International Journal of Evaluation and Research in Education (IJERE)
Vol. 13, No. 6, December 2024, pp. 3741~3751
ISSN: 2252-8822, DOI: 10.11591/ijere.v13i6.29443

Design and validation of research capability test for educators

Jay-cen Tamayo Amanonce[1,2], Rudolf Tamangen Vecaldo[1,2], Maribel Fontiveros Malana[1], Conchita Malenab Temporal[1,2], Antonio Ingaran Tamayao[1,2], Jhoanna Battung Calubaquib[1,2], Marie Claudette Maramag Calanoga[1,2], Ria Arellano Tamayo[1,2]
[1]College of Teacher Education, Cagayan State University-Andrews Campus, Tuguegarao City, Philippines
[2]Graduate School, Cagayan State University-Andrews Campus, Tuguegarao City, Philippines

Article Info

Article history:

Received Dec 1, 2023
Revised Apr 8, 2024
Accepted May 7, 2024

Keywords:

Educators
Reliability
Research capability
Research tool
Validity

ABSTRACT

Assessing the research capability of educators yields implications for bridging the gap and addressing concerns along research. Accordingly, the Commission on Higher Education (CHED) highlighted substantive research concerns which need to be given attention specifically in enhancing research capabilities of graduate students, research staff, and faculty which calls for a well-designed and validated research capability test (RCT). Thus, this study established the RCT's validity and reliability to gauge the knowledge of educators in conceptualizing, implementing, and publishing research. The developed RCT underwent series of critiquing, content validation from experts, and internal and external pilot testing. The study has generated 80 items with desirable difficulty index (50.76), reasonably good discrimination index (0.28), and large functioning distractors (78.97% distractor efficiency). Remarkably, there is a significant association between distractor efficiency and difficulty index of the items with an inter-item consistency of α=0.904. Thus, the RCT is a valid and reliable instrument to gauge the research capability of educators.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Jay-cen Tamayo Amanonce
College of Teacher Education, Graduate School, Cagayan State University – Andrews Campus
Caritan, Tuguegarao City, Cagayan, 3500, Philippines
Email: jaycen.amanonce@gmail.com

1. INTRODUCTION

The indispensability of research in academic institutions is beyond doubt as global demands for university ranking performance mark it a key criterion for quality assurance. As Sarwar [1] and Dowsett [2] posit the phenomenon of a new era in higher education in which university ranking systems heavily marked by research performance have assumed an importance and have changed the landscape of higher education which is likely to continue to influence further development internationally and nationally. These rankings are often the go-to resource for researchers, students, potential faculty members, funders, parents, and other stakeholders who seek an objective insight into the university's performance and ability [3]–[6]. The aspiration for world-class universities is espoused with stringent measures to produce research assessment indicators and universities around the world have never been pushed as much as today by worldwide rankings [7], [8].

Outcome studies and quality improvement initiatives in the research enterprise are urgently needed, given the importance of university innovation based on research performance. This paves for the remarkable and overwhelming attention among academicians and practitioners on research capability which is an expedient response as regulating bodies such as Commission on Higher Education (CHED) and Department of Education (DepEd) encourage educators to engage in research to support evidence-based decision-making, practice, program, and policy development apart from the teachers' need to progress personally and professionally [9]–[11].

For higher education institutions, CHED is accountable for the advancement of research functions in the Philippines, as stated in Section 8 of Republic Act (RA) No. 7722 [12]. CHED is also directed to facilitate expertise sharing and collaboration between and among higher education institutions (HEIs) as it emphasizes the role of quality academic research which are practical and of immediate usefulness. However, the Commission takes a position that in general, HEIs in the Philippines are continually facing pressing concerns at various stages of the research enterprise. As such, it highlighted substantive concerns that need to be addressed along research such as but not limited to enhancing research capabilities, building up generation of researcher, instilling a research culture, raising research quality and impact, and increasing research productivity [13].

In the context of teacher education institutions (TEIs), among the four-fold functions of teachers namely instruction, extension, research, and production, it is research that is considered as a crucial criterion for the grant of center of excellence (COE) and center of development (COD) status to the TEIs' curricular programs. Research and publication have the second highest percentage (30%) in evaluating TEIs for COE and COD, next to instructional quality (45%), followed by extension and linkages (20%), and institutional qualification (5%). It is indicated that at least 50% of the faculty should have published researches in reputable journals. Also, at least three institutional researches should have been conducted which shall be used for institutional policy development and program improvement [14]. Furthermore, local and international accrediting agencies consider research development as a key performance indicator for quality assurance. Thus, the capability of the faculty to conduct research is an advantage in ensuring research-based and quality extension and instruction [11].

Premised on the cited context, research indubitably plays important roles both in instruction and extension. Effectiveness and efficiency of teachers' teaching strategies and practices, instructional materials, and assessment tools could be determined and monitored through educational research. Students' learning styles, interests, attitudes, self-efficacy, motivations, and achievements towards learning could also be ascertained through research. Expressly, teachers utilize research as a source of new concepts and advancements to develop worthy teaching-learning strategies and practices [15]. Ultimately, research results serve as bases in improving the quality of instruction offered to students. The importance of research is also realized in extension or community services as it serves its purpose in finding solutions to societal problems and eventually improve a community or the quality of people's lives.

The emphasis on research capabilities has garnered a great deal of attention in the literature during the previous few decades due to the present emphasis on the importance of research. Since DepEd institutionalized research, the majority of studies identified the research capability of teachers inside the department [14], [16], [17]. There have been numerous foreign and local studies on research capability of teachers in the different levels of education [11], [17]–[19]. However, those studies are pure perceptual and survey type. Whereas, the present study dealt with the development and validation of research capability tool which will be used to test the research capability of educators.

Though the significance of research in education is evident, there is a dearth of available validated research capability assessment tools particularly an objective test gauging educators' knowledge, skill, and attitude in conceptualizing, implementing, and publishing research. It is necessary to develop research capability assessment tools to determine the gaps in research among educators. Determining the capability of teachers is a direct way of assessing their needs along research. It is on this premise that this study was conducted. Its specific objectives were to: i) design and construct a research capability test (RCT) for educators that is gender-sensitive, criterion-referenced, and contextualized; ii) determine the RCT's difficulty, discrimination, and distractor indices and efficiencies; iii) ascertain the relationship between the RCT indices; and iv) establish its reliability.

Along with instruction, extension and production, research is one of the core functions of a HEI. The capability and competency of an instructor to do research is expedient in addressing the most pressing societal concerns and issues, starting in the classroom. Hence, this study developed a well-designed and validated RCT to measure the instructors' competencies and capabilities in constructing a good title; in making logical and organized abstract and introduction; in examining the measurability of the specific research questions; in designing a research paradigm; in organizing a thematic presentation of literature review; and in employing hypothesis testing procedures, appropriate research design, method, data gathering technique, validation procedures and statistical tools. Further, this RCT also measures the instructors' capability in presenting results of data analyses through tables and graphs; in writing a discussion of research findings; in writing a brief, concise, impactful, and rousing conclusion and recommendation for a research study; and in employing the APA style format in writing references and in-text citations. In this study, the RCT items were fundamentally crafted based on the research competencies which aim to determine the teachers' capability to conceptualize, implement, and publish research. The RCT underwent rigid validation to ensure its appropriateness and comprehensibility. This tool measures the knowledge, skill, and attitude of teachers in dealing with research. The study's conceptual paradigm is shown in Figure 1.

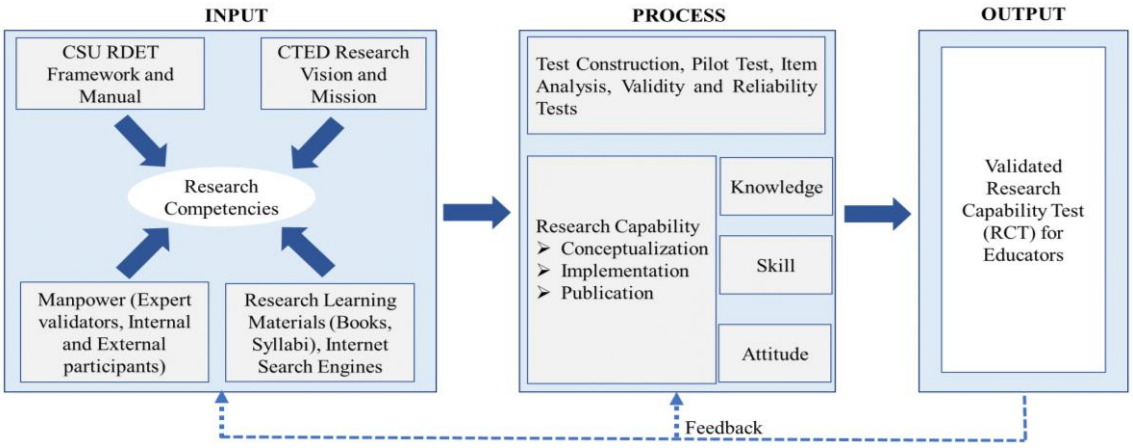


Figure 1. A paradigm showing the input-process-output model of the study

2. METHOD

2.1. Research design

Quantitative design was employed in the study since its main focus is to establish the credibility of the developed RCT through numerical validation procedures. Specifically, the difficulty index, discrimination index, distractor efficiency, and reliability coefficient of the RCT were ascertained. These indices and values are substantial in establishing the efficiency of the RCT as a criterion-referenced test.

2.2. Sampling procedure and research participants

Purposive sampling was utilized in the first and second pilot tests to determine the participants. The participants were selected considering the following inclusion criteria: i) They have at least one year of teaching experience in the university; ii) handle general education or teacher education courses; iii) they are permanent faculty members; and iv) willingly participate in the study.

To determine the sample size for the study, the researchers focused on faculty members from public HEIs in the Cagayan Valley Region, Philippines. In the first pilot test, the study included 52 teachers from the province of Cagayan as internal participants. This initial sample size aimed to provide a sufficient representation of faculty members within the region and gather diverse perspectives from a sizable group. For the second pilot test, the study selected 34 teachers from the province of Isabela as external participants. This additional sample aimed to diversify the study's sample pool by including faculty members from a different province within the same region. In brief, the sample size determination is guided by the need to encompass a representative sample of faculty members from public HEIs in the Cagayan Valley Region, with the specific numbers chosen based on considerations such as feasibility, resource availability, and the desired level of statistical power for the study [20].

To control the confounding variables, random assignment of respondents to different academic ranks was done. This move ensured that the academic ranks from instructors to professors were proportionately distributed. By doing so, the study established that the respondents' answers in the tests could be more confidently attributed to their knowledge in research rather than external factors related to academic ranks.

2.3. Research instrument

The first version of the research capability pilot test was a 120-item team-made tool designed to evaluate individual research capability of teacher educators. Each multiple-choice item has four options. On the other hand, the second version of the tool for RCT has 100 items after the first item analysis and the third version and final form of the RCT has 80 items. The instrument underwent face and content validation which resulted to an overall scale (S-CVI) of 0.956 suggesting that the items are highly relevant to the identified competencies. In addition, the reliability of the instrument was evaluated based on its internal consistency which resulted to Cronbach's alpha value of 0.904 indicating that the items are consistent in measuring the intended constructs.

2.4. Research procedure

The development and validation of the research capability test (RCT) involved several crucial stages. Initially, the planning phase involved identifying key competencies essential for evaluating educators' research capability, drawing from syllabi, textbooks, and other educational materials. A table of specifications was

created, and a comprehensive literature review was conducted to generate RCT items, resulting in a shortlist of 120 questions based on research competencies. Thereafter, content validation was performed by a panel of four experts with extensive qualifications in research and teaching. These experts assessed the items for face validity, clarity, and relevance, leading to the retention of 99 items and revision of 21 others based on their feedback.

Subsequently, the test underwent a first try-out administered to internal teacher-participants, with data gathered used for initial item analysis. This analysis evaluated each item's difficulty and discrimination indices, guiding revisions to ensure only desirable items covering specific competencies were retained. Following this, a second try-out test was conducted among teacher education faculty from another institution to further examine validity and reliability.

Similarly, the results of the second try-out test were subjected to second item analysis which employed criteria and methods similar to the first, and the Cronbach's alpha coefficient was computed to establish internal consistency. Based on the findings of the second item analysis, the test items underwent further revision, focusing on improving distractors and modifying stems and options where necessary. Finally, before reproduction, the RCT underwent expert proofreading, layout adjustments, and design enhancements to finalize the test for use in evaluating educators' research capability.

2.5. Ethical considerations

Approval was sought from the campus executive officers for the pilot testing of the RCT, indicating that the study underwent an ethical review process. Participants were provided with written information detailing the purpose and procedures of the study, ensuring transparency about what their involvement would entail. The concept of free, prior, and informed consent (FPIC) was explained to the participants, indicating that they were given the opportunity to participate voluntarily after fully understanding the implications of their involvement. Consent was obtained from participants during both the first and second pilot tests ensuring participants' autonomy and understanding at different stages of the study. Participants were assured that any information they provided would be handled with utmost confidentiality and anonymity, emphasizing the importance of protecting their privacy rights. Participants were informed that their participation was entirely voluntary as the researchers acknowledge respect to their autonomy and a commitment to avoiding coercion or undue influence.

These measures taken to ensure the validity of the data collected, the ethical practices outlined, such as informed consent, confidentiality, and voluntary participation, contribute to the validity of the data. When participants are fully informed, willingly participate, and trust that their data will be handled confidentially, it enhances the likelihood that the data collected will accurately reflect their experiences and perspectives. Overall, the study demonstrates a commitment to ethical principles in research, particularly regarding participant welfare, autonomy, and data confidentiality, which are crucial for ensuring the validity and reliability of the study outcomes.

2.6. Data analysis

The difficulty index, discrimination index, and distractor efficiency were interpreted using descriptive statistics like mean, standard deviation, frequency count, and percentage distribution. Item difficulty was obtained using Crocker and Algina [21] (Difficulty index=Number of examinees who answered an item correctly divided by overall number of examinees) multiplied by 100. The computed difficulty index was categorized based on the guideline of Gregory [22], namely: i) 86-100% (Very easy); ii) 71%-85% (Easy); iii) 40%-70% (Desirable); iv) 15%-39% (Difficult); and v) 14% and below (Very difficult). To determine the discrimination coefficient for every item, point-biserial correlation was employed. Since this statistical test includes all examinees in the analysis, it is considered to be the most effective way of determining item discrimination [23]. The discrimination index was construed along the following descriptive values: i) greater than +0.30 (Very good); ii) between +0.09 and +0.30 (Reasonably good); and iii) less than +0.09 (Poor) [24]. The cross-tabulation of the difficulty and discrimination indices presented in Table 1 shows the researchers' basis in retaining and rejecting items in the RCT.

The distractor efficiency was interpreted using the criteria utilized by Mahjabeen *et al.* [25], to wit, if less than 5% of the examinees chose the incorrect answers, the distractors are identified as non-functional distractors (NFD). Moreover, the distractor efficiency was interpreted along the following descriptive values: i) 3 NFD: 0% Distractor efficiency; ii) 2 NFD: 33.3% Distractor efficiency; iii) 1 NFD: 66.6% Distractor efficiency; and iv) 0 NFD: 100% Distractor efficiency. Finally, Chi-Square was employed in testing the relationship between RCT item indices and Cronbach's alpha was utilized to ascertain the reliability coefficient of the RCT. A test is considered reliable if the computed value is greater than .70 [21].

Table 1. Interpretation of the cross-tabulation of the difficulty and discrimination indices

Df/Ds	Poor	Reasonably good	Very good
Very difficult	Reject	Reject	Reject
Difficult	Reject	Retain	Retain
Desirable	Reject	Retain	Retain
Easy	Reject	Retain	Retain
Very easy	Reject	Reject	Reject

3. RESULTS AND DISCUSSION

3.1. Item analysis of the first pilot test

Table 2 indicates that the 120-item RCT exhibited a desirable difficulty, with an average difficulty index of 54.05 (range=11.54 to 94.23). Among these items, scores ranged from 30 to 92, showcasing that while most questions were suitably challenging, a few could benefit from revisions. Notably, the majority of items met the criteria for being reasonably good, as evidenced by an average discrimination index of 0.25 (range=-0.25 to 0.63). This suggests that most questions effectively differentiated between high and low performers. Additionally, the computed distractor efficiency of 79.42 (range=0-100%) showed that almost all items included effective or functional distractors.

Table 2. Features of the RCT items according to the evaluation criteria

Parameters	Result
Count of exam takers	52
Count of RCT items	120
Total points	120
Points obtained	
Mean±SD	64.87±14.24
Range	30 – 92
Index of difficulty	
Mean±SD	54.05±18.41
Range	11.54 - 94.23
Index of discrimination	
Mean±SD	0.25±0.18
Range	-0.25 - 0.63
Efficiency of distractors	
Mean±SD	79.42±25.28
Range	0-100

It is shown in Table 3 that the 120-item RCT has a large number of functional distractors. Considerably, 64 items possessed 100% distractor efficiency, 40 items had 66.60 distractor efficiency, 14 exhibited 33.30% distractor efficiency, and interestingly, only two items had 0% distractor efficiency. This finding implies that RCT included effectively feasible answers that examinees may select when making misjudgments or having partial knowledge or understanding of the competency, reducing the possibility of passing the test by merely guessing the test options [26].

Table 3. Types of the distractors by distractor efficiency

Type of NFD	Frequency	Percent	Distractor efficiency
0 NFD	64	53.33	100%
1 NFD	40	33.33	66.60%
2 NFD	14	11.67	33.30%
3 NFD	2	1.67	0%
Total	120	100	

NFD=Non-functional distractors

Table 4 presents that in terms of item difficulty, more than a majority, or 64 out of 120 items, were desirable. Significantly, among the items, 30 were considered challenging, 18 were categorized as easy, seven as very easy, and one as very difficult. This observation reflects that test-takers generally considered the RCT as quite challenging, as shown in the computed difficulty index of 54.05% Table 2. This finding could be explained by the nature of the RCT being a criterion-referenced test that measures competencies in conceptualizing, conducting, and writing research within teacher education context, which is reasonably challenging.

Additionally, through the utilization of point biserial coefficients, 51 items demonstrated a reasonably good level, while 50 items were very good. In contrast, 19 items were poor and consequently eliminated from consideration. With this finding, the researchers revisited each item's difficulty and discrimination indices to

improve the test items further. This procedure ushered in the dropping of additional eight items. After thoroughly reviewing the test questions, six items were revised to ensure that all the 54 competencies indicated in the TOS were represented. Subsequently, the RCT proceeded into the second phase of pilot testing with 100 items. This adjustment allowed for a shorter examination duration while maintaining the essential competencies assessed. Shortening the test could potentially alleviate test anxiety and the sense of burden for test-takers while still providing viable choices in assessment design [27].

Table 4. Contingency table of the difficulty and discrimination indices

Df/Ds	Poor	Reasonably good	Very good	Total
Very difficult	0#	0#	1#	1
Difficult	10#	17*	3*	30
Desirable	7#	22*	35*	64
Easy	1#	9*	8*	18
Very easy	1#	3#	3#	7
Total	19	51	50	120

*: Retained; #: Rejected

3.2. Item analysis of the second pilot test

Table 5 illustrates that 34 participants engaged in the second pilot test, consisting of 100 items with a score range of 24 to 77. The results indicated that the difficulty level of the RCT items was desirable as reflected in the mean difficulty of 50.76 (range=2.94-97.06). Notably, this finding aligns with Sahoo and Singh [28] assertion that an appropriate difficulty index for an item falls within the range of 30% to 70%. In a criterion-referenced test like the RCT, a desirable difficulty index is deemed acceptable as it emphasizes the mastery of competencies [29]. Additionally, criterion-referenced test items typically encompass a spectrum of difficulty levels since educational outcomes encompass a wide range of cognitive processes, ranging from fundamental to advanced [30].

Table 5. Features of the RCT items according to the evaluation criteria

Parameters	Result
Count of exam takers	34
Count of RCT items	100
Total points	100
Points obtained	
Mean±SD	49.35±14.43
Range	24 – 77
Index of difficulty	
Mean±SD	50.76±18.46
Range	2.94 - 97.06
Index of discrimination	
Mean±SD	0.28±0.19
Range	-0.18 - 0.71
Efficiency of distractors	
Mean±SD	78.97±24.02
Range	0-100

As shown in Table 6, among the 100 items in the RCT, 50 items displayed 100% distractor efficiency. Thirty-eight items contained one non-functional distractor, achieving a 66.6% distractor efficiency rate, while ten items had two non-functional distractors, totaling a 33.3% efficiency. Only two items featured three non-effective distractors, reaching 0% distractor efficiency as they did not attract any examinees. This finding suggests that, on the whole, the RCT incorporates a substantial number of effective distractors. Consequently, the test maintains an acceptable level of distractor efficiency. Prior research has affirmed that as the count of plausible distractors increases, a test tends to enhance its accuracy, validity, and consistency [31].

Table 6. Types of the distractors by distractor efficiency

Type of NFD	Frequency	Percent	Distractor efficiency
0 NFD	50	41.67	100%
1 NFD	38	31.67	66.60%
2 NFD	10	8.33	33.30%
3 NFD	2	1.67	0%
Total	100	100	

NFD=Non-functional distractors

Table 7 presents that a majority, specifically 59 out of 100 items, were deemed desirable. Notably, examinees perceived 27 items as difficult, while ten were categorized as easy, and four as very easy; none were deemed very difficult. This observation indicates that overall, test-takers found the RCT to be generally challenging based on the computed difficulty index of 50.76 earlier presented in Table 5. Furthermore, the assessment using point-biserial coefficients indicated that 52 items demonstrated a very good discriminatory ability, while 32 exhibited a reasonably good capacity to differentiate among examinees. This finding implies that the RCT effectively discerns examinees based on their research capabilities as gauged by the test items. Examinees achieving high scores tended to respond correctly to most items, whereas those with lower scores tended to answer most items incorrectly.

Significantly, the researchers assessed the quality of the items by cross-tabulating the difficulty index and discrimination index. A total of 82 items were considered fit for retention, while 18 items were rejected. The researchers further revisited all the items relative to the competencies and results of distractor efficiency. Two more items were removed from the test. Hence, the RCT had a total of 80 items based on the 54 competencies of the research capability of the educators.

Table 7. Contingency table of the difficulty and discrimination indices

Df/Ds	Poor	Reasonably good	Very good	Total
Very difficult	0#	0#	0#	0
Difficult	4#	13*	10*	27
Desirable	9#	16*	34*	59
Easy	1#	3*	6*	10
Very easy	2#	0#	2#	4
Total	16	32	52	100

*Retained; #Rejected

3.3. Relationship between the RCT quality indices

As shown in Table 8, the association between the RCT quality indices based on the result of the second item analysis. The finding reveals that the difficulty index is significantly related to distractor efficiency (x2=67.353; p =0.000). This result implies that RCT items with more effective distractors tend to possess higher difficulty indices. Put differently, when distractors are more efficient, it becomes more challenging for less-prepared examinees to identify the correct answer. Earlier studies have confirmed that test items featuring options that strongly attract test-takers are often associated with increased difficulty [32]. Hence, examinees who are able to answer more difficult items are more likely to overcome more functional distractors.

Table 8. Association of difficulty index, discrimination index, distractor efficiency

Indices	Difficulty index	Chi-square value	P-value
Difficulty index and discrimination index	8	11.464	0.177
Difficulty index and distractor efficiency	12	67.353	0.000
Discrimination index and distractor efficiency	6	5.711	0.456

3.4. Validity and reliability of the RCT

This study established the content and face validity of the RCT. There is a high level of content validity as indicated by the overall scale (S-CVI) of 0.956 suggesting that the items in the scale are highly relevant to the identified competencies. The closer the S-CVI is to 1.0, the better the content validity [33]. Moreover, the face validity of the RCT was ensured by seeking feedback from the panel of validators. These validators provided comments and suggestions, helping to refine the test to ensure that it measures what it is supposed to measure. By incorporating the feedback from the panel, the researchers enhanced the face validity, making the test more acceptable and credible to potential users. In summary, the study has demonstrated that the RCT has both strong content validity (as indicated by the high S-CVI) and adequate face validity (as ensured through the input of the panel of validators), making it a quality tool for measuring the identified competencies.

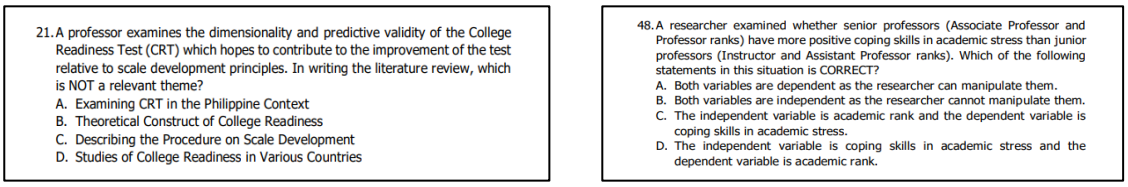
On the other hand, the reliability of the RCT questionnaire was assessed using Cronbach's alpha, a statistical measure commonly used to evaluate the internal consistency of a questionnaire or test. The computed value of Cronbach's alpha (α=0.904) indicates a high level of internal consistency within the RCT questionnaire [21], [34]. Typically, Cronbach's alpha values range between 0 and 1, with higher values indicating greater internal consistency. A value of 0.904 suggests that the items in the questionnaire are strongly correlated with each other, implying that the questionnaire is reliable in measuring the intended construct - in this case, the research capabilities of teacher education faculty. The statement also aligns with Neuman [35] perspective, which emphasizes that a reliable test should consistently yield similar results when administered

repeatedly to the same subjects. Therefore, the high Cronbach's alpha value (0.904) suggests that the RCT questionnaire is reliable and capable of delivering consistent outcomes when assessing the research capabilities of teacher education faculty.

3.5. Gender-sensitive and contextualized RCT items

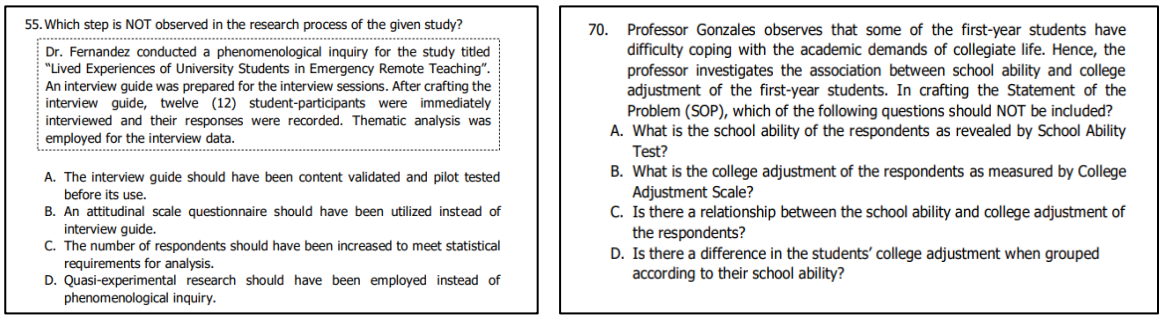
One of the features of the developed RCT is its gender-sensitivity. In Figure 2, the use of a "professor" as seen in Figure 2(a) and a "researcher" as seen in Figure 2(b) is gender neutral which suggests treating men and women evenly. The use of gender fair language lessons gender discrimination and stereotyping [36].

Figure 3 illustrates that the RCT items were situated within the contemporary landscape of the education field. For instance, item number 55 as shown in Figure 3(a) centered around Emergency Remote Teaching, a topic resonating with educators due to their experiences during the COVID-19 Pandemic. Additionally, item 70 as shown in Figure 3 (b) delved into the adjustment of first-year students in school and college, a relevant subject among educators at the tertiary level. Contextualizing the RCT in teacher education per se can aid the faculty in comprehending each test item more effectively, given that crafting questions aligned with real-world contexts has become a common practice in the academe [37].



(a) (b)

Figure 2. Sample items in the RCT which are gender neutral like the use of (a) a "professor" and (b) a "researcher"



(a) (b)

Figure 3. Sample items in the RCT which are contextualized such as topics on (a) emergency remote teaching and (b) school and college adjustment

4. CONCLUSION

The RCT is a valid test that measures the research capability of educators. It possesses desirable item difficulty, a reasonably good discrimination index, and high distractor efficiency. Notably, the RCT exhibits excellent internal reliability, and a significant association between its difficulty index and distractor efficiency, thus, affirming the test's soundness and credibility. On the other hand, the study has some limitations. First, the scope of RCT is limited to basic research competencies. Second, the established interpretation of the test is based only on test scores. Third, the scope of the pilot testing includes teachers only from public HEIs. Hence, it is suggested that the RCT may be subjected to further validation tests in other HEIs at the regional and national levels to ensure the generalizability of findings. Pilot testing the RCT to a larger scope may also establish an alternative way of interpreting results of the RCT through norm-reference. Moreover, the HEIs may use the RCT to determine the research capability of their faculty members as a baseline in formulating their college research roadmap. The graduate school may utilize the RCT as an admission or a complement to their current graduate school admission test in accepting students with sufficient research capability.

ACKNOWLEDGEMENTS

The researchers express their sincerest gratitude to Cagayan State University Research Development and Extension (CSU RDE) office for approving and funding this research (Special Order No. OP-5005-SO-2023-02-155). They are equally grateful to the expert validators who thoroughly critiqued the items in the RCT and the teacher educators who participated in the first and second pilot testing of the RCT.

REFERENCES

[1] R. Sarwar, A. Zia, R. Nawaz, A. Fayoumi, N. R. Aljohani, and S.-U. Hassan, "Webometrics: evolution of social media presence of universities," *Scientometrics*, vol. 126, no. 2, pp. 951–967, Feb. 2021, doi: 10.1007/s11192-020-03804-y.
[2] L. Dowsett, "Global university rankings and strategic planning: a case study of Australian institutional performance," *Journal of Higher Education Policy and Management*, vol. 42, no. 4, pp. 478–494, Jul. 2020, doi: 10.1080/1360080X.2019.1701853.
[3] University Rankings, "A closer look for research leaders," *Elsevier*, 2021. [Online]. Available: https://rb.gy/7lb85 (accessed May 16, 2023).
[4] C. Burmann, F. García, F. Guijarro, and J. Oliver, "Ranking the performance of Universities: The role of sustainability," *Sustainability*, vol. 13, no. 23, Nov. 2021, doi: 10.3390/su132313286.
[5] F. Selten, C. Neylon, C.-K. Huang, and P. Groth, "A longitudinal analysis of university rankings," *Quantitative Science Studies*, vol. 1, no. 3, pp. 1109–1135, Aug. 2020, doi: 10.1162/qss_a_00052.
[6] M. M. Vernon, E. A. Balas, and S. Momani, "Are university rankings useful to improve research? A systematic review," *PLOS ONE*, vol. 13, no. 3, Mar. 2018, doi: 10.1371/journal.pone.0193762.
[7] D. Li and Y. Li, "'Trust the people, not the institution': university preparation for the Research Assessment Exercise 2020 in Hong Kong," *Higher Education Research & Development*, vol. 42, no. 3, pp. 633–648, Apr. 2023, doi: 10.1080/07294360.2022.2096574.
[8] J. Li, "The global ranking regime and the reconfiguration of Higher education: Comparative case studies on research assessment exercises in China, Hong Kong, and Japan," *Higher Education Policy*, vol. 29, no. 4, pp. 473–493, Dec. 2016, doi: 10.1057/s41307-016-0015-7.
[9] M. Vallez, C. Lopezosa, and R. Pedraza-Jiménez, "A study of the Web visibility of the SDGs and the 2030 Agenda on university websites," *International Journal of Sustainability in Higher Education*, vol. 23, no. 8, pp. 41–59, Dec. 2022, doi: 10.1108/IJSHE-09-2021-0361.
[10] G. Chen and L. Chan, *University rankings and governance by metrics and algorithms*. Research Handbook on University Rankings, Edward Elgar Publishing Ltd, 2021.
[11] M. Caingcoy, "Research capability of teachers: Its correlates, determinants and implication for continuing professional development," *Journal of World Englishes and Educational Practices*, vol. 2, no. 5, pp. 1–11, Oct. 2020, doi: 10.32996/jweep.2020.2.5.1.
[12] "Section 8 of Republic Act (RA) No. 7722." [Online]. Available: https://rb.gy/hv0qb (accessed May 16, 2023).
[13] "Commission on higher education (CHED) memorandum order no. 52, series of 2016." [Online]. Available: https://rb.gy/utxt0 (accessed May 16, 2023).
[14] "Commission on higher education (CHED) memorandum order no. 16, series of 2015," *Amendment to CMO No. 26, s. 2007 – COE-COD for Teacher Education*. [Online]. Available: https://ched.gov.ph/2015-ched-memorandum-orders/ (accessed May 16, 2023).
[15] G. Ion and R. Iucu, "The impact of postgraduate studies on the teachers' practice," *European Journal of Teacher Education*, vol. 39, no. 5, pp. 602–615, Oct. 2016, doi: 10.1080/02619768.2016.1253674.
[16] M. Ulla, K. Barrera, and M. Acompanado, "Philippine classroom teachers as researchers: Teachers' perceptions, motivations, and challenges," *Australian Journal of Teacher Education*, vol. 42, no. 11, pp. 52–64, Nov. 2017, doi: 10.14221/ajte.2017v42n11.4.
[17] T. M. Willemse and F. Boei, "Teacher educators' research practices: an explorative study of teacher educators' perceptions on research," *Journal of Education for Teaching*, vol. 39, no. 4, pp. 354–369, Oct. 2013, doi: 10.1080/02607476.2013.797292.
[18] I. B. Gonzales, D. A. Corpuz, and R. M. Dellosa, "Research capabilities of public elementary school teachers and management support of the schools division of Nueva Vizcaya, Philippines," *Humanities & Social Sciences Reviews*, vol. 8, no. 4, pp. 258–266, Jul. 2020, doi: 10.18510/hssr.2020.8427.
[19] V. E. Tamban and O. B. Maningas, "Research capability of public school teachers: A basis for research capability enhancement program," *PEOPLE: International Journal of Social Sciences*, vol. 6, no. 1, pp. 222–235, 2020, doi: 10.20319/pijss.2020.61.222235.
[20] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate data* analysis, 7th edition. Pearson, 2009.
[21] L. Crocker and J. Algina, *Introduction to classical and modern test theory*. Cengage Learning, 2008.
[22] R. J. Gregory, *Psychological testing: History, principles, and applications*, 7th edition. Pearson, 2013.
[23] C. B. Essen and G. Akpan, "Analysis of difficulty and point- biserial correlation indices of 2014 Akwa Ibom State mock multiple choices mathematics test," *International Journal of Education and Evaluation*, vol. 4, no. 5, pp. 1–11, 2018.
[24] "EAC visual data," *Educational Assessment Corporation 2009*, 2009. [Online]. Available: https://www.edassess.net/eacs/pointbiserial.aspx (accessed Jul. 03, 2019).
[25] W. Mahjabeen *et al.*, "Difficulty index, discrimination index and distractor efficiency in multiple choice questions," *Annals of Pakistan Institute of Medical Sciences*, vol. 13, no. 4, pp. 310–315, 2017.
[26] N. Thompson, "Distractor analysis for test items," *ASC*, 2022. [Online]. Available: https://assess.com/distractor-analysis-test-items/
[27] B. P. Foley, "Getting lucky: How guessing threatens the validity of performance classifications," *Practical Assessment, Research, and Evaluation*, vol. 21, no. 3, 2016.
[28] D. P. Sahoo and R. Singh, "Item and distracter analysis of multiple choice questions (MCQs) from a preliminary examination of undergraduate medical students," *International Journal of Research in Medical Sciences*, vol. 5, no. 12, pp. 5351–5355, Nov. 2017, doi: 10.18203/2320-6012.ijrms20175453.
[29] M. H. Al-Habashneh, M. H. Al-Habashneh, and D. N, "Constructing a criterion reference test to measure the research and statistical competencies of graduate students at the Jordanian Governmental Universities," *Journal of Education and Practice*, vol. 8, no. 2, pp. 67–81, 2017.
[30] K. Burton, "Designing criterion-referenced assessment," *Journal of Learning Design*, vol. 1, no. 2, 2012, doi: 10.5204/jld.v1i2.19.
[31] T. Puthiaparampil and M. Rahman, "How important is distractor efficiency for grading best answer questions?" *BMC Medical Education*, vol. 21, no. 1, p. 29, Dec. 2021, doi: 10.1186/s12909-020-02463-0.
[32] M. R. Hingorjo and F. Jaleel, "Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency," *Journal of the Pakistan Medical Association*, vol. 62, no. 2, pp. 142–147, 2012.

[33] M. S. B. Yusoff, "ABC of content validation and content validity index calculation," *Education in Medicine Journal*, vol. 11, no. 2, pp. 49–54, Jun. 2019, doi: 10.21315/eimj2019.11.2.6.
[34] K. S. Taber, "The use of cronbach's alpha when developing and reporting research instruments in science education," *Research in Science Education*, vol. 48, no. 6, pp. 1273–1296, Dec. 2018, doi: 10.1007/s11165-016-9602-2.
[35] W. L. Neuman, *Social research methods: qualitative and quantitative approaches*. UK: Pearson, Essex, 2014.
[36] S. Sczesny, M. Formanowicz, and F. Moser, "Can gender-fair language reduce gender stereotyping and discrimination?" *Frontiers in Psychology*, vol. 7, Feb. 2016, doi: 10.3389/fpsyg.2016.00025.
[37] A. Ahmed and A. Pollitt, "Improving the quality of contextualized questions: an experimental investigation of focus," *Assessment in Education: Principles, Policy & Practice*, vol. 14, no. 2, pp. 201–232, Jul. 2007, doi: 10.1080/09695940701478909.

BIOGRAPHIES OF AUTHORS

Jay-cen Tamayo Amanonce is an Assistant Professor at the College of Teacher Education and Graduate School, Cagayan State University, Tuguegarao City, Cagayan, Philippines. He is a graduate of Doctor of Philosophy (PhD) in Mathematics Education. His previous administrative designations are Campus Quality Assurance Coordinator, Campus ISO Focal Person, and College Document Control Officer. His research interests are research capability, teacher education, mathematics education, online teaching and learning, college readiness, and licensure examination for teachers. He published several research articles in Scopus-indexed and ASEAN Citation Index journals. Also, he presented his scholarly works in national and international scientific conferences and fora. Notably, he was a recipient of Outstanding Faculty Member of Cagayan State University in 2021 and 2022. He can be contacted through his email address jaycen.amanonce@gmail.com.

Rudolf Tamangen Vecaldo is an Accredited Full-fledged Professor at the College of Teacher Education, Cagayan State University, Tuguegarao City, Cagayan, Philippines. He is a graduate of Doctor of Education (EdD) and a regular member of the Division VIII (Social Sciences) of the National Research Council of the Philippines. He is a research professor in the College of Teacher Education and the Graduate School. Impressively, he was awarded the Outstanding Researcher in 2021. His research interests are college readiness, teacher education, beliefs system, culture, phenomenological inquiry, and research capability. He published numerous research articles in Scopus-indexed, Web of Science, and ASEAN Citation Index journals. He presented his scholarly works in national and international scientific conferences and fora. He can be contacted through his email address rudvectop2@gmail.com.

Maribel Fontiveros Malana is an Accredited Full-fledged Professor at the College of Teacher Education, Cagayan State University, Tuguegarao City, Cagayan, Philippines. She is a graduate of Doctor of Philosophy in Education (PhD) major in Educational Management. She also finished a Diploma Course in Language Studies Education as a scholar of the University of the Philippines-Department of Education Open University. She has been teaching language and professional education courses. She is the research coordinator of the College of Teacher Education. Her research interests are language education, student leadership, professional education, and research capability. She had published articles in Scopus-indexed and presented paper in local and international conferences and fora. She can be contacted through her email address mfontiveros.malana@gmail.com.

Conchita Malenab Temporal is an Accredited Full-fledged Professor at the College of Teacher Education, Cagayan State University, Tuguegarao City, Cagayan, Philippines. She is a graduate of Doctor of Philosophy in Education (PhD) major in Language Education and handles language and literature courses both in the undergraduate and graduate programs. She has been in the teaching profession for two decades now and counting and has been a recipient of the Philippine Education Leadership Award 2019. Her research interests are language teaching and learning, applied linguistics, cross-cultural communication, translation studies, and research capability. She has carried out research reviews and has presented researches in national and international fora and published researches in peer-reviewed journals. She can be contacted through conchitatemporal@csu.edu.ph.

Antonio Ingaran Tamayao is an Accredited Full-fledged Professor at the College of Teacher Education, Cagayan State University, Tuguegarao City, Cagayan, Philippines. He is a graduate of Doctor of Philosophy in Education (PhD) major in Development Education. He is the project leader of a funded research by the Commission on Higher Education (CHED). He is a prolific book writer of professional education and social sciences courses. He is a passionate professor, mentor, and researcher. His research interests are along college readiness, Ibanag and Ilocano confluence, migration, and other related sociological issues in education. His research papers are published in different international refereed journals. Few of these papers were presented in Shanghai University, University of Hawaii, University of Queensland, University Kebangsaan Malaysia, and prestigious universities in the Philippines. He can be contacted through toni_tamayao@yahoo.com.

Jhoanna Battung Calubaquib is an Accredited Full-fledged Professor at the College of Teacher Education, Cagayan State University, Tuguegarao City, Cagayan, Philippines. She is a graduate of Doctor of Philosophy in Education (PhD) major in Science Education. She is an associate member of the Division X (Chemical Sciences) of the National Research Council of the Philippines. She is the University Research Director of Cagayan State University. Her research interests focus on Plant Science, Biochemistry and Environmental Science, Health & Nutrition, and research management and capability. She published various research articles in Scopus-indexed and Web of Science journals. Also, she presented her scholarly works in national and international scientific conferences and fora. Remarkably, she was a recipient of Outstanding Technology Transfer in 2021 and Research Innovation Award in 2022. She can be contacted through jb_120771@csu.edu.ph.

Marie Claudette Maramag Calanoga is an Accredited Full-fledged Professor at the College of Teacher Education, Cagayan State University, Tuguegarao City, Cagayan, Philippines. She is a graduate of Doctor of Philosophy in Education (PhD) major in Development Education. She is the University Dean of the College of Teacher Education and the University Director of Quality Assurance. She has been teaching language and research courses in the undergraduate and graduate programs. Her research interests center on language, teacher education, licensure examination, blended learning, and research capability. She published research articles in reputable international refereed journals and presented papers in national and international conferences and fora. She can be contacted through her email address acemarie365@csu.edu.ph.

Ria Arellano Tamayo is an Accredited Full-fledged Professor at the College of Teacher Education, Cagayan State University, Tuguegarao City, Cagayan, Philippines. She is a graduate of Doctor of Philosophy in Education (PhD) major in Language Education. She teaches language and professional education courses in the undergraduate programs. She has been the detailed secretary at the office of the Vice President for Academic Affairs. She is a regular member of the Association of Language and Literature Teachers of the Philippines. Her research interests include linguistics, structural analysis, discourse analysis, language policy making, research capability, language education & literacy, and teacher education. She presented papers in international conferences and published several research articles in international peer-reviewed journals. She can be contacted through her email address ria100581@gmail.com.