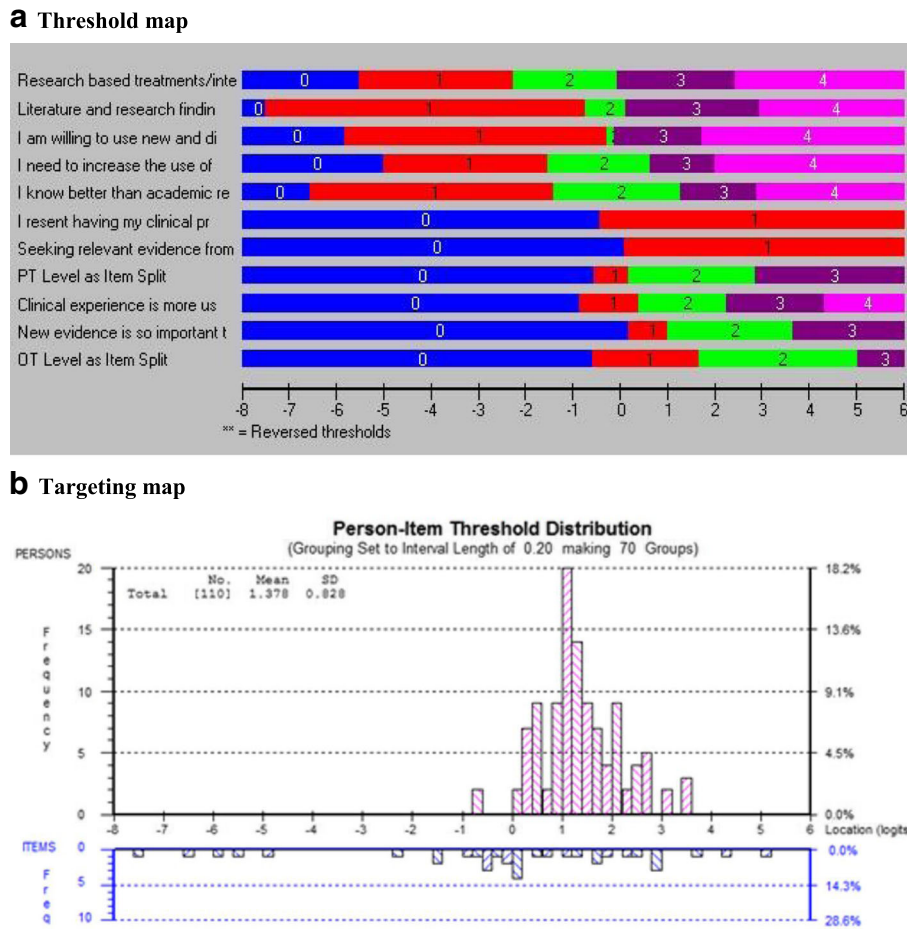


Fig. 2 a Threshold map and b targeting map of "attitudes" construct

$p = 0.82$). Figure 2a shows the gradient across item-thresholds, and Fig. 2b shows that the sample is reasonably well targeted by the items as the mean location of the sample is 0.52 (expectation 0) with SD of 1.46 (expectation 1.0).

Knowledge

The knowledge items are shown in Table 8. Eleven items were tested, all measured on a 5-point ordinal scale of declarative statements related to scientific terminology needed for EBP. Four items had disordered thresholds and needed rescoring. The categories "never heard the term" and "have heard it but don't understand" were collapsed for items 1, 7, 9, and 10. All 11 items fit the Rasch model after rescoring. Two items showed dependency with at least 1 other item, and the best-worded item was kept resulting in the deletion of items 2 and 11. Three items showed DIF: items 4 and 9 by profession and item 7 by language. In order to include all possible knowledge items, these items were split; however, the DIF by language for

item 7 was likely due to poor translation (*English: number-needed-to-treat (NNT); French: nombre de sujet à traiter*). The final set of 9 items best reflecting the knowledge all fit the Rasch model and formed a measure ($\chi^2 = 25.13$, $df = 18$, $p = 0.12$). The threshold map in Fig. 5a shows a hierarchy across knowledge items with the easiest item familiarity with reliability and the hardest item treatment effect size. Figure 5b shows that the people are not well targeted by the items (mean location 2.25, SD 1.45; expectation mean 0, SD 1).

Resources

The resources items are shown in Table 9. Fifteen items were tested, all measured on a 5-point Likert Scale for the degree to which the respondent agrees or disagrees with the statement with additional response options for "My work does not involve clinical care" as per the attitudes items. This latter option was merged with the "neutral" option. All 15 items had disordered thresholds and needed rescoring. The categories "strongly disagree"

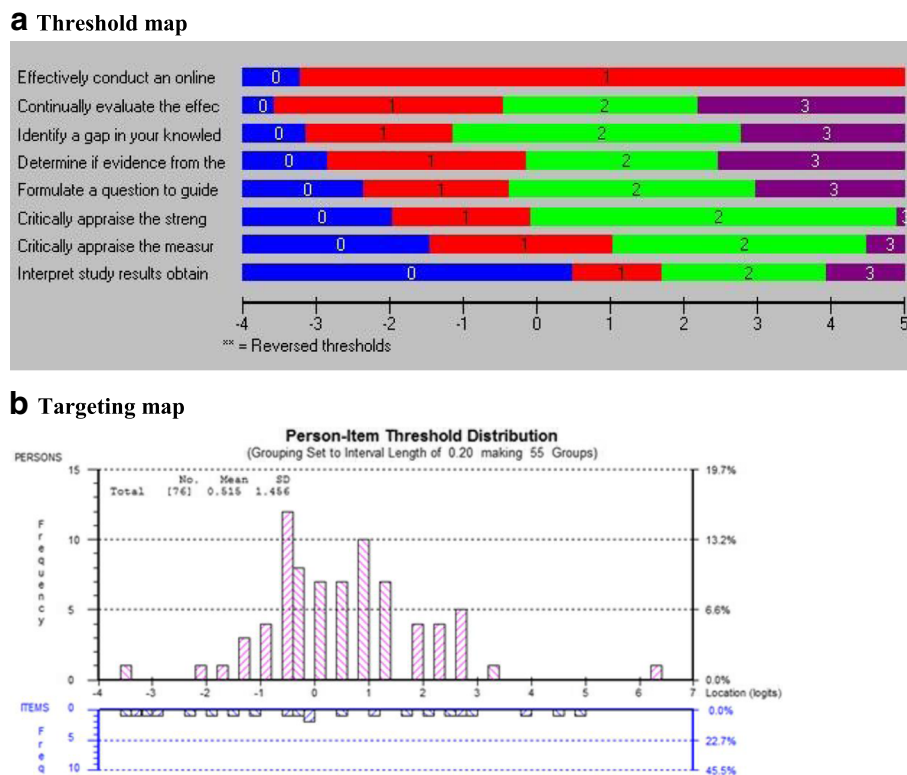


Fig. 3 a Threshold map and b targeting map of “self-efficacy towards EBP” construct

and “disagree” were collapsed for all items. After re-scoring, all but 1 item (item 12) fit the Rasch model and it was deleted. Only 1 item showed dependency on another item, and the best-worded item was kept resulting in the deletion of item 14. None of the items showed DIF. The final set of 13 items best reflecting the resources all fit the Rasch model, but the global fit was poor ($\chi^2 = 56.92$, $df = 26$, $p = 0.00$). Figure 3a shows the hierarchy across the items and Fig. 3b shows that the people are reasonably well targeted by the items, but there are many items at the low end of the resource continuum and no people.

Discussion

This study illustrated the psychometric steps needed to test the extent to which a set of items from existing measures of EBP form measures designed for PTs and OTs. The original 181 items obtained from 5 measures represented the core constructs of EBP. Theoretical and methodological processes were then applied for item reduction which resulted in 68 unique candidate items. Two constructs fit a formative conceptual model: self-use of EBP (9 items) and EBP activities (7 items). Four constructs fit a reflective conceptual model: attitudes towards EBP ($n = 17$

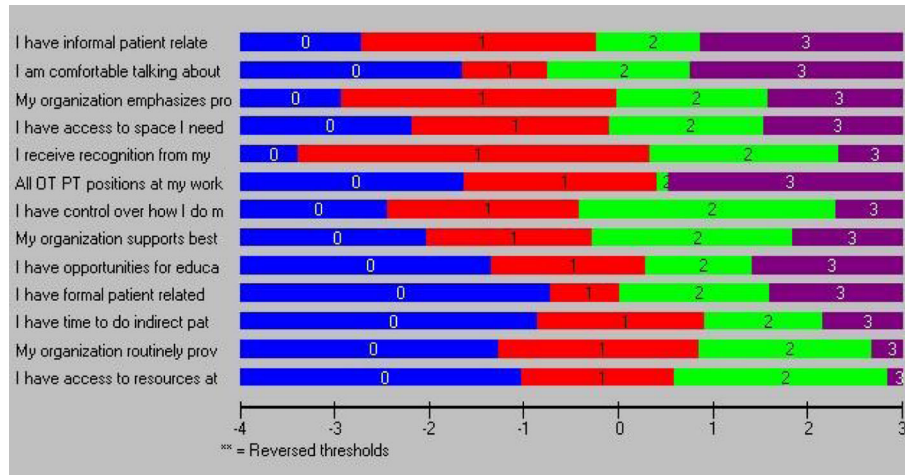
items), self-efficacy ($n = 9$ items), knowledge ($n = 11$ items), and resources ($n = 15$ items).

Rasch analysis showed that items reflecting four core constructs around EBP formed measures: attitudes towards EBP ($n = 10$ items), self-efficacy ($n = 8$ items), knowledge ($n = 9$ items), and resources ($n = 13$ items) (see Table 10). We discuss each one next.

Attitudes towards EBP

The items fit the Rasch model but were poorly targeted to our participants primarily because the mean fit for the participants (> 0) suggested that their overall scores on attitudes towards the EBP construct were greater than what would be expected in the general population. This was not surprising as our participants are recent graduates of rehabilitation programmes with a strong focus on EBP. This finding is congruent with previous studies showing recent graduates are more likely to report positive attitudes towards EBP than more seasoned clinicians and those with bachelor’s level training in OT and PT [57–59]. Scale reliability for this construct was unacceptable using person separation index (PSI) = 0.63, indicating that the items inadequately divided our participants along a continuum. RMT was applied to testing attitudes towards EBP in a previous study [27], and the results showed that 10 items fit the Rasch model with

a Threshold map



b Targeting map

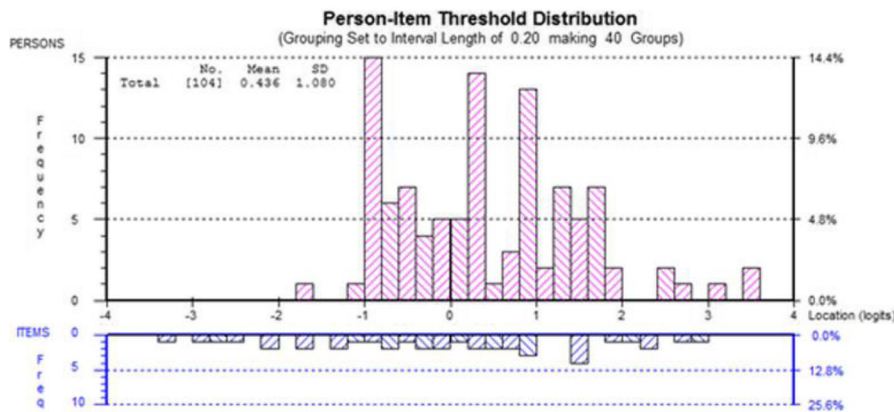


Fig. 4 a Threshold map and b targeting map of “resources” construct

good reliability (0.83); however, the authors did not test DIF as the sample size was small ($n = 110$).

Three attitudes items showed DIF: item 2 showed DIF by profession; PTs are more likely than OTs to change their practice because of the evidence they find. Previous research on the use of and attitudes towards EBP showed that OTs score significantly lower than other health care providers which might indicate that this professional group faces different challenges related to EBP compared to other health care providers [60, 61]. Another explanation might be that OT, as a profession, is considered heavily client-centred and as such may value patient preferences as a contributor to EBP more so than the level of evidence [62, 63].

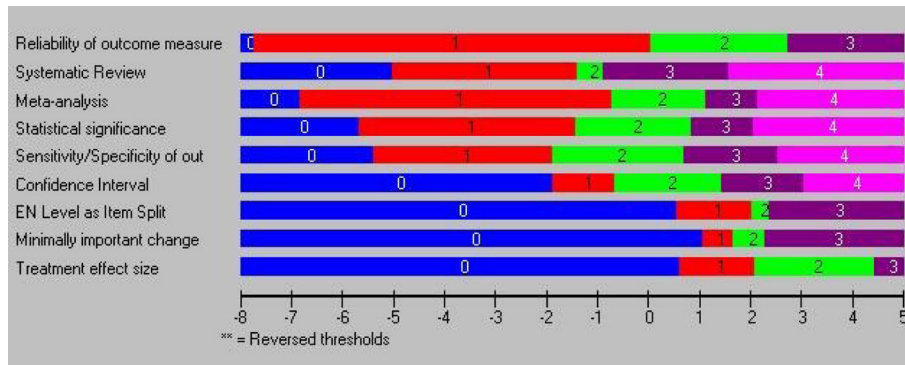
Items 5 and 11 showed DIF by language. Upon review, the English wording was idiomatic and the French translation did not reflect the correct meaning of the idiom. This might be attributed to the lack of proficiency in the English language [64]. The occurrence of DIF in this construct was considered

infrequent (2/17 items) and comparable with other findings [65–67]. Future measure development should use simultaneous translation as it has advantages over forward translation particularly in that idiomatic wording; linguistic discrepancies between languages can be minimized or resolved during the item generation process [64, 68].

Self-efficacy

The final eight items fit the Rasch model and were reasonably well targeted to participants; however, the mean value for the participants was > 0 suggesting that the participants had a higher level of self-efficacy towards EBP than expected from the items. Self-efficacy has good internal reliability ($PSI = 0.80$), indicating that the items adequately separated the sample along the self-efficacy continuum. The internal consistency reliability value we found is similar to the original source of our items which is the evidence-based

a Threshold map



b Targeting map

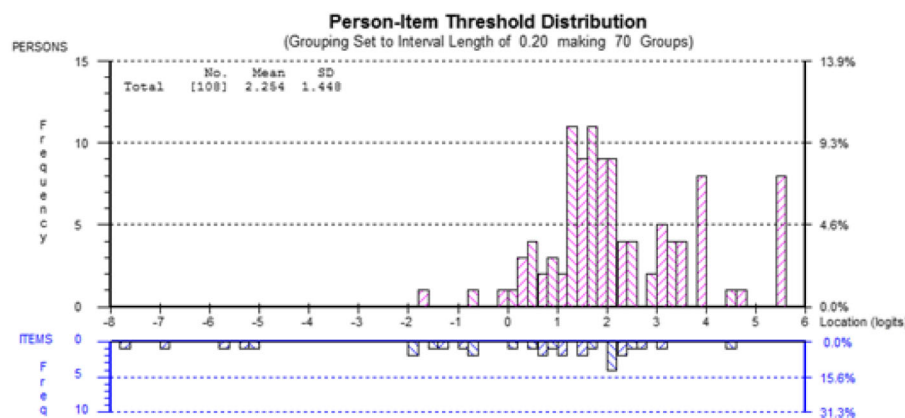


Fig. 5 a Threshold map and **b** targeting map of “knowledge” construct

practice confidence (EPIC) scale [45]. RMT has been applied to testing self-efficacy towards EBP showing similar findings among different targeted populations with larger sample sizes [31, 32].

Knowledge

Nine items fit the Rasch model, but the sample was not well targeted by the items with mean fit for the participants (> 0); this suggests that the overall level of knowledge about EBP of our participants was greater than what would be expected from the items. The knowledge construct has a good internal reliability (PSI = 0.81), indicating that the items adequately separated our participants along the measurement continuum. RMT has been used by others [27, 28, 30] to validate knowledge of EBP, and this research has yielded very similar findings to ours.

Three items showed DIF: items 4 which showed that OTs’ familiarity with the statistical term “meta-analysis” is greater than that of PTs’. Item 9 showed that PTs are more aware of the statistical term “minimally important change (MIC)” than OTs. This might be explained by OTs’ use of a combination of different assessment and

treatment approaches and their focus on client-centred strategies. Standardized tools are used when relevant and available [69–71]. PTs might know more the measurement-related statistical terms like “minimally important change (MIC)” as they mainly use the standardized tool over qualitative approaches. Also, the professional education and exposure to concepts associated with EBP may vary across programmes and between the two professions. Item 7 showed that francophones are more familiar with the term “number-needed-to-treat (NNT)” than anglophones. However, the DIF here was more likely due to the poor translation rather than real differences because “nombre de sujet à traiter” is associated with “caseload” in French than with a statistical term. Harris-Haywood et al. [30] evaluated the measurement properties of the Cultural Competence Health Practitioner Assessment Questionnaire across health care professionals which includes the following domains: knowledge, adapting practice, and promoting health for culturally and linguistically diverse populations. The authors reported DIF by almost all 23 knowledge items: 17 items had DIF by race, 4 by gender, and 6 by profession.

Table 8 Results of Rasch analysis for the original items of the construct knowledge of EBP

Instructions: Please indicate your level of agreement with the following statements with respect to your organization or workplace setting: 5-point Likert Scale

(Directives: Veuillez indiquer à quel point vous êtes en accord avec les énoncés suivants)

Item #	Description of item	Response option rescored	Result		
			Item misfit	Local item dependency	DIF
1	E Reliability of outcome measures	✓	No	No	No
	F Fidélité de la mesure des résultats				
2	E Validity of outcome measures		No	Yes with item 10 (item 2 deleted)	No
	F Validité de la mesure des résultats				
3	E Sensitivity/specificity of outcome measures		No	No	No
	F Sensibilité/spécificité de la mesure des résultats				
4	E Meta-analysis		No	No	Profession (left)
	F Méta-analyse				
5	E Confidence interval		No	No	No
	F Intervalle de confiance				
6	E Systematic review		No	No	No
	F Revue systématique				
7	E Number needed to treat	✓	No	No	Language (item split, French item deleted)
	F Nombre de sujets à traiter				
8	E Statistical significance		No	Yes	No
	F Signification statistique				
9	E Minimally important change (MIC)	✓	No	No	Profession (left)
	F Différence minimale cliniquement importante (DMCI)				
10	E Treatment effect size	✓	No	Yes	No
	F Ampleur de l'effet du traitement				
11	E Randomized controlled trial (RCT)		No	Yes with item 8 (item 11 deleted)	No
	F Essai clinique randomisé				

DIF differential item functioning, E English, F French translation

Resources

The 13 items reflecting resources available for EBP fit the Rasch model and reasonably well for the targeted participants. The mean value was >0 , but there are many items at the low end of the resource continuum with no participants. This suggested that the resources available for our participants were greater than what would be expected from the items included.

This might suggest a selection bias towards people who are good in EBP [72]. Although all items fit, there was still evidence of global misfit, indicating that a review of the items in this measure is warranted. This construct has a good internal reliability (PSI = 0.86), indicating that the items adequately separated participants along the measurement continuum. Bobiak et al. [29] showed almost similar findings in terms of the threshold range of the items and PSI values.

This project would not have been possible without the wealth of items and measures already developed, tested, and used in the study of EBP. However, in order to implement these measures in a practical way, through a survey of EBP of busy practitioners, we had to ensure a minimum number of non-redundant items. Otherwise, the response rate would be nil. This process, which was started for practical reasons, took the field of implementation science a step forward by optimizing the measurement of these important constructs so that it is easier to detect differences between groups and change over time.

Implementation of research in practice is considered to be a multidimensional process that involves clients, practitioners, and organizations [73]. Therefore, the development of measures that can reflect these dimensions is important because it provides tools that help the ultimate goal of implementation for behaviour change (i.e. the use of EBP).

Table 9 Results of Rasch analysis for the original items of the construct resources

Instructions: Please indicate your level of agreement with the following statements with respect to your organization or workplace setting: 5-point Likert Scale
(Directives: Veuillez indiquer à quel point vous êtes en accord avec les énoncés suivants)

Item #	Description of item	Response option rescored	Result		
			Item misfit	Local item dependency	DIF
1	E I am comfortable talking about patient/client care issues with those in charge at the organization F Je suis à l'aise de parler de problèmes liés au soin d'un patient ou client avec les responsables de l'organisation	✓	No	No	No
2	E I receive recognition from my manager(s)/supervisor(s) about my work F Mon ou mes supérieurs ou superviseurs apprécie mon travail	✓	No	No	No
3	E I have control over how I do my work F J'exerce un contrôle sur la façon dont je fais mon travail	✓	No	No	No
4	E My organization emphasizes productivity F Mon organisation valorise la productivité	✓	No	No	No
5	E My organization supports best practice F Mon organisation soutient les pratiques optimales	✓	No	No	No
6	E I have opportunities for educational activities in my organization F J'ai l'occasion d'assister à des activités de perfectionnement dans mon organisation	✓	No	No	No
7	E I have formal patient/client related discussions with peers or colleagues (e.g. continuing education, patient rounds, team meetings) in my organization F J'ai des discussions structurées sur les patients ou clients avec mes pairs ou collègues (ex. formation continue, tournée des patients, réunions d'équipe) dans mon organisation	✓	No	No	No
8	E I have informal patient/client related discussions with peers or colleagues (e.g. other health care providers, informal bedside teaching) in my organization F J'ai des discussions informelles sur les patients ou clients avec mes pairs ou collègues (ex. autres professionnels de la santé, enseignement au chevet du patient) dans mon organisation	✓	No	No	No
9	E My organization routinely provides information/feedback on my practice (e.g. audits, performance reviews) F Je reçois régulièrement de mon organisation une rétroaction ou des commentaires (ex. audit, évaluation de rendement) au sujet de ma pratique	✓	No	No	No
10	E I have access to resources at my workplace to help deliver quality care for my patients/clients (e.g. databases, libraries, equipment) F Dans mon milieu de travail, j'ai accès à des ressources qui aident à améliorer la qualité des soins prodigués à mes patients ou clients (ex. bases de données, bibliothèque, équipement).	✓	No	No	No
11	E All OT/PT positions at my workplace are currently filled F Tous les postes en physiothérapie ou ergothérapie sont comblés en ce moment dans mon milieu de travail	✓	No	No	No
12	E There is a high OT/PT clinician staff turnover rate at my workplace F Il y a un fort roulement du personnel clinique d'ergothérapie et de physiothérapie dans mon milieu de travail	✓	Yes with a value of 4.41 (item 12 deleted)	No	No
13	E I have access to space I need to do my job well at my workplace F J'ai accès à la place dont j'ai besoin pour bien faire mon travail	✓	No	Yes	No
14	E There is an appropriate space to provide quality care F Les locaux sont appropriés à la prestation de soins de qualité	✓	No	Yes with item 13 (item 14 deleted)	No
15	E I have time to do indirect patient activities (e.g. talk about a plan of care, look up something in a journal, get involved in new initiatives at work)	✓	No	No	No

Table 9 Results of Rasch analysis for the original items of the construct resources (*Continued*)

Instructions: Please indicate your level of agreement with the following statements with respect to your organization or workplace setting: 5-point Likert Scale
 (Directives: Veuillez indiquer à quel point vous êtes en accord avec les énoncés suivants)

Item #	Description of item	Response option rescored	Result		
			Item misfit	Local item dependency	DIF
	in my practice				
F	J'ai le temps de faire des activités indirectement liées aux patients (ex. discuter d'un plan de soin, chercher dans une revue, participer à de nouvelles initiatives professionnelles) dans ma pratique				

DIF differential item functioning, E English, F French translation

Furthermore, these measures will ease the study of behaviour change mechanisms like those presumed in the Theoretical Domains Framework [74] or Theory of Planned Behaviour [75]. For instance, these mechanisms may include factors such as attitudes towards EBP or knowledge of EBP that in turn impact the intention to use EBP in daily practice. As a result, if we can provide measures of attitudes towards EBP or knowledge of EBP, this then will predict EBP behaviours (i.e. actual use of EBP) [73]. The final set of items provided in this work for each construct fairly covers the full range of the latent variable from “easier” to “difficult” items which will allow researchers to capture different abilities of clinicians regarding aspects of EBP. That by itself will support researchers in the design and implementation of robust knowledge translation interventions in a gradual manner aimed at targeting different levels of clinician abilities instead of designing one-size-fits-all interventions.

Strengths

Our study has several strengths. The analyses provide a single legitimate total score for four constructs around EBP on a continuum from very low to very high. These measures can be used to identify determinants of good EBP, to guide the development of implementation interventions, and to evaluate sustainability. The validation of such indices around

EBP addresses the lack of standardized tools for EBP educational evaluation [19, 76–78]. This study tackled a specific gap in EBP measures for groups like PTs and OTs in two languages, English and French. The analyses provided a mathematically sound way of reducing the multitude of items to only those that behave in the manner that sufficiently assesses the constructs around EBP while maintaining strong theoretical and methodological processes. We used data from an inception cohort which will allow us to follow participants during the next 3 years to gauge the changes over time. We recruited participants from several Canadian universities in English and French languages which allow for generalizing the findings. The use of simultaneous translation helped to resolve idiomatic wording and linguistic discrepancies between languages compared to sequential translation where items are first generated in only one language followed by subsequent translation into another language [64, 68]. Future work is required to validate these constructs over time, and there is a need to address the gaps in the lower and upper ends of the scales. Further, the correlation between different EBP constructs will allow for greater understanding of the EBP constructs. This will be described in a subsequent descriptive analysis of this large data set and though the ongoing longitudinal study of these graduates into practice.

Table 10 Summary of the results of the Rasch analysis yielding four measures of EBP

Construct	Items		Thresholds	N at ceiling	p value for global fit	alpha (α)	Threshold range	Item location (SD)	Person location mean (SD)
	Start	Finish							
Attitudes	17	10	35	110 (86%)	0.29	0.63	– 7.51 to 5.04	1.29	– 0.41 (1.41)
Self-efficacy	9	8	22	76 (59%)	0.82	0.80	– 3.70 to 4.91	1.60	– 0.34 (1.12)
Resources	15	13	39	104 (81%)	0.00*	0.86	– 3.38 to 2.86	0.51	– 0.34 (1.73)
Knowledge	11	9	32	108 (84%)	0.12	0.81	– 7.85 to 4.50	1.56	– 0.26 (1.23)
All	52	40							

*All items fit the Rasch model, but improvements are still needed

Limitations

Our study has some limitations. First, the data were from a cross-sectional study, and unnecessarily, the data reflect all time points, a weakness reported in previous work [27–32]. Therefore, it is essential to revalidate these constructs with subsequent waves of data to explore stability, intensity, and direction of EBP constructs. Second, the participants were new graduates, and 43% reported that they were not working in clinical practice which limits the use of EBP. Third, the sample size was generally small and diversified across the EBP constructs due to missing data. Fourth, the number of administered items was much less than other questionnaires targeting several constructs; however, it requires about 15 min, which might influence the return rates especially for future time points. Fifth, it was noticed that the level of EBP knowledge for our participants was high which causes a ceiling effect that may explain the high mean person location (> 0) for all EBP constructs. Sixth, despite the fact that the items capture a wide range of difficulty, gaps appear in some locations, making the estimation of the ability of persons located near or within those gaps less precise. Lastly, the knowledge items may not be the true reflection of the participants' EBP knowledge level since this construct elicits knowledge about statistical and methodological terms, one aspect of EBP. Some items were missed in our knowledge item sets such as knowing how to search literature and different literature appraisal issues.

Conclusion

This study provides evidence supporting the construct and content validity and internal consistency reliability of four measure to assess key constructs around EBP. Our use of a strong theoretical and methodological processes was helpful in reducing the number of items for further analysis, and the use of Rasch analysis was critical and helpful in choosing the best fitting items for each construct. However, a number of gaps in the measures were uncovered indicating a further refinement of the content and wording is needed.

Additional file

Additional file 1: A list of included and excluded items. (XLSX 31 kb)

Abbreviations

DIF: Differential item functioning; EBP: Evidence-based practice; OT: Occupational therapy; PT: Physical therapy; RMT: Rasch Measurement Theory

Acknowledgements

The authors would like to thank Carly Goodman, MSc, for helping with sending invitations, developing the e-survey, data collection, data processing, and taking field notes. We wish to thank all physiotherapists and occupational therapists who took part in our study for their time. FAZ acknowledges the

generous support of the postdoctoral bursary from the Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR).

Funding

This study was funded by the Canadian Institutes of Health Research (CIHR). The funding body did not influence the study results and was not involved in the study in any way.

Availability of data and materials

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

All authors participated in adapting the study design, contributed to the development of the item set, and analysed and interpreted the data. FAZ wrote the initial manuscript. NEM and AT oversaw all phases of the study. The authors contributed to the study design, development of the item set, and analyses and interpretation of the data. All authors read and approved the final manuscript.

Authors' information

FAZ has clinical training in physical therapy with an MSc degree in Professional Physical Therapy, a thesis-based MSc in rehabilitation science and a PhD in Rehabilitation Science from the School of Physical and Occupational Therapy at McGill University. NEM is the James McGill Professor in the Department of Medicine and the School of Physical and Occupational Therapy at McGill University. She is one of Canada's best-known researchers in stroke rehabilitation. She is a health outcomes, health services, and population health researcher with interests in all aspects of disability and quality of life in people with chronic diseases and the elderly. AR has clinical training in occupational therapy and is a Professor at the School of Rehabilitation at the University of Montreal. AT has clinical training in occupational therapy and is an Associate Professor at the School of Physical and Occupational Therapy at McGill University with a cross-appointment at the Center for Medical Education, Faculty of Medicine, McGill University.

Ethics approval and consent to participate

All participants provided informed consent. Approval for this project was granted by all Institutional Review Board belonging to each university.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹McGill University, Montreal, QC, Canada. ²Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Montreal, QC, Canada. ³Center for Outcomes Research and Evaluation (CORE), Research Institute of McGill University Health Center, Montreal, QC, Canada. ⁴Université de Montréal, Montreal, QC, Canada. ⁵Centre for Medical Education, Montreal, Canada.

Received: 20 July 2018 Accepted: 28 November 2018

Published online: 18 December 2018

References

- Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71–2.
- Feng X, Liu C, Guo Q, Bai Y, Ren Y, Ren B, et al. Research progress in rehabilitation treatment of stroke patients: a bibliometric analysis. *Neural Regen Res*. 2013;8(15):1423–30. <https://doi.org/10.3969/j.issn.1673-5374.2013.15.010>.
- Frontera WR. Some reflections on the past, present, and future of physical and rehabilitation medicine (on the occasion of the 30th SOFMER congress). *Ann Phys Rehabil Med*. 2016;59(2):79–82. <https://doi.org/10.1016/j.rehab.2016.02.002>.

4. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. *J Speech Lang Hear Res*. 2008;51(1):S225–39. [https://doi.org/10.1044/1092-4388\(2008/018\)](https://doi.org/10.1044/1092-4388(2008/018)).
5. Winstein CJ, Kay DB. Translating the science into practice: shaping rehabilitation practice to enhance recovery after brain damage. *Prog Brain Res*. 2015;218:331–60. <https://doi.org/10.1016/bs.pbr.2015.01.004>.
6. Moore JL, Shikako-Thomas K, Backus D. Knowledge translation in rehabilitation: a shared vision. *Pediatr Phys Ther*. 2017;29(Suppl 3 Supplement, IV Step 2016 Conference Proceedings Supplement):S64–s72. <https://doi.org/10.1097/pep.0000000000000381>.
7. World Health Organisation. Rehabilitation: WHO framework on rehabilitation services: expert meeting. 2017.
8. Korner-Bitensky N, Desrosiers J, Rochette A. A national survey of occupational therapists' practices related to participation post-stroke. *J Rehabil Med*. 2008;40(4):291–7. <https://doi.org/10.2340/16501977-0167>.
9. Salls J, Dolhi C, Silverman L, Hansen M. The use of evidence-based practice by occupational therapists. *Occup Ther Health Care*. 2009;23(2):134–45. <https://doi.org/10.1080/07380570902773305>.
10. Petzold A, Korner-Bitensky N, Rochette A, Teasell R, Marshall S, Perrier MJ. Driving poststroke: problem identification, assessment use, and interventions offered by Canadian occupational therapists. *Top Stroke Rehabil*. 2010;17(5):371–9. <https://doi.org/10.1310/tsr1705-371>.
11. Menon-Nair A, Korner-Bitensky N, Ogoturova T. Occupational therapists' identification, assessment, and treatment of unilateral spatial neglect during stroke rehabilitation in Canada. *Stroke*. 2007;38(9):2556–62. <https://doi.org/10.1161/strokeaha.107.484857>.
12. Accreditation Council for Canadian Physiotherapy Academic Programs. Essential competency profile for physiotherapists in Canada. 2009. <http://www.physiotherapyeducation.ca/Resources/Essential%20Comp%20PT%20Profile%202009.pdf>. Accessed 15 April 2018.
13. Canadian Association of Occupational Therapy. Profile of practice of occupational therapists in Canada. 2012. <https://www.caot.ca/document/3653/2012otprofile.pdf>. Accessed 15 April 2018.
14. Godin G, Belanger-Gravel A, Eccles M, Grimshaw J. Healthcare professionals' intentions and behaviours: a systematic review of studies based on social cognitive theories. *Implement Sci*. 2008;3:36. <https://doi.org/10.1186/1748-5908-3-36>.
15. Thompson-Leduc P, Clayman ML, Turcotte S, Legare F. Shared decision-making behaviours in health professionals: a systematic review of studies based on the Theory of Planned Behaviour. *Health Expect*. 2015;18(5):754–74. <https://doi.org/10.1111/hex.12176>.
16. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol Rev*. 2016;10(3):277–96. <https://doi.org/10.1080/17437199.2016.1151372>.
17. Newhouse RP. Instruments to assess organizational readiness for evidence-based practice. *J Nurs Adm*. 2010;40(10):404–7. <https://doi.org/10.1097/NNA.0b013e3181f2ea08>.
18. Gawlinski A. Evidence-based practice changes: measuring the outcome. *AACN Adv Crit Care*. 2007;18(3):320–2.
19. Shaneyfelt T, Baum KD, Bell D, Feldstein D, Houston TK, Kaatz S, et al. Instruments for evaluating education in evidence-based practice: a systematic review. *JAMA*. 2006;296(9):1116–27. <https://doi.org/10.1001/jama.296.9.1116>.
20. Cadarin L, Bagnasco A, Tolotti A, Pagnucci N, Sasso L. Instruments for measuring meaningful learning in healthcare students: a systematic psychometric review. *J Adv Nurs*. 2016;72(9):1972–90. <https://doi.org/10.1111/jan.12926>.
21. Buchanan H, Siegfried N, Jelsma J. Survey instruments for knowledge, skills, attitudes and behaviour related to evidence-based practice in occupational therapy: a systematic review. *Occup Ther Int*. 2016;23(2):59–90. <https://doi.org/10.1002/oti.1398>.
22. Leung K, Trevena L, Waters D. Systematic review of instruments for measuring nurses' knowledge, skills and attitudes for evidence-based practice. *J Adv Nurs*. 2014;70(10):2181–95. <https://doi.org/10.1111/jan.12454>.
23. Sick J. Rasch measurement in language education part 5: assumptions and requirements of Rasch measurement. *Shiken*. 2010;14(2):23–9.
24. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med*. 2012;44(2):97–8. <https://doi.org/10.2340/16501977-0938>.
25. Andrich D. Rasch models for ordered response categories. *Encyclopedia Stat Behav Sci*. 2005;4:1698–707. <https://doi.org/10.1002/0470013192.bsa541>.
26. Traub RE. Classical Test Theory in historical perspective. *Educ Meas*. 1997;16(4):8–14. <https://doi.org/10.1111/j.1745-3992.1997.tb00603.x>.
27. Akram W, Hussein MS, Ahmad S, Mamat MN, Ismail NE. Validation of the knowledge, attitude and perceived practice of asthma instrument among community pharmacists using Rasch analysis. *Saudi Pharm J*. 2015;23(5):499–503. <https://doi.org/10.1016/j.jpsps.2015.01.011>.
28. Zoanetti N, Griffin P, Beaves M, Wallace EM. Rasch scaling procedures for informing development of a valid Fetal Surveillance Education Program multiple-choice assessment. *BMC Med Educ*. 2009;9:20. <https://doi.org/10.1186/1472-6920-9-20>.
29. Bobiak SN, Zyzanski SJ, Ruhe MC, Carter CA, Ragan B, Flocke SA, et al. Measuring practice capacity for change: a tool for guiding quality improvement in primary care settings. *Qual Manag Health Care*. 2009;18(4):278–84. <https://doi.org/10.1097/QMH.0b013e3181bee2f5>.
30. Harris-Haywood S, Goode T, Gao Y, Smith K, Bronheim S, Flocke SA, et al. Psychometric evaluation of a cultural competency assessment instrument for health professionals. *Med Care*. 2014;52(2):e7–e15. <https://doi.org/10.1097/MLR.0b013e31824df149>.
31. Blackman IR, Giles T. Psychometric evaluation of a self-report evidence-based practice tool using Rasch analysis. *Worldviews Evid Based Nurs*. 2015;12(5):253–64. <https://doi.org/10.1111/wvn.12105>.
32. Hasnain M, Gruss V, Keehn M, Peterson E, Valenta AL, Kottorp A. Development and validation of a tool to assess self-efficacy for competence in interprofessional collaborative practice. *J Interprof Care*. 2017;31(2):255–62. <https://doi.org/10.1080/13561820.2016.1249789>.
33. Thomas A, Rochette A, Lapointe J, O'Connor K, Ahmed S, Bussi eres A, et al. Evolution of evidence-based practice: evaluating the contribution of individual and contextual factors to optimize patient care. Canadian Institutes for Health Research, Project scheme grant; 2016–2022. 2016. http://webapps.cihr-irsc.gc.ca/funding/detail_e?pResearchId=8239783&p_version=CRIS&p_language=E&p_session_id=
34. Fekete C, Boldt C, Post M, Eriks-Hoogland I, Cieza A, Stucki G. How to measure what matters: development and application of guiding principles to select measurement instruments in an epidemiologic study on functioning. *Am J Phys Med Rehabil*. 2011;90(11 Suppl 2):S29–38. <https://doi.org/10.1097/PHM.0b013e318230fe41>.
35. Geyh S, Cieza A, Kollerits B, Grimby G, Stucki G. Content comparison of health-related quality of life measures used in stroke based on the international classification of functioning, disability and health (ICF): a systematic review. *Qual Life Res*. 2007;16(5):833–51. <https://doi.org/10.1007/s11136-007-9174-8>.
36. Estabrooks CA, Squires JE, Cummings GG, Birdsell JM, Norton PG. Development and assessment of the Alberta Context Tool. *BMC Health Serv Res*. 2009;9:234. <https://doi.org/10.1186/1472-6963-9-234>.
37. Delbecq AL, Van de Ven AH. A group process model for problem identification and program planning. *J Appl Behav Sci*. 1971;7(4):466–92. <https://doi.org/10.1177/002188637100700404>.
38. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health*. 1984;74(9):979–83.
39. McEvoy MP, Williams MT, Olds TS, Lewis LK, Petkov J. Evidence-based practice profiles of physiotherapists transitioning into the workforce: a study of two cohorts. *BMC Med Educ*. 2011;11:100. <https://doi.org/10.1186/1472-6920-11-100>.
40. McEvoy MP, Williams MT, Olds TS. Development and psychometric testing of a trans-professional evidence-based practice profile questionnaire. *Med Teach*. 2010;32(9):e373–80. <https://doi.org/10.3109/0142159x.2010.494741>.
41. Aarons GA, Glisson C, Hoagwood K, Kelleher K, Landsverk J, Cafri G. Psychometric properties and U.S. national norms of the Evidence-Based Practice Attitude Scale (EBPAS). *Psychol Assess*. 2010;22(2):356–65. <https://doi.org/10.1037/a0019188>.
42. Aarons GA. Mental health provider attitudes toward adoption of evidence-based practice: the Evidence-Based Practice Attitude Scale (EBPAS). *Ment Health Serv Res*. 2004;6(2):61–74.
43. Aarons GA, McDonald EJ, Sheehan AK, Walrath-Greene CM. Confirmatory factor analysis of the evidence-based practice attitude scale in a geographically diverse sample of community mental health providers. *Admin Pol Ment Health*. 2007;34(5):465–9. <https://doi.org/10.1007/s10488-007-0127-x>.
44. Salbach NM, Jaglal SB. Creation and validation of the evidence-based practice confidence scale for health care professionals. *J Eval Clin Pract*. 2011;17(4):794–800. <https://doi.org/10.1111/j.1365-2753.2010.01478.x>.

45. Salbach NM, Jaglal SB, Williams JI. Reliability and validity of the evidence-based practice confidence (EPIC) scale. *J Contin Educ Heal Prof.* 2013;33(1):33–40. <https://doi.org/10.1002/chp.21164>.
46. Clyde JH, Brooks D, Cameron JI, Salbach NM. Validation of the evidence-based practice confidence (EPIC) scale with occupational therapists. *Am J Occup Ther.* 2016;70(2):7002280010p1–9. <https://doi.org/10.5014/ajot.2016.017061>.
47. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16(3):297–334. <https://doi.org/10.1007/bf02310555>.
48. Squires JE, Hayduk L, Hutchinson AM, Mallick R, Norton PG, Cummings GG, et al. Reliability and validity of the Alberta Context Tool (ACT) with professional nurses: findings from a multi-study analysis. *PLoS one.* 2015;10(6):e0127405. <https://doi.org/10.1371/journal.pone.0127405>.
49. Lynn MR. Determination and quantification of content validity. *Nurs Res.* 1986;35(6):382–5.
50. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).* 2000;25(24):3186–91.
51. Gibbons CJ, Mills RJ, Thornton EW, Ealing J, Mitchell JD, Shaw PJ, et al. Rasch analysis of the hospital anxiety and depression scale (HADS) for use in motor neurone disease. *Health Qual Life Outcomes.* 2011;9:82. <https://doi.org/10.1186/1477-7525-9-82>.
52. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum.* 2007;57(8):1358–62. <https://doi.org/10.1002/art.23108>.
53. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. Third ed. New York: Routledge; 2015.
54. Linacre JM. Sample size and item calibration [for person measure] stability. *Rasch measurement transactions;* 1994. <http://www.rasch.org/rmt/rmt74m.htm>. Accessed 15 May 2018.
55. Andrich D, Sheridan B, Luo G. Rasch unidimensional measurement models (RUMM) 2030. Perth: RUMM Pty Ltd; 2008.
56. Statistical Analysis System Institute Inc. SAS® 9.4 Software. 2014. <http://support.sas.com/software/94/index.html>. 15 Mar 2018.
57. Morrison T, Robertson L. New graduates' experience of evidence-based practice: an action research study. *Br J Occup Ther.* 2016;79(1):42–8. <https://doi.org/10.1177/0308022615591019>.
58. Powell CA, Case-Smith J. Information literacy skills of occupational therapy graduates: promoting evidence-based practice in the MOT curriculum. *Med Ref Serv Q.* 2010;29(4):363–80. <https://doi.org/10.1080/02763869.2010.518923>.
59. Thomas A, Law M. Research utilization and evidence-based practice in occupational therapy: a scoping study. *Am J Occup Ther.* 2013;67(4):e55–65. <https://doi.org/10.5014/ajot.2013.006395>.
60. Wilkinson SA, Hinchliffe F, Hough J, Chang A. Baseline evidence-based practice use, knowledge, and attitudes of allied health professionals: a survey to inform staff training and organisational change. *J Allied Health.* 2012;41(4):177–84.
61. Brown T, Tseng MH, Casey J, McDonald R, Lyons C. Research knowledge, attitudes, and practices of pediatric occupational therapists in Australia, the United Kingdom, and Taiwan. *J Allied Health.* 2010;39(2):88–94.
62. Mroz TM, Pitonyak JS, Fogelberg D, Leland NE. Client centeredness and health reform: key issues for occupational therapy. *Am J Occup Ther.* 2015;69(5):6905090010p1–8. <https://doi.org/10.5014/ajot.2015.695001>.
63. Alsop A. Evidence-based practice and continuing professional development. *Br J Occup Ther.* 1997;60(11):503–8. <https://doi.org/10.1177/030802269706001112>.
64. Rogers WT, Lin J, Rinaldi CM. Validity of the simultaneous approach to the development of equivalent achievement tests in English and French. *Appl Meas Educ.* 2010;24(1):39–70. <https://doi.org/10.1080/08957347.2011.532416>.
65. Kwakkenbos L, Willems LM, Baron M, Hudson M, Cella D, van den Ende CHM, et al. The comparability of English, French and Dutch scores on the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F): an assessment of differential item functioning in patients with systemic sclerosis. *PLoS One.* 2014;9(3):e91979. <https://doi.org/10.1371/journal.pone.0091979>.
66. Delisle VC, Kwakkenbos L, Hudson M, Baron M, Thombs BD. An assessment of the measurement equivalence of English and French versions of the Center for Epidemiologic Studies Depression (CES-D) Scale in systemic sclerosis. *PLoS One.* 2014;9(7):e102897. <https://doi.org/10.1371/journal.pone.0102897>.
67. Arthurs E, Steele RJ, Hudson M, Baron M, Thombs BD. Are scores on English and French versions of the PHQ-9 comparable? An assessment of differential item functioning. *PLoS One.* 2012;7(12):e52028. <https://doi.org/10.1371/journal.pone.0052028>.
68. Askari S, Fellows L, Brouillette MJ, Moriello C, Duracinsky M, Mayo NE. Development of an item pool reflecting cognitive concerns expressed by people with HIV. *Am J Occup Ther.* 2018;72(2):7202205070p1–9. <https://doi.org/10.5014/ajot.2018.023945>.
69. Stapleton T, McBrearty C. Use of standardised assessments and outcome measures among a sample of Irish occupational therapists working with adults with physical disabilities. *Br J Occup Ther.* 2009;72(2):55–64. <https://doi.org/10.1177/030802260907200203>.
70. Bowman J. Challenges to measuring outcomes in occupational therapy: a qualitative focus group study. *Br J Occup Ther.* 2006;69(10):464–72. <https://doi.org/10.1177/030802260606901005>.
71. Atwal A, McIntyre A, Craik C, Hunt J. Occupational therapists' perceptions of predischarge home assessments with older adults in acute care. *Br J Occup Ther.* 2008;71(2):52–8. <https://doi.org/10.1177/030802260807100203>.
72. Fowler F. Survey research methods. 4th ed. Thousand Oaks; 2009. <https://doi.org/10.4135/9781452230184>.
73. Moullin JC, Ehrhart MG, Aarons GA. Development and testing of the Measure of Innovation-Specific Implementation Intentions (MISII) using Rasch measurement theory. *Implement Sci.* 2018;13(1):89. <https://doi.org/10.1186/s13012-018-0782-1>.
74. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implement Sci.* 2012;7:37. <https://doi.org/10.1186/1748-5908-7-37>.
75. Ajzen I. The theory of planned behavior. *Organ Behav Hu Dec Processes.* 1991;50(2):179–211. doi:[https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T).
76. Stevenson K, Lewis M, Hay E. Do physiotherapists' attitudes towards evidence-based practice change as a result of an evidence-based educational programme? *J Eval Clin Pract.* 2004;10(2):207–17. <https://doi.org/10.1111/j.1365-2753.2003.00479.x>.
77. Stevenson K, Lewis M, Hay E. Does physiotherapy management of low back pain change as a result of an evidence-based educational programme? *J Eval Clin Pract.* 2006;12(3):365–75. <https://doi.org/10.1111/j.1365-2753.2006.00565.x>.
78. Flores-Mateo G, Argimon JM. Evidence based practice in postgraduate healthcare education: a systematic review. *BMC Health Serv Res.* 2007;7:119. <https://doi.org/10.1186/1472-6963-7-119>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

