PJHRD
Philippine Journal of Health Research and Development
University of the Philippines Manila - The Health Sciences Center
Information, Publication and public affairs Office (IPPAO)
8/F Philippine General Hospital Complex, Taft Avenue, Manila 1000 Philippines

Online ISSN: 2783-042X
April-June
VOL. 29 NO. 2 2025
PAGE 54-59

RESEARCH ARTICLE

Utilizing cognitive interview in the item refinement of the Blended Teaching Assessment Tool (BTAT) for Health Professions Education

Maria Teresita B. Dalusong[1]*, Glenda Sanggalang Ogerio[2], Valentin C. Dones III[3], Maria Elizabeth M. Grageda[2]

[1]College of Physical Therapy, Pamantasan ng Lungsod ng Maynila (University of the City of Manila), Philippines
[2]National Teacher Training Center for the Health Professions, University of the Philippines Manila, Manila, Philippines
[3]College of Rehabilitation Science, University of Santo Tomas, Philippines

ABSTRACT

Background: A high-quality measurement tool is essential to accurately assess the innovative teaching strategies in health professions education. The Blended Teaching Assessment Tool (BTAT) aims to evaluate quality blended teaching or instructional delivery in Philippine health science programs. However, there is a lack of studies examining students' cognitive processes to support the validity of questionnaires.

Methodology: Cognitive interviewing (CI) was employed to determine whether students interpreted and responded to the items correctly. Content analysis was done using Tourangeau's Cognitive framework. Four CIs were conducted by an expert moderator and note-taker with a total of 8 health science students (2 groups with 3 members, and 2 one-on-one interview) for around 1 to 2 hours via Zoom following a retroactive approach with verbal and spontaneous probing, guided by a semi-structured interview questionnaire.

Results: Various issues related to comprehension, retrieval, judgment, and response were identified, leading to significant revisions of the tool from 82 items across 8 dimensions to 53 items across 5 dimensions. The challenges included unfamiliar terminology, ambiguous phrasing, complex statements, inconsistencies and irrelevance to students' real-life experiences. These findings emphasize the importance of students' feedback in enhancing the validity and reliability of assessment tools.

Conclusion: The Cognitive Interview identified crucial issues in comprehension, retrieval, judgment, and response, making it essential for developing the Blended Teaching Assessment Tool and ensuring valid responses on the quality of blended teaching and learning delivery.

Introduction

Developing high-quality measurement tools is essential to accurately measure the quality of instructional delivery like Blended Teaching-Learning (BTL) and provide reliable and valid evaluation data [1,2]. The measurement instrument development is a complex and multi-step process that requires theoretical, methodological, and statistical competencies [3]. It involves creating appropriate items and measurement scales tailored to specific target population [4]. Apart from psychometric considerations, the process must also address item clarity, comprehensibility, appropriate response options, and feasibility in real-world settings. Neglecting these factors can compromise data quality and the instrument's reliability and validity.

Cognitive interviewing (CI) is a qualitative research method used to develop and refine assessments across various fields, enhancing their validity, comprehensibility, and cultural relevance [5,6]. Guided by the Tourangeau's four steps of cognitive operation: 1) comprehension, 2) retrieval, 3) judgment, and 4) response, CI explores how target participants interpret, recall relevant and comprehensiveness of their experiences and respond to assessment items, helping to identify and address biases, ambiguities, or cultural nuances of the items [7,8]. This has been an essential phase in the development and refinement of instrument used in health care setting and in health professions education which determined alignment of participants' interpretations and responses with the item's or question's original intention [5,9-17].

However, in Philippines, few studies reported the use of CI in the development of patients' self-reported questionnaire, while no local study was found utilizing CI in the development of instruments for assessment of teaching and learning methods due to various challenges in Health Professions [18-21].

The BTL, which combines online and face-to-face instruction, has gained traction in health professions education globally and locally, particularly during the COVID-19 pandemic [21-27]. In the Philippine context, emergence of BTL presents specific challenges, including the need for curriculum recalibration, enhanced teacher training, and validated, and reliable assessments of BTL quality [28-33].

Teaching quality in higher education is commonly measured using student evaluation of teaching (SET) or students' faculty evaluation (SFE) tools [34]. While these tools are widely used, their effectiveness may be limited by the newly adapted innovations in Blended approach and the resulting changes in student experiences, which can introduce response bias [35,36]. Dones' (2024) systematic review of the development and psychometric properties of blended teaching evaluation instruments identified only two out of the five qualified studies assessing BTL showed "adequate" content validity, while the others were deemed "doubtful" according to the COSMIN criteria [32,37]. This inadequacy stems from an emphasis on expert validation without incorporating CIs to enhance relevance, clarity, and comprehensiveness for the targeted university students [32,33].

Among the recently developed and published blended teaching assessment tools [38-44], pre-testing was limited to asking students on items' clarity and comprehension [39-41]. No CI on thinking process were reported among these instruments. The blended teaching quality evaluation scale (BTQS) was the only instrument developed for undergraduate Nursing [40], while the Blended Learning Usability Evaluation – Questionnaire (BLUE-Q), was developed for health care professionals [43,44]. These confirm the limited validated tools for assessing blended teaching in health professions education and the underuse of CI. Thus, Blended teaching assessment tool (BTAT) was developed to evaluate the quality of blended teaching approach conducted in health science courses.

BTAT followed a tool development model that involves 3 different phases (Fig. 1) including phase 1's item development; phase 2's scale development; and phase 3's tool evaluation [45,46]. Item development employed literature

Corresponding author's email address:
mtbdalusong@plm.edu.ph
Keywords: Cognitive Interview, Blended Learning, Assessment tool, Content validity
Date submitted: November 7, 2024
Date accepted: June 17, 2024

review and focus group discussions (FGDs) with experts and students followed by experts' content validation. Phase 1's literature review conceptualized quality blended teaching as the alignment, suitability and integration of teaching skills or practices while employing the appropriate technology within a combined online and face-to-face instructional delivery to optimize learning experiences and effectively attain student outcomes. Content and thematic analysis of literature and FGDs resulted to the development of the initial draft of BTAT with 156 key indicators for quality blended instructional approach within 14 dimensions. This went through experts' content validation using 3 rounds of Delphi approach and CVI analysis that reduced BTAT's items to 82.

Phase 2's scale development involved pre-testing with the target population followed by item reduction through factor analysis. This study focuses on the pre-testing phase using CI as a means to validate BTAT and reduce response error. The CI determined relevance, comprehension, and comprehensiveness issues that arise in each item and was used to refine BTAT before it was administered to a large group. This phase ensured and supported the content validity of BTAT based on target participants' understanding of item statements, perceptions and experiences of quality blended instructional delivery. Results presented in this study covered only one of the rigorous steps of tool development process, whose overall goal is to develop a valid and reliable instrument that will measure quality blended teaching in health professions education.

Methodology

2.1 Study design

This study phase employed a qualitative narrative design to enhance the content validity of the BTAT through CI s. The initial BTAT (Appendix I), developed after content validation by 16 experts over three Delphi rounds, consisted of 8 dimensions and 82 items. During the CI phase, target participants were interviewed to assess their cognitive processes while responding to the initial BTAT items, following Tourangeau's framework [7].

2.2 Ethics Approval:

The protocol has been granted approval and 'exemption' (protocol number UPMREB 2022-0259-EX) from ethics review evaluation by the University of the Philippines Manila's Research Grants Administration Office (RGAO). Those interested to participate who answered the google form informed consent were invited for an orientation on the CI process, instructions, and details of ethics guidelines such as confidentiality, anonymity, data privacy, voluntary participation and right to withdraw, and scheduling the conduct of CI.

2.3 Participants:

Purposive sampling was used to invite eligible participants such as university students currently enrolled in health science programs implemented in blended approach. Their program or department heads recommended about five students each. An invitation letter, flyer, and informed consent form were emailed. Of the invited students, 12 (4 Physical Therapy, 4 Nursing, 1 Psychology, 2 Speech Pathology, 1 Occupational Therapy) attended the orientation on the CI process, but only 8 (6 female, 2 males; mean age = 20) completed it.

2.4 Data Collection Method:

Using the CI approach, the interviews were conducted and recorded via Zoom with 8 health science students (physical therapy, speech therapy, occupational therapy, psychology). The first two sessions had 3 students each, while the last 2 sessions had one participant. The primary investigator (MTD) conducted the interviews, with another team member (RA, VCD) taking notes using a common template.

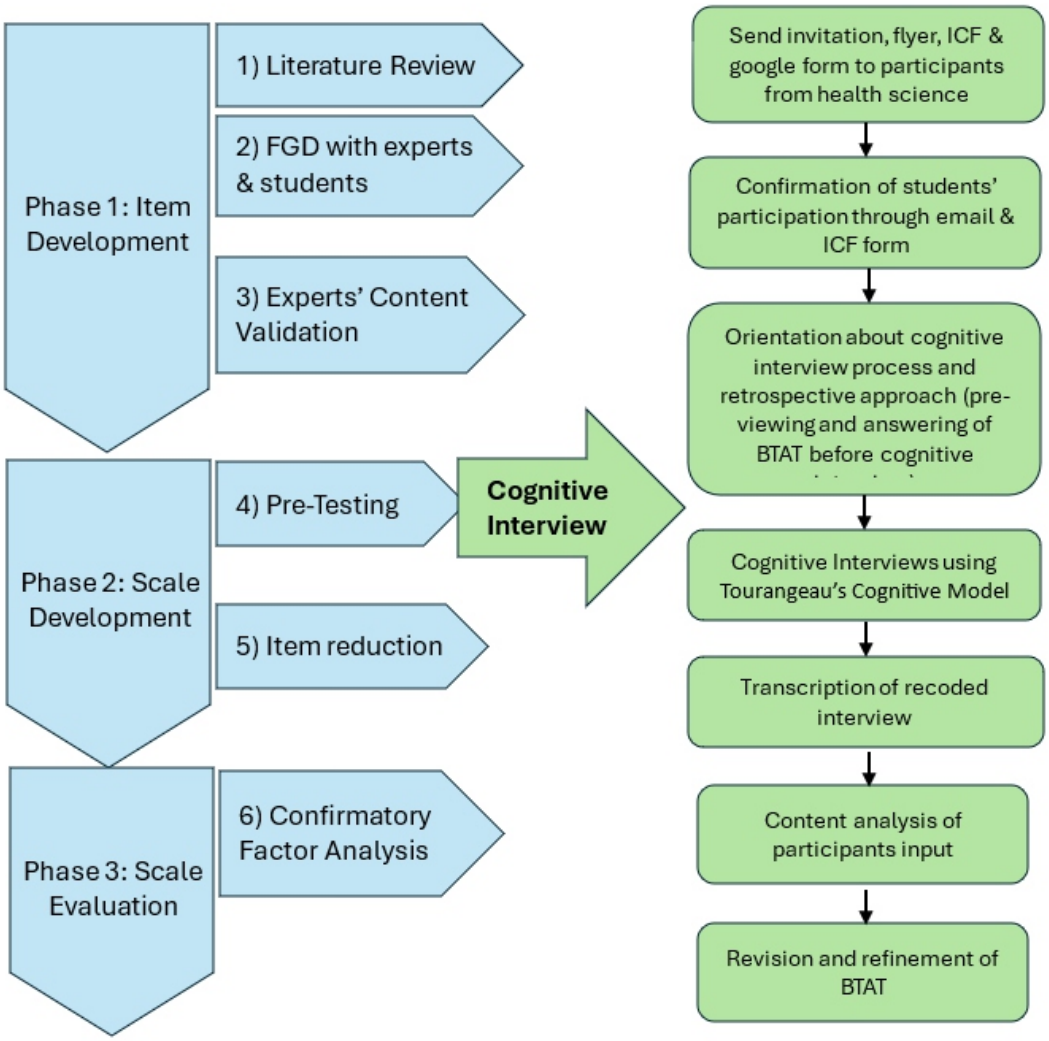

Figure 1. BTAT Development Phases and Steps and Cognitive Interview Process

Table 1. Cognitive interview guide based on Tourangeau's Cognitive Model [7].

Cognitive interview process	Objectives	Sample questions
Comprehension	Check participants' understanding or perception of the item (comprehension) Determine the meaning of some words or phrases in the items	How do you understand this item? What do you think this item talks about? What thought came to mind when you read this item? Are there any difficult/confusing words or phrases in the item? What does it mean to you? What words do you suggest so that you can understand it easily?
Recall or retrieval	Determine participants' ability to recall experience related to the item to respond to it. (relevance to their experience). Obtain participants' comprehensiveness of the items in each domain.	What experience do you easily remember when you read this item? Is it easy for you to recall an experience related to this item? Are the items complete to assess the dimension? Do you think there are still lacking for the dimension?
Judgement	Determine if participants can easily and confidently respond or not to each item	Is it easy for you to select or decide on an answer from the choices? Why?
Response	Determine how participant chooses their response and distinguish between choices	Why did you choose this answer? How do you distinguish them from other choices?

In retrospective approach, students initially answered the BTAT followed by CI using a semi-structured interview questionnaire guide (Table 1). Verbal probing approach explored participants' understanding of each item and its relevance to their recalled experiences, with spontaneous probe questions to clarify their understanding on item or words that the participants found confusing.

Since there were 82 items reviewed during the CI, 2 separate sessions of interviews were conducted to the groups with 3 participants to avoid fatigue. The 4 CIs for the first 2 groups lasted for 2 - 3 hours per session with in between breaks. While the last 2 scheduled CIs were attended only by one participant each and managed to complete it for 2 - 2 ½ hours per session with breaks in-between as preferred by the participants.

The recorded sessions were transcribed and were independently analyzed per session by 2 research members. After the analysis of sessions, the 2 researchers met to review and compare their initial findings. Saturation point was achieved when no other similar issues, comments or suggestions emerged per item on the last sessions of CI.

2.5 Data Analysis

Transcripts were combined and analyzed per item. Two researchers (MTD & GSA) independently coded each item using Tourangeau's four steps of cognitive operation: 1) comprehension, 2) retrieval, 3) judgment, and 4) response [7]. This framework helped identify and revise issues for each item. After independent analysis, the two researchers reviewed and compared their findings. A third researcher with a PhD and expertise in BTL delivery was consulted to resolve disagreements and refine the items with issues.

To ensure validity of the analyzed results, member checking was employed where revised BTAT in google format with face validation questions was sent back to the 8 CI participants. Three participants answered and returned the form with 3 items suggestions on typographical error and grammar.

Results

3.1 Comprehension Issues

The comprehension stage was the initial step in the cognitive process where participants interpreted the statements before recalling their relevant experiences. Researchers assessed whether participants understood the statement's intent. Analysis showed that certain BTAT items had comprehension issues due to:

3.1.1 Unfamiliar words or statements. Some participants found words whose meaning they were not familiar with (e.g. "seamlessly", "assessment tools", "rapport", "formative and summative", "mechanism and system" and "fair evaluation practices"). These words were either removed or revised based on the suggestions of participants and agreement of the research team.

3.1.2 Ambiguity (different or unclear interpretation). Ambiguity arose from complex syntax, leading to varying interpretations of some items. For example, item 3, "The LMS is linked to storage applications (e.g., One Drive, Google Drive) that provide sufficient storage for the necessary files uploaded to the cloud," was interpreted by some as student's personal storage provided by the institution, while others viewed it as storage for the LMS. To clarify, "personal" was added before "storage." In item 10, "The recorded videos have good visual quality," one participant interpreted it as video perspective showing skills demonstration, while another thought it meant 'high-definition (HD) videos.' This was revised to "clear video output and readable text" based on multimedia principles.

3.1.3 Wordy statements. Participants observed that some items contained lengthy statements and suggested making them more concise. For example, item 56, "My teacher shows expertise in the subject matter by clearly explaining and simplifying complex lessons; clarifying confusions and answering questions during online synchronous and/or face-to-face sessions," was simplified by removing the phrase "shows expertise in the subject matter" because the phrases "clearly explaining and simplifying complex lessons, clarifying confusions and answering questions" were specific manifestations of teacher's expertise.

3.1.4 Similar meaning and experiences which can be merged. Participants noticed some overlapping items that they suggested to be integrated as one item. For example, items 37: "My teacher appropriately uses a variety of educational technologies, software application and instructional media that facilitate my learning during synchronous or asynchronous sessions" and item 39: "My teacher uses educational technology or instructional media (e.g., recorded videos) that encourages me to engage independently in my learning and think critically" which both concern teacher's appropriate use of diverse educational technology and instructional media to facilitate learning, were merged. For students, "facilitate learning" was synonymous to independent learning, while "critical thinking" was described as "thinking deeply or actively, especially when answering exams." The terms "facilitate learning," "independent learning" and "critical learning" were simplified as "to engage me in my learning" which focuses more on the general purpose of the educational technology and instructional media. These items were merged and revised with additional examples as "My teacher appropriately uses a variety of educational technologies (LMS, interactive TV, 3D models, simulations, Kahoot, Padlet, & other hard/software applications) and instructional media (PowerPoint, videos) to engage me in my learning during synchronous or asynchronous sessions." Researchers' analysis of most items with similar narration of experiences were merged as one item. Similarly, dimensions 2 (course syllabus), dimension 3 (course module), and dimension 4 (course orientation) were merged as one dimension as course direction and pathway. While the original dimension 7 (quality assessment) and dimension 8 (quality feedback) were merged as one dimension (teacher's evaluation of student's performance) including some similar items.

3.2 Retrieval issues

Retrieval issues occur when the items presented to the participants fail to recall related experiences. Both lack of understanding and lack or insufficient experience, even if cues or examples are given, can affect retrieval of relevant experiences which influenced their responses. For certain items, participants indicated that they had not personally experienced them but recognized their importance in BTL. Consequently, they recommended that these items be retained in the assessment tool. These are the specific recall issues mentioned by participants:

3.2.1. Inconsistencies with BLT experiences. Participants shared their

Table 2. Sample items revised

Original Items with comprehension issues	**Revised statements**
Item 43. My teacher builds engaging connections and fosters respectful or appropriate **rapport** with students in both online and face-to-face classes.	My teacher builds engaging connections and respectful or appropriate **communication** with students in both online and face-to-face classes.
Item 46. My teacher regularly monitors and **provides feedback** and reminder on the completion of asynchronous and face-to-face activity requirements.	My teacher regularly monitors and **provides reminders on the status of completion** of asynchronous and face-to-face activity requirements.
Item 56. My teacher shows expertise in the subject matter by clearly explaining and simplifying complex lessons; clarifying confusions and answering questions during online synchronous and/or face-to-face sessions.	My teacher clearly or concisely explains and simplifies complex lessons; clarifies confusions and answers questions during synchronous and/or face-to-face sessions.
Original items with retrieval issues	**Revised statements**
Item 41: My teacher permits access to recorded synchronous meetings within a specific timeframe (e.g. weeks or 1 - 2 months accessible) in compliance with school's policies, and when conditions like cancellation of classes, school events, or other related factors occur.	My teacher permits access to recorded synchronous meetings.
Original item with judgement issue	
Item 3: The LMS is linked to storage applications (e.g., One Drive, Google Drive) that provide sufficient storage for the necessary files uploaded to the cloud.	The LMS provides sufficient personal storage (e.g., One Drive, Google Drive) for the storing academic files in the cloud.

Appendix II shows the complete revisions of items per cognitive process.

real-life experiences which were related but not exactly the way some items stated. For instance, they greatly valued the availability of recorded lecture videos and synchronous meetings in the BTL, as it allowed them to review the sessions. In item 41 stating, "My teacher permits access to recorded synchronous meetings within a specific timeframe in compliance with school policies," participants noted that recordings are often provided regardless of conditions like class cancellations. Most students found this item important and relevant even during normal conduct of combined online and face to face set up, leading to its revision as "My teacher permits access to recorded synchronous meetings."

3.2.2. Recall difficulties. These items had recall difficulties because of limited examples provided, and few experiences. For item 29 about orientation on "Technology requirements such as minimum gadget/device specifications, operating system, data consumption, applications, or software," some participants didn't recall their teachers discussing these requirements because they never experienced it. Others remembered mentions of minimum requirements and new applications but not the details. Despite their limited experience, participants found this topic relevant, especially with new software. One suggested including the number of gadgets needed, as they experienced during online quizzes. Ultimately, participants agreed to retain this item, as supported by the research team. Examples were also provided for other items with similar issues.

3.2.3. Irrelevant to experience. For item 23, "The course modules motivate me to prepare and take responsibility for my own learning process or involvement before the actual class," participants agreed on the importance of providing a module. Some participants expressed that there are other factors that motivates them to study, not necessarily the presence of the module. One commented that once the module is provided, they just take on the tasks of studying it for the next class. Researchers revised the item to, "My teacher provides clear direction or instructions in a course module or unit/topic activity guide that helps me prepare well and take responsibility for my own learning before the actual classes."

3.3 Judgement

Issues on judgement happen when participants have trouble evaluating and deciding on their responses. Judgment issues arose particularly in item with two different considerations in one statement. For instance, in Item 3, which states, "The LMS is linked to storage applications (e.g., One Drive, Google Drive) that provide sufficient storage for the necessary files uploaded to the cloud," some participants found the first part relevant but failed to consider the second part regarding storage sufficiency.

3.4 Response

Response issues arose when participants struggled to choose an answer from the provided scale. Some students suggested including a "not applicable" as one of the options since some items were not experienced in BTL, arguing that it might be unfair to rate a teacher poorly for items not implemented. However, researchers decided against including "not applicable" as the items are seen as quality indicators in the BTL approach.

For almost all items, participants did not have problems with the response choices or scale provided. The university students explained their responses based on their experiences and were able to clearly differentiate between the response categories, from strongly agree to strongly disagree.

3.5 Overall BTAT Survey

Participants found the tool comprehensive, comprehensible, and relevant for evaluating instructional quality in the BTL approach in a university setting, appreciating the clear instructions and user-friendly Google Form format. Although they felt the initial 82 items of the BTAT were lengthy, the clustering by dimension helped them relate their experiences better. While they recognized the BTAT as an effective assessment instrument, they expressed concerns about it being overwhelming if used for evaluating all their teachers at the semester's end. They recommended shortening the survey, suggesting that some items and dimensions could be combined.

The CI provided significant information for the revision of BTAT. From the initial 8 Dimensions with 82 items, the revised BTAT was reduced to 5 Dimensions with 53 items (Appendix III on revised BTAT).

Discussion

The BTAT is designed to evaluate the quality of instructional delivery in the BTL approach. Content validity ensures that the tool's items accurately measure the relevant constructs [45]. Although experts have validated the tool, conducting CI with target participants adds an essential layer of validation and improves validity of the university students' responses. Existing literature emphasizes expert validation while underrepresenting participant input [33]. However, incorporating participant feedback is vital, as it provides insights into the constructs and items,[3,33,47] which ultimately strengthens content validity [45].

The CI is a qualitative method in scale development that assesses whether participants capture the intended meaning of each item and evaluates the rationale behind their responses [48,49]. Using Tourangeau's (1984) cognitive interview model, researchers identified challenges in the four-step cognitive response process: comprehension, memory retrieval, judgment, and selection of response options [7,16]. This approach proved invaluable in understanding how students interpreted the survey questions, helping to identify response errors and improve the BTAT.

There is no established standard for CIs, but it is recommended to have a clear aim, a useful framework, and a structured protocol [8,49-51]. The retrospective interview method, combining scripted questionnaires with spontaneous probing questions, allowed participants to articulate their thought processes, enhancing researchers' understanding of their cognition. This approach has also been used effectively for longer survey items to clarify participants' thoughts during interviews [5,16]. Consequently, researchers gained insights into comprehension levels, as a basis for item refinement.

During the pre-testing stage of the BTAT using CIs, researchers found that university students enrolled in blended health sciences programs could effectively respond to and comprehend most of the items. Similarly, CI studies involving high school and university students indicated higher comprehension compared to younger participants [51-53]. CI also confirmed healthcare students' comprehension and experience in the use of immersive technology for healthcare education measure [14]. Health Science students'

comprehension provided valuable feedback on items that could be merged based on their experiences, improving the BTAT's relevance, comprehensiveness, and comprehensibility.

Researchers identified terms that participants did not understand and items that didn't match their course experiences, hindering their ability to respond effectively. The main issue highlighted was comprehension, particularly with unfamiliar or complex terminology related to educational concepts such as 'formative', 'evaluation' and 'assessment'. Such comprehension challenges are common in CIs and can significantly affect participants' recall and judgment processes [16,51,53,54]. Addressing these issues is essential for improving the BTAT's effectiveness.

Retrieval issues in the BTAT emerged from incongruities in statements relative to participants' experiences, compounded by a lack of familiarity and limited cues. Similar challenges were observed in CIs with students using a well-being instrument [53]. Most retrieval problems stemmed from participants' limited experiences and [51,53], and difficulties recalling specific events from years ago [12,16]. However, recall was less significant for BTAT, as it is administered when blended instruction is implemented.

Judgment issues in BTAT were primarily associated with compound sentence structures experienced differently by participants. Similarly, Sports Psychology students similarly found ambiguous, interchanging and complex phrases like "feeling confident" and "feeling secure" in the Competitive State Anxiety Inventory [52]. Researchers agree that the judgment process depends on having clear and relevant information for assessment [11,50,55].

No issues were noted with BTAT's response choices, as participants could align their answers with response options effectively and meaningfully differentiated their answers [50,55]. While response issues are less frequent than comprehension issues [16,50,55], some studies also reported no issues at this stage of the cognitive process [11,53].

Conclusion

Cognitive interview provided valuable feedback for content validation tool which helped researchers in making improvements to the BTAT that are relevant to the tool's target users. The challenges related to comprehension, retrieval, judgement and response were revealed through the process of CI, making it a valuable step in the process of development of high-quality measurement instruments to evaluate the quality of blended instruction in the health sciences.

Limitation of the study

Some participants coming from similar universities may affect their comprehension and understanding of the items based on similar experiences. Participants who consented to participate in the CIs may not necessarily be the most articulate and critical individuals capable of evaluating the items in the BTAT.

Acknowledgement

The authors thank the health science students who participated in the cognitive interview and their institutions' program directors who facilitated in the recruitment process.

Declaration of interest

The authors declare no conflict of interests.

Funding

This tool development project is funded by the National Research Council of the Philippines but had no role in the study design, data collection or interpretation throughout the project.

Appendices (Retrieval is subject to authors' approval)

Appendix I Original BTAT used for Cognitive Interview
Appendix II Complete revisions of items per cognitive process
Appendix III Revised BTAT after Cognitive Interview

References

1. Bass KM, Drits-Esser D, Stark LA. (2016) A primer for developing measures of science content knowledge for small-scale research and instructional use. CBE Life Sci Educ., 2016;15(2):1–14, doi: 10.4236/psych.2018.911145.
2. Streiner DL, Kottner J. (2014) Recommendations for reporting the results of studies of instrument and scale development and testing. J Adv Nurs., 70(9):1970–9, doi: 10.1111/jan.12402.
3. Carpenter S. (2017) Ten Steps in Scale Development and Reporting: A Guide for Researchers. Commun Methods Meas [Internet], 12(1):25–44. https://doi.org/10.1080/19312458.2017.1396583
4. Kyriazos TA, Stalikas A. (2018) Applied Psychometrics: The Steps of Scale Development and Standardization Process. Psychology., 09(11):2531–60, https://doi.org/10.4236/psych.2018.911145
5. Shiyanbola OO, Bolt D, Tarfa A, Brown C, Ward E. (2019) A content validity and cognitive interview process to evaluate an Illness Perception Questionnaire for African Americans with type 2 diabetes. BMC Res., 12(1):1–6, https://doi.org/10.1186/s13104-019-4342-9
6. Farmer N, Powell-Wiley TM, Middleton KR, Brooks AT, Mitchell V, Troncoso M, *et al*. (2022) Use of a focus group-based cognitive interview methodology to validate a cooking behavior survey among African-American adults. Front Nutr., 9(December):1–14, https://doi.org/10.4236/psych.2018.911145
7. Willis GB. (2004) Cognitive Interviewing Revisited: A Useful Technique, in Theory? Methods Test Eval Surv Quest., (August 2004):23–43, doi: 10.1002/0471654728.ch2.
8. Boeije H, Willis G. (2013) The cognitive interviewing reporting framework (CIRF): Towards the harmonization of cognitive testing reports. Methodology, 9(3):87–95, doi: 10.1027/1614-2241/a000075.
9. Scott K, Ummer O, Lefevre AE. (2021) The devil is in the detail: Reflections on the value and application of cognitive interviewing to strengthen quantitative surveys in global health. Health Policy Plan., 36(6):982–95, doi: 10.1093/heapol/czab048.
10. Balza JS, Cusatis R, McDonnell SM, Basir MA, Flynn KE. (2022) Effective questionnaire design: How to use cognitive interviews to refine questionnaire items. J Neonatal Perinatal Med., 15(2):345–9, doi: 10.3233/NPM-210848.
11. Schildmann EK, Groeneveld EI, Denzel J, Brown A, Bernhardt F, Bailey K, *et al*. (2016) Discovering the hidden benefits of cognitive interviewing in two languages: The first phase of a validation study of the Integrated Palliative care Outcome Scale. Palliat Med., 30(6):599–610, doi: 10.1177/0269216315608348.
12. Kamp K, Wyatt G, Dudley-Brown S, Brittain K, Given B. (2018) Using cognitive interviewing to improve questionnaires: An exemplar study focusing on individual and condition-specific factors. Appl Nurs Res., 43:121–5, doi: 10.1016/j.apnr.2018.06.007.
13. Sigl H, Kriston L, Scholl I, Härter M, Hahlweg P. (2025) Response process validity of the 9-item shared Decision-Making Questionnaire (SDM-Q-9) in a cognitive interview study with patients with cancer. Sci Rep., 15(1):1–9, doi: 10.1038/s41598-025-99640-2.
14. Jacobs C, Wheeler J, Williams M, Joiner R. (2023) Cognitive interviewing as a method to inform questionnaire design and validity - Immersive Technology Evaluation Measure (ITEM) for healthcare education. Comput Educ X Real [Internet]., 2(March):100027, https://doi.org/10.1016/j.cexr.2023.100027.
15. Teal R, Enga Z, Diehl SJ, Rohweder CL, Kim M, Dave G, *et al*. (2015) Applying cognitive interviewing to inform measurement of partnership readiness: A new approach to strengthening community–academic research. Prog Community Heal Partnerships Res Educ Action., 9(4):513–9, doi: 10.1353/cpr.2015.0083.
16. Spark MJ, Willis J. (2014) Application of cognitive interviewing to improve self-administered questionnaires used in small scale social pharmacy research. Res Soc Adm Pharm., 10(2):469–74, http://dx.doi.org/10.1016/j.sapharm.2013.06.007.
17. Shin S, Jung D, Kim S. (2015) Validation of a Clinical Critical Thinking Skills Test in Nursing. J Educ Eval Heal Prof., 12:1–6, doi: http://dx.doi.org/10.3352/jeehp.2015.12.1
18. Pajanel MR, Gonzales IJ. (2025) Validation of the Filipino Version of the Diabetes Distress Scale for Adult Patients with Diabetes seen at the Outpatient Department of a Tertiary Government Hospital in Quezon City , Philippines. J ASEAN Fed Endocr Soc.,

40(1):61–71, doi: https://doi.org/10.15605/jafes.040.01.05.
19. Melgar MIE. (2022) Translating COGNISTAT and the Use of the Cognitive Interview Approach: Observations and Challenges. Madridge J Behav Soc Sci., 5(1):89–95, doi: 10.18689/mjbss-1000116.
20. Ablao CJ, Diaz DC, Macapagal J. (2022) Challenges of Distance Education Assessment in the Health Professions during the COVID-19 Pandemic: A Philippine Reflection in the Rapid Review of International Context. Acta Med Philipp., 56(8):49–57, doi: 10.47895/amp.vi0.3145.
21. Baticulon RE, Sy JJ, Alberto NRI, Baron MBC, Mabulay REC, Rizada LGT, *et al*. (2021) Barriers to Online Learning in the Time of COVID-19: A National Survey of Medical Students in the Philippines. Med Sci Educ., 31(2):615–26, https://doi.org/10.1007/s40670-021-01231-z.
22. Rabe A, Sy M, Cheung WYW, Lucero-Prisno DE. (2020) COVID-19 and Health Professions Education: A 360° View of the Impact of a Global Health Emergency. MedEdPublish., Jul 17;9:148, doi: 10.15694/mep.2020.000148.1.
23. Lavadia WT, Sana EA, Salvacion MLDS. (2022) Smart phone chat apps for teaching and assessment in Orthopaedic Residency Training. Malaysian Orthop Journal, Off Press., 17(1):45–60.
24. Fontanilla RC, Guzman CR. (2023) The Quality of Online Learning Delivery and The Learning Management System of The Medical Colleges of Northern Philippines (MCNP) And International School of Asia and The Pacific (ISAP): A Correlational Study. Seybold Rep ISSN., (September):1533–921, doi: 10.17605/OSF.IO/P4NX8.
25. Ashraf MA, Yang M, Zhang Y, Denden M, Tlili A, Liu J, *et al*. (2021) A systematic review of systematic reviews on blended learning: Trends, gaps and future directions. Psychol Res Behav Manag., 14(October):1525–41, doi: 10.2147/PRBM.S331741.
26. Smith K, Hill J. (2019) Defining the nature of blended learning through its depiction in current research. High Educ Res Dev., 38(2):383–97, https://doi.org/10.1080/07294360.2018.1517732.
27. Mirriahi N, Alonzo D, Fox B. (2015) A blended learning framework for curriculum design and professional development. Res Learn Technol., 23(1063519), doi: 10.3402/rlt.v23.28451.
28. Joaquin JJB, Biana HT, Dacela MA. (2020) The Philippine Higher Education Sector in the Time of COVID-19. Front Educ., 5, https://www.frontiersin.org/articles/10.3389/feduc.2020.576371.
29. Cobo-Rendón R, Bruna Jofre C, Lobos K, Cisternas San Martin N, Guzman E. (2022) Return to University Classrooms With Blended Learning: A Possible Post-pandemic COVID-19 Scenario. Front Educ., 7., https://www.frontiersin.org/articles/10.3389/feduc.2022.957175.
30. Dayagbil FT, Palompon DR, Garcia LL, Olvido MMJ. (2021) Teaching and Learning Continuity Amid and Beyond the Pandemic. Front Educ., 6, https://www.frontiersin.org/articles/10.3389/feduc.2021.678692
31. Tupas FP, Linas-Laguda M. (2020) Blended learning - An approach in Philippine basic education curriculum in new normal: A review of current literature. Univers J Educ Res., 8(11):5505–12, doi: 10.13189/ujer.2020.081154.
32. Dones VC, Dalusong MTB, Manlapaz D, Rojas JA, Ho KC, Reyes JJ, *et al*. (2024) Evidence on psychometric properties of self-report questionnaires in evaluating blended learning in health sciences university students: research protocol for systematic review and meta-analysis. BMJ Open., Aug 1;14(8):e075266, http://bmjopen.bmj.com/content/14/8/e075266.abstract
33. Morgado FFR, Meireles JFF, Neves CM, Amaral ACS, Ferreira MEC. (2017) Scale development: Ten main limitations and recommendations to improve future research practices. Psicol Reflex e Crit., 30(1):1–20, doi: 0.1186/s41155-016-0057-1.
34. Dodeen H. (2013) College students ' evaluation of effective teaching : Developing an instrument and assessing its psychometric properties. Res High Educ J., 21:1–12.
35. Bangert AW. (2008) The development and validation of the student evaluation of online teaching effectiveness. Comput Sch., 25(1–2):25–47, doi: 10.1080/07380560802157717.
36. Matosas-López L, Aguado-Franco JC, Gómez-Galán J. (2019) Constructing an Instrument with Behavioral Scales to Assess Teaching Quality in Blended Learning Modalities. Vol. 8, JNew App Ed Res., 142–65, doi: 10.7821/naer.2019.7.410.
37. Terwee CB, Prinsen CA, Chiarotto A, Cw De Vet H, Bouter LM, Marjan JA, *et al*. (2018) COSMIN methodology for assessing the content validity of PROMs: User manual. Circulation [Internet]. 120(9):0–70. Available from: www.cosmin.nl
38. Martín-martínez L, Rodríguez-legendre F, Vitoria UF De, Management K, International EA, Km J, *et al*. (2020) Evaluation of a blended learning model for pre-service teachers. Knowl Manag E-learning, 12(2):147–64, doi: 10.34105/j.kmel.2020.12.00.
39. Panigrahi R, Nihar KL, Singh N. (2024) Quality Measurement of the Blended Learning Model in Higher Education: Scale Development and Validation. High Learn Res Commun., 14(1):86–102, doi: 10.18870/hlrc.v14i1.1467.
40. Zhao Y, Li W, Jiang H, Siyiti M, Zhao M, You S, *et al*. (2024) Development of a blended teaching quality evaluation scale (BTQES) for undergraduate nursing based on the Context, Input, Process and Product (CIPP) evaluation model: A cross-sectional survey. Nurse Educ Pract., 77(February):103976, https://doi.org/10.1016/j.nepr.2024.103976
41. Han X. (2023) Evaluating blended learning effectiveness: an empirical study from undergraduates' perspectives using structural equation modeling. Front Psychol., 14(May):1–13, doi: 10.3389/fpsyg.2023.1059282.
42. Medallon MC. & Dimaculangan GA. (2022) SET PH Fac performance Eval in Flexible Env Scale Dev .pdf. PAPSCU Excell Acad Res Link Bull., 3(1).
43. Arora AK, Myers J, Apramian T, Kulasegaram K, Bainbridge D, Seow H. (2025) Reliability and construct validation of the Blended Learning Usability Evaluation–Questionnaire with interprofessional clinicians in Canada: a methodological study. J Educ Eval Health Prof., 22:1–8, doi: 10.3352/jeehp.2025.22.5.
44. Arora AK, Rodriguez C, Carver T, Zhang H, Schuster T. (2024) Validation of the blended learning usability evaluation – questionnaire (BLUE-Q) through an innovative Bayesian questionnaire validation approach: a methodological study. J Educ Eval Health Prof., 21:31, doi: 10.3352/jeehp.2024.21.31.
45. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. (2018) Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Front Public Heal., 6(June):1–18, doi: 10.3389/fpubh.2018.00149.
46. DeVellis RF. (2012) Scale Development: Theory and applications. 3rd ed. London: Sage Publication.
47. Willis G. (2012) Cognitive Interviewing Training Guide., 56. https://www.researchgate.net/publication/241686481_Cognitive_Interviewing Consultado em 12nov2023, 11:10
48. Kalkbrenner MT. (2021) A Practical Guide to Instrument Development and Score Validation in the Social Sciences: The Measure Approach. Pract Assessment, Res Eval., 26:1–18, doi: https://doi.org/10.7275/svg4-e671.
49. Massey M. (2016) Analysis of Cognitive Interview Testing of Questions on Cognitive Functioning.
50. Beatty PC, Willis GB. (2007) Research synthesis: The practice of cognitive interviewing. Public Opin Q., 71(2):287–311, doi: 10.1093/poq/nfm006.
51. Patel-Syed Z, Becker S, Olson M, Rinella H, Scott K. (2024) What do you think it means? Using cognitive interviewing to improve measurement in implementation science: description and case example. Implement Sci Commun., 5(1), doi: https://doi.org/10.1186/s43058-024-00549-0
52. LaPietra E, Brown Urban J, Linver MR. (2020) Using Cognitive Interviewing to Test Youth Survey and Interview Items in Evaluation: A Case Example. J Multidiscip Eval., 16(37):74–96, doi: 10.56645/jmde.v16i37.651.
53. Dietrich H, and Ehrlenspiel F. (2010) Cognitive Interviewing: A Qualitative Tool for Improving Questionnaires in Sport Science. Meas Phys Educ Exerc Sci., Jan 21;14(1):51–60, https://doi.org/10.1080/10913670903455025.
54. Susanto H, Wedyaswari M, Dalimunthe KL. (2023) A content validity and cognitive interview to develop the HARMONI items: Instrument measuring student well-being in West Java, Indonesia. Front Educ., 8(March):1–9, doi: 10.3389/feduc.2023.1132031.
55. McClam M, Workman L, Dias EM, Walker TJ, Brandt HM, Craig DW, *et al*. (2023) Using cognitive interviews to improve a measure of organizational readiness for implementation. BMC Health Serv Res. V:23(1):1–10, https://doi.org/10.1186/s12913-022-09005-y
56. Ryan K, Gannon-Slater N, Culbertson MJ. (2012) Improving Survey Methods With Cognitive Interviews in Small- and Medium-Scale Evaluations. Am J Eval., 33(3):414–30, doi: 10.1177/1098214012441499.