Research Article

Development and Validation of an Aptitude Test in Research Productivity

Joy B. Araza [1]

[1] *Associate Professor, Samar State University Catbalogan City, Samar Philippines. Email: joy_arazajoy60@yahoo.com*

ARTICLE INFO	ABSTRACT
	The creation and validation of the Aptitude Test for Research Productivity (ATRP), a tool for predicting higher education faculty members' research capacity and productivity, are presented in this paper. ATRP is positioned as a tool to help with the strategic recruitment and development of faculty members who are engaged in research. Research used a mixed-methods strategy, thorough literature analysis conducted in the first phase to determine the essential competencies linked to high research productivity, refined into quantifiable attributes through expert interviews, modified into a pilot test version that included several psychometric features meant to assess cognitive and non-cognitive abilities relevant to research. Sample faculty members from three state universities participated in a multi-step process that involved item analysis, reliability testing, and construct validation to validate the ATRP. The study evaluated the validity and reliability of the ATRP using Rasch measurement models and traditional test theory. Ordinal logistic regression used to assess the test's predictive ability and determine how well it could differentiate between various degrees of research productivity. The validation process's results showed that the ATRP can accurately distinguish between faculty members who are likely to be highly and poorly productive researchers. ATRP items successfully capture fundamental components of research aptitude, evidenced by the test's strong construct validity and internal consistency. The study suggests using ATRP in conjunction with a comprehensive evaluation plan for hiring and training new faculty members. Subsequent investigations on enhancing the ATRP by the integration of adaptive testing technologies and investigating the effects of cross-disciplinary applications. **Keywords:** Research Productivity, Aptitude Testing, Higher Education, Psychometric Validation, Faculty Development.

INTRODUCTION

The development and validation of an aptitude test in research productivity is underscored by the intricate dynamics of the multifaceted attributes that contribute to Research productivity, which has seen increased emphasis over the past two decades across state universities, colleges, higher education institutions, and the global research community. Understanding research productivity entails looking into various factors that influence it.

In essence, the rationale for developing an aptitude test for assessing research productivity is essential in today's landscape of education both in basic and higher education. Perspectives on aptitude as an inherent or learned competence suggest that these can measured to surface a range of capabilities crucial for research success. Validated aptitude tests could aid in matching individuals with suitable opportunities, guiding investments in skill development, and optimizing institutional support systems to foster a thriving research environment conducive to innovation and knowledge advancement.

Aptitude tests, known for predicting future performances under new sets of conditions (Maggay, 2017; cited by Macklem, 1990), can offer insights into a researcher's potential to navigate the complex landscape of modern research.

Aptitude test could measure a range of capabilities, from the ability to acquire new knowledge to the potential for contributing meaningfully to one's field of study (Ballado, Morales, and Ortiz, 2014; cited by Aiken, 1985; Kubiszyn and Borich, 2003).

Research productivity is complex and rooted on several factors (Brew, Boud, Namgung, Lucas & Crawford, 2015; cited by Lee & Bozeman, 2005).

Statement of the Problem

This study aimed to explore the comprehensive process of test development and validation to create effective, results-driven tests that accurately assessed an examinee's aptitudes in research productivity.

This study aimed to explore the comprehensive process of test development and validation to create effective, results-driven tests that accurately assessed an examinee's aptitudes in research productivity. Specifically it answers the following questions:

(1) What dimensions emerge as indicators of research productivity?

(2) What is the Content Validity Index (CVI) of the test items developed for measuring research productivity?

(3) What is the Reliability Index of the scale developed?

(4) What is the dimensionality of the test?

(5) How well does the Aptitude Test for Research Productivity (ATRP) conform to the Rasch Measurement Model, indicating the test's fit and its ability to measure the construct with precision and accuracy?

(6) How well does the Aptitude Test for Research Productivity (ATRP) estimates research productivity.

The study leverages four theoretical perspectives—Psychometric Theory, Validity Theory, Factor Analysis Theory, and Motivation Theory—to enrich understanding and development of an aptitude test for research productivity. These theories contribute to constructing a test that not only predicts aptitudes effectively but also correlates with actual performances, thereby offering a more accurate, effective, and comprehensive approach to assessing research productivity.

The conceptual framework for developing an aptitude test in research productivity integrates theories along with Researcher Personal Profile, Professional Background, and Personality Character. These elements together with the literature review helped identify behaviors indicative of aptitude in research productivity, aiding the development of the test. Additionally, the Test Validity Theory and Theory of Factor Analysis underpin the instrument's development and validation, ensuring a rigorous approach to discerning indicators of varying levels of research productivity.
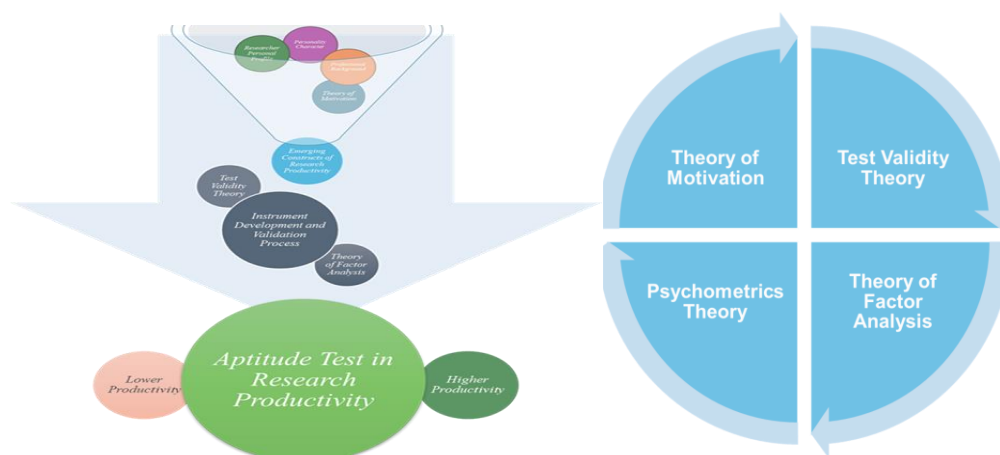

Figure 1. Conceptual Framework

LITERATURE REVIEW

While acknowledging the diverse factors impacting research productivity, this study focuses solely on cognitive aspects in assessing aptitude, highlighting the complexity of capturing all contributors within a single test. While crucial, cognitive abilities may not fully encompass motivational drivers, experiential backgrounds, and personality traits. This study lays the foundation for future research to explore a broader range of variables, aiming for a more holistic assessment of aptitude in research productivity.

Concept	References	Synthesis
Theory of Motivation	- Hardré, 2011, Mawoki & Babandako, 2011, Greenberg & Baron, 2010, Beck, 2004	These studies explore how motivation, both intrinsic and extrinsic, influences research productivity. They collectively highlight the pivotal role of motivational factors in driving academic research efforts and outcomes.
Researcher Personal Profile	- Nguyen, Quy Huu, 2015, Meyer and Allen, 1997	This theme underscores the impact of personal attributes on research productivity. Factors such as gender, tenure, and academic rank are found to influence research output, suggesting a need for personalized approaches in the development of aptitude tests.
Professional Background	- Tien & Blackburn, 1996, Chen et al, 2006, Chen et al, 2010	The literature suggests that the professional background of researchers, including their field of study and professional experiences, significantly affects their research productivity. This informs the development of aptitude tests by highlighting the importance of contextual and disciplinary considerations.
Personality Character	- Hunter & Kuh, 1987, Tien, 2000	Studies indicate that personality traits and characteristics, such as curiosity and adaptability, are crucial predictors of research productivity. This emphasizes the importance of including personality assessments in the aptitude test development to capture a comprehensive view of potential research productivity.
Test Validity Theory	- Wolming, Simon & Wikström, 2010, American Psychological Association (various years)	The evolution of validity theory from a simple, straightforward concept to a complex, multifaceted approach underscores the importance of developing aptitude tests that accurately measure what they intend to. This evolution supports the need for a rigorous validation process in the development of the aptitude test for research productivity.
Theory of Factor Analysis	- Cronbach, 1949 - Thurstone, 1947 - Cattell (Child, 1998)	Factor analysis theory helps in identifying specific abilities or factors that contribute to research productivity. This theory supports the use of statistical techniques to design aptitude tests that can discern distinct capacities relevant to research productivity.

The literature review reveals a comprehensive understanding of the various elements that contribute to research productivity, including motivation, personal profiles, professional backgrounds, and personality characteristics. However, this study acknowledges a notable limitation: it will primarily focus on the cognitive aspects of assessing aptitude, positioning it as the sole variable to be tested in the model. This decision underlines the complexity of encapsulating the full spectrum of factors influencing research productivity within a single aptitude test. While

cognitive abilities are crucial, this approach may not fully account for other significant contributors, such as motivational drivers, experiential backgrounds, and personality traits, that also play pivotal roles in determining research effectiveness. Recognizing this constraint, the current study serves as an initial step toward validating the aptitude test's effectiveness. It sets the groundwork for future research to explore predictive testing further and incorporate a broader range of variables, offering a more holistic assessment of aptitude in research productivity.

METHODOLOGY

Research productivity defined as the number of completed and published research projects by an individual over a certain period, with emphasis on peer-reviewed and Scopus publications in three Levels:
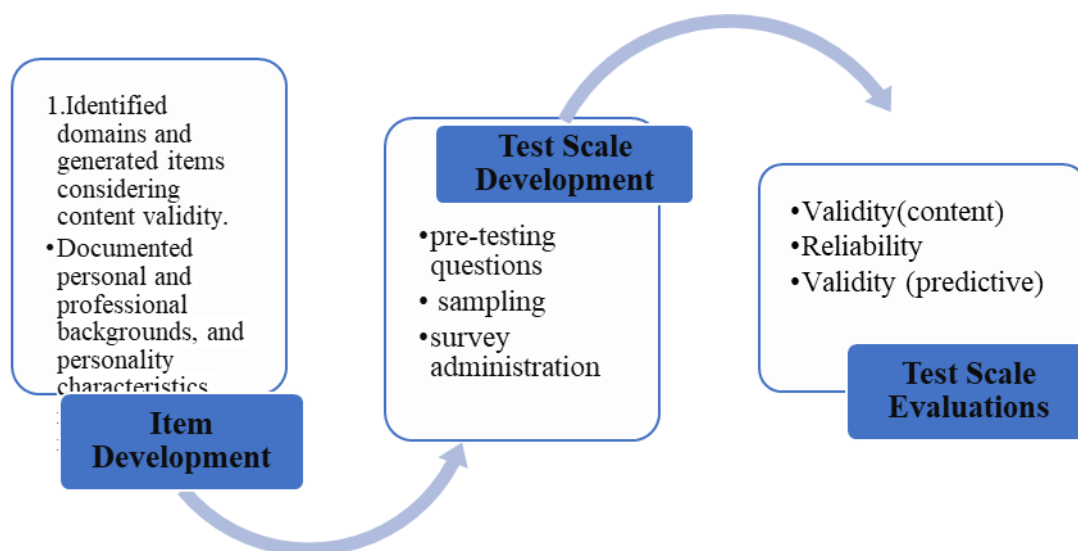

Figure 2. Test Development Procedure

Level 3- highly productive, with Scopus Publications

Level 2 - moderate productive with International referred Publications

Level 1 - low productive, with no publication to local publications

The methodology for developing an Aptitude Test in Research Productivity involves a structured multi-phase process focusing on the measurement of variables, participant selection, research instrument development, design and procedures, and data processing. First we define research productivity as the number of completed and published research projects by an individual over a certain period, with emphasis on peer-reviewed and Scopus publications,

The design of the study is **Exploratory Sequential Mixed Method Research Design** employed for test development, including item development, test scale development, and evaluation phases, alongside Instrumentation research for test creation.

Specifically: 1, **Item Development** which includes: Identified domains and generated items considering content validity and Documented personal and professional backgrounds, and personality characteristics relevant to research.

Second, Test Scale Development which includes: Involved pre-testing questions, sampling, survey administration, item reduction, and extraction of latent factors, and Third, **Test Scale Evaluations** specifically on tests for validity and reliability.

KII – 12 SUC Faculty enriched with 89 Document Review (Books)

Content Validation – Three Experts

Pre-Testing – 100 SUC faculty

Reliability and Validity Testing – 837 public school and SUC Faculty

Table 1: Respondents of the study

University/School Division	No. of Faculty	No. of Target Respondents	No. of actual Respondents
SUC 1	238	238	237
SUC 2	218	218	200
SUC 3	276	276	200
High School 1	436	200	200
Total	**1163**	**932**	**837**

The Participants of the Study involved *12 SUC Faculty for the KII enriched with 89 Document Review (Books), 3 experts for Content Validation, 100 SUC Faculty in the Pre-Testing and for* Reliability and Validity Testing a total of 837 public school and SUC Faculty shown in Table 1 on the screen.

Research Instruments

Developed a multiple-choice Aptitude Test based on literature review, in-depth interviews, and analysis of existing tests.

Data Analyses Procedure

- Utilized thematic data analysis for qualitative data and Content Validity Index (CVI) for test item quality.

- Conducted item analysis for test refinement and employed Rasch Model Measurement (RMM) for quantitative analysis, focusing on validity and reliability aspects.

- Ordinal and Binary Logistic Regression were used to test the predictive capacity of the test regarding research productivity levels.

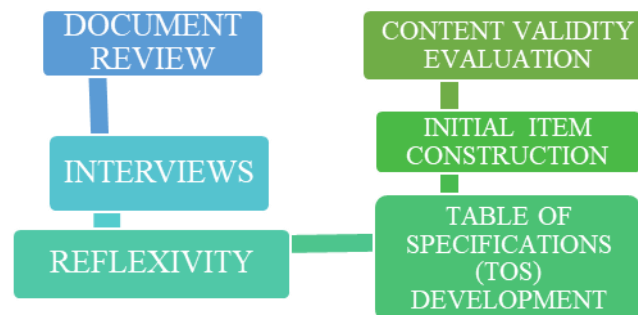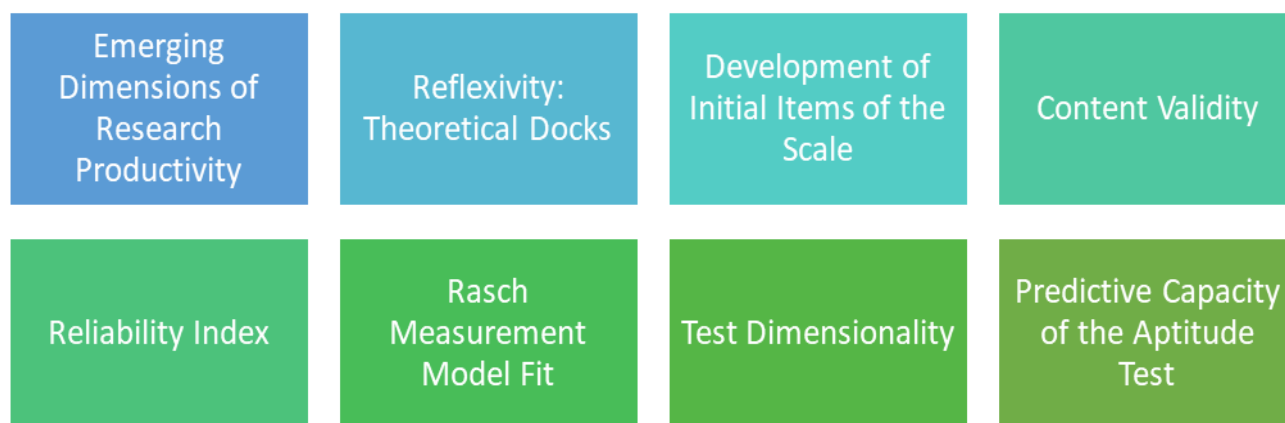**Data Gathering Procedure: Scale Development and Validation**

Figure 3. Data Gathering Procedure

In the development of our research instruments, a meticulous process was undertaken to construct a multiple-choice Aptitude Test aimed at gauging research productivity. This endeavor was rooted in a comprehensive literature review, insightful in-depth interviews, and a critical analysis of existing tests. Our data gathering procedure was systematically designed to ensure the robustness of our scale development and validation process. It encompassed a variety of methodologies, including document reviews, interviews to ensure reflexivity, the development of a Table of Specifications (TOS) to guide the construction of our test items, and rigorous evaluations of content validity.

When it came to data analysis, our approach was two-pronged. Qualitatively, we leaned on thematic data analysis to distill insights from our data, ensuring that the nuances of the responses were captured and interpreted accurately. Quantitatively, we adopted the Content Validity Index (CVI) to ascertain the quality of our test items, ensuring they met high standards of relevance and accuracy. Further refining our test, we conducted item analyses to improve its precision and reliability. The Rasch Model Measurement (RMM) was pivotal in our quantitative analysis, allowing us to delve deeply into the validity and reliability aspects of our instrument. Lastly, to assess the predictive capacity of our Aptitude Test concerning research productivity levels, we employed Ordinal and Binary Logistic Regression analyses. This comprehensive methodology underscores our commitment to developing a rigorously validated tool that can accurately predict research productivity among academics

RESULTS AND DISCUSSION

Emerging Dimensions of Research Productivity	Reflexivity: Theoretical Docks	Development of Initial Items of the Scale	Content Validity
Reliability Index	Rasch Measurement Model Fit	Test Dimensionality	Predictive Capacity of the Aptitude Test

The results of the study are presented as Emerging Dimensions of Research Productivity; Reflexivity: Theoretical Docks; Development of Initial Items of the Scale; Content Validity; Reliability Index; Rasch Measurement Model Fit; Test Dimensionality; and Predictive Capacity of the Aptitude Test

In our exploration of the emerging dimensions of research productivity, we've identified eight key areas that significantly influence an individual's ability to contribute effectively to their field. These dimensions range from the initial attitudes toward research to the concrete outputs that define productivity.

1. **Attitude at Starting the Research Productivity Journey**: Highlights the critical role of intrinsic curiosity and a proactive approach towards research from an early stage. This foundation is pivotal for engaging deeply in research endeavors.

2. **Motivation in Engaging and Being Productive in Research**: Unveils the varied motivations behind research productivity, including financial incentives, academic recognition, and a deep-seated passion for discovery and contribution to knowledge.

3. **Personal and Professional Life Advantages**: Emphasizes the dual benefits of research productivity on both personal growth and professional advancement, such as income improvement, career progression, and enhanced credibility in one's field.

4. **Challenges or Hurdles in Research Productivity**: Acknowledges the barriers to research productivity, including time management issues, bureaucratic obstacles, and the balancing of workload, underscoring the need for supportive frameworks.

5. **How to Be Productive in Research**: Outlines strategies for enhancing research productivity through continuous learning, collaboration, and a disciplined research approach, pointing towards the importance of staying current and engaged in one's field.

6. **Preparation Needed to Engage in Research Productivity**: Stresses the significance of thorough preparation through academic training, conference participation, and practical engagement in research-related activities for effective research productivity.

7. **Skills Contributing to Research Productivity**: Identifies essential skills for research productivity, such as adaptability, diligence, creativity, and perseverance, highlighting the blend of soft and hard skills required for successful research endeavors.

8. **Definition of Research Productivity**: Defines research productivity in terms of impactful outputs and the ability to produce work that significantly advances one's field, emphasizing the value of high-quality publications and practical applications of research findings.

Reflexivity: Theoretical Docks

Our exploration into research productivity reveals its strong ties to key theoretical frameworks: Personal Profile, Theory of Motivation, Personality Character, and Professional Background. Key takeaways include:

- **Personal Agency**: Initial positive attitudes toward research, driven by curiosity and exploration (aligned with Personal Profile theory), are fundamental for setting the stage for research productivity.

- **Motivation**: The drive for research productivity is fueled by both intrinsic desires and extrinsic rewards, showing the significant role of motivation in engaging with research.

- **Character of Productivity**: Effective research productivity stems from proactive engagement and preparation, including literature review and strategic planning, which are essential for overcoming research challenges.

- **Professional Background**: Navigating the path of research involves facing and overcoming various obstacles, emphasizing the importance of a supportive environment and individual resilience.

In our investigation into the landscape of research productivity, we've distilled our findings into several key dimensions that align closely with established theoretical frameworks, offering a comprehensive view of what drives and influences the capacity for research excellence.

Our study on research productivity reveals key dimensions shaped by personal drive, motivation, personality, and professional experiences. Starting with a researcher's innate curiosity and motivation, these personal attributes lay the foundation for embarking on a productive research journey. Motivation, both from internal passion and external rewards, is crucial for engagement and productivity. Personality traits like resilience fuel perseverance through professional challenges, while the background sets the stage for overcoming obstacles and preparing for success. In summary, the interplay of these factors illustrates the complex nature of achieving research excellence, highlighting the importance of a supportive and adaptive approach in the pursuit of academic and professional growth.

### Development of Initial Items of the Scale

In our study, we developed a comprehensive scale to measure teachers' research productivity, drawing from literature and in-depth interviews with experienced researchers. This process led to 277 initial test items, emphasizing the teachers' experiences and perceptions of research productivity. The scale, contextualized to reflect highly productive researchers' experiences, spans various competencies crucial for research aptitude, such as the research process, data collection methods, and analysis, accounting for the complexities of academic research. The Table of Specifications outlines these competencies, ensuring the scale comprehensively covers the essential aspects of research productivity, from methodology to ethics, highlighting the multifaceted nature of conducting impactful research.

Table 2: Table of Specification for Aptitude Test for Research Productivity

Competencies/ Topics	Number of Items (Item Location)	Number of items	Percentage (%)
Type of Research	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,16, 17, 18, 19, 20, 21, 22	22	7.94
Research Process	23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71	49	17.69
Data Collection method	72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106	35	12.64
Level of measurement / Rating scale	107, 108, 109, 110, 111, 112, 113, 114, 115, 116	10	3.61
Reliability and Validity	117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130	14	5.05
Research Data Analysis	131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158	28	10.48
Literature Review	159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188	30	10.10
Research Proposal	189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208	20	7.22
Research Terminal Report	209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222	14	5.05
Citation and referencing	223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235	13	4.69
Researcher Qualities and Character	236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251	16	5.77
Research Ethics	252, 253, 254, 255, 256, 257, 258, 259, 260, 261	10	3.61
Research Productivity	262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277	16	5.77
Total		277	100

Content Validity

Through the initial validation process involving expert review and content validity index analysis, 62 items were identified for revision or elimination from the original 277-item aptitude test, leading to a refined set of 215 items that better align with established content validity standards.
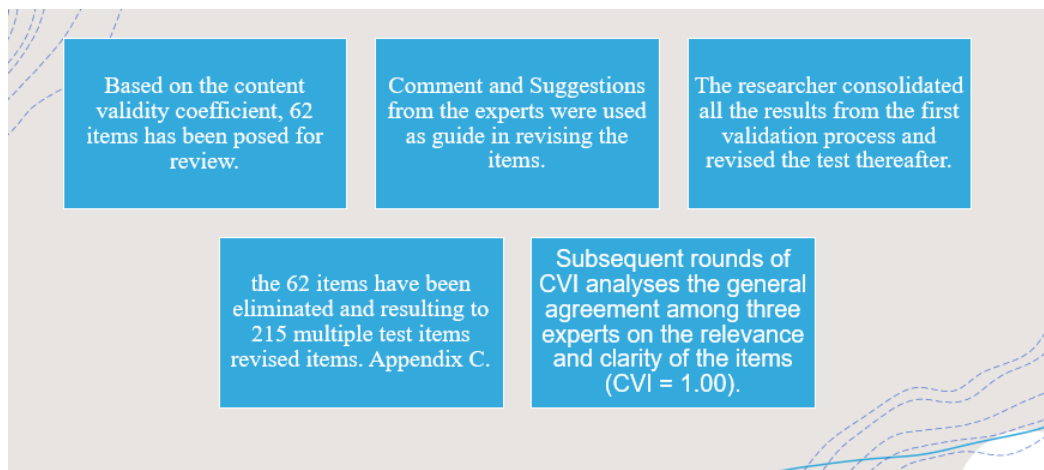
Reliability Index
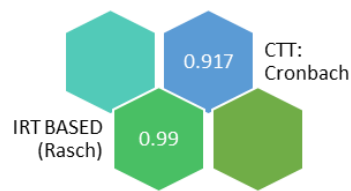
The reliability of the revised aptitude test for measuring teacher research productivity, initially pretested with 100 respondents, was confirmed through item analysis and Cronbach's alpha coefficient (.917), leading to a refined set of 75 items. Further testing with 850 respondents using the Rasch Measurement Model yielded an exceptionally high reliability index of 0.99, indicating the test's stability and consistency in evaluating the construct of research productivity.

Rasch Measurement Model Fit

Table 3: Rasch Unidimensionality coefficient, overall Fit statistics, and reliability coefficient

Measures	Min	Max	Mean	SD
Infit MNSQ	.81	1.22	0.99	.01
Outfit MNSQ	.54	1.50	1.05	.03
Rasch Unimensionality Coefficient			1.10	
Item Reliability			.99	

Model fit in this study shown in Table 3, determined by outfit and infit MNSQ values between 0.5 and 1.5, confirms the suitability and accuracy of person and item alignment with the Rasch model, indicating the test scale appropriately measures the intended construct.

The item-person map visually represents the match between test item difficulties and respondent abilities on a scale of -3 to +4 logits, with item difficulty and respondent ability levels directly compared to identify whether a respondent's ability is above, below, or at the average difficulty level of the test items.
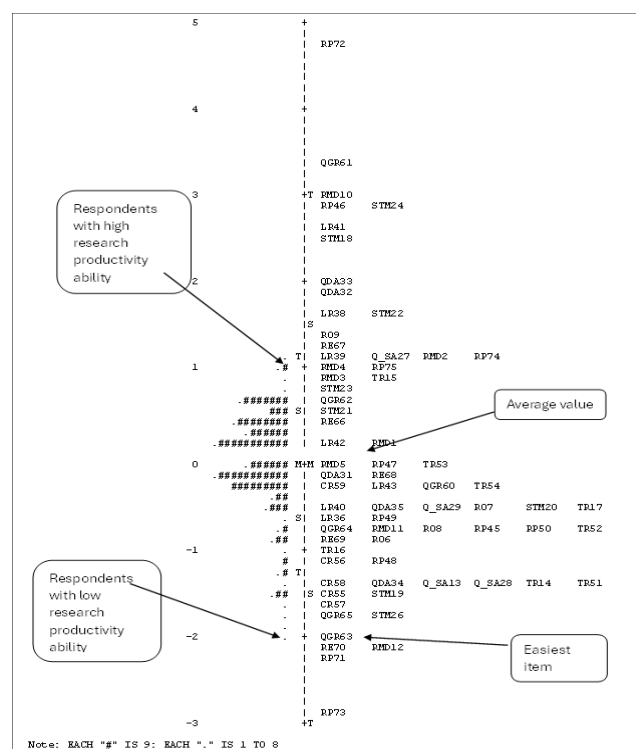

Figure 4. Person-Item Map

Test Dimensionality

Table 4: Rasch Unidimensionality coefficient, overall Fit statistics, and reliability coefficient

Measures	Min	Max	Mean	SD
Infit MNSQ	.81	1.22	0.99	.01
Outfit MNSQ	.54	1.50	1.05	.03
Rasch Unimensionality Coefficient			1.10	
Item Reliability			.99	

The study's Rasch unidimensionality coefficient of 1.10 indicates a clearly unidimensional variable, affirming the test scale's focus on measuring a single construct, research productivity, as per psychometric requirements.

Predictive Capacity of the Aptitude Test

Table 5: Model fitting information

Model Fitting Information				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	212.963			
Final	172.825	40.138	23	.015
Goodness of Fit				
Pearson		29.043	23	.179
Deviance		31.058	23	.121

The regression model's effectiveness is highlighted by the significant improvement in model fit when including the aptitude test scores as predictors of research productivity levels, demonstrating that scores are a useful, though not exhaustive, indicator of productivity. The model's pseudo R-squared values, while low, indicate that the aptitude test captures a portion of the variance in research productivity. Furthermore, the model's goodness-of-fit tests confirm an adequate fit to the observed data, validating the ordinal logistic regression approach for this analysis. A significant observation also highlights that that the threshold, pairwise, in a step process provides cut points between the categories of the dependent variable. The findings show that there exists a significant threshold between "Low Productivity" and other categories indicating a statistical separation. The lack of comparison between Level 2 and Level 3 in the table underscores that there is no-significant threshold between these levels, as such of less distinction. The analyses then leads to explore two categories of predictor cobining the Level 2 and Level 3 as they are not statistically distinguishable.

Table 6: Model Estimation

Pseudo R-Square	Cox and Snell	Nagelkerke	McFadden
	.047	.055	.025
Estimates			

Parameter Estimates		Level 1	Level 2
Estimate		-2.309	0.432
Standard Error	Threshold	0.619	0.612
Wald		13.925	0.498
df		1	1

Test of Parallel Lines				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	172.825			
General	141.767	31.058	23	.121

Predictive Capacity of the Aptitude Test

Table 7: Preliminary analyses results for logistic regression

Observed		Predicted		
	Level 1	Level 2+Level 3		Percentage Correct
Level 1	0	127		0.00
Level 2 + Level 3	0	710		100.00
Overall Percentage				84.80
Variables in the Equation					
	B	S.E.	Wald	df	Sig.	Exp(B)
Constant	1.721	.096	319.108	1	.000	5.591
Variables not in the Equation				
		Score	df	Sig.
Variables	Score	10.174	1	.001
	Overall Statistics	10.174	1	.001

The analysis found a clear distinction in predictive capacity between Level 1 and Level 2+Level 3 using ATRP scores, leading to further exploration focused on a binary classification of productivity levels. Assessment of the logistic regression model revealed significant predictors, as indicated by odds ratios, and confirmed the model's superiority over an intercept-only model, although its explanatory power was limited. The model is presented in the general formula above: $\log(p/(1-p)) = 0.028 \cdot$ ATRP Score $+ 0.495$

Table 8: Logistic regression model fit

Omnibus Tests of Model Coefficients			
Specification	Chi-square	df	Sig.
Step	10.236	1	.001
Block	10.236	1	.001
Model	10.236	1	.001
Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	702.390	.012	.021
Classification Table			
Observed		Predicted	
	Level 1	Level 2+Level 3	Percentage Correct
Not Productive	0	127	.0
Productive	0	710	100.0
Overall Percentage			84.8
Variables in the Equation					
Step 1	B	S.E.	Wald	df	Sig.	Exp(B)
Score	.028	.009	10.028	1	.002	1.028
Constant	.495	.389	1.617	1	.203	1.640

Table 9: Logistic regression model fit

Omnibus Tests of Model Coefficients			
Specification	Chi-square	df	Sig.
Step	10.236	1	.001
Block	10.236	1	.001
Model	10.236	1	.001
Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	702.390	.012	.021
Classification Table			
Observed		Predicted	
	Level 1	Level 2+Level 3	Percentage Correct
Not Productive	0	127	.0
Productive	0	710	100.0
Overall Percentage			84.8
Variables in the Equation					
Step 1	B	S.E.	Wald	df	Sig.	Exp(B)
Score	.028	.009	10.028	1	.002	1.028
Constant	.495	.389	1.617	1	.203	1.640

ATRP: Indicative Norm

Table 9: Norm of the ATRP

Level of Productivity				Classification
Ordinal Regression	Range of Score	Logistic Regression	Range of Score	
Level 1	Less than 45.00	Level 1	Below 45.00	Uncertain Productivity
Level 2	45.01 to 46.35	Level 2+	45 and above	Likely Productive
Level 3	More than 46.35			

NB: Score is expressed as percentage of correct items over total number of items.

Level 1: This category is for scores below 45.00, classified as having uncertain productivity. It indicates that individuals with scores in this range are less likely to be considered productive based on the criteria set by the ATRP.

Level 2: Scores in this range are considered likely productive but not as high as those in Level 3. This indicates a moderate level of productivity.

Level 3: This is the highest productivity level indicating that scores above 46.35 are associated with a significantly high level of productivity.

Two Tier Norm:

Level 1: Associated with uncertain productivity. Level 2+: A level of productivity that is likely or very high.

In summary, the model reveals that "ATRP Score" significantly predicts "Productivity Levels," albeit with low overall explanatory power. While the predictor "ATRP Score" influences the likelihood of productivity outcomes, its practical significance and predictive ability are limited.

Nonetheless, based on findings, the norms for Level 1 and Level 2+Level 3 classifications are established, indicating uncertain productivity for scores below 45.00 (Level 1), likely productivity for scores between 45.01 and 46.35 (Level 2), and significantly high productivity for scores above 46.35 (Level 3). Additionally, a two-tier classification system is proposed, designating Level 1 as uncertain productivity and Level 2+ as likely or very high productivity.

Table 10: Table of Specification for Aptitude Test for Research Productivity

Competencies/Topics	Number of Items (Item Location)	Number of items	Percentage (%)
Research Terminal Report	209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222	14	5.05
Citation and referencing	223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235	13	4.69
Researcher Qualities and Character	236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251	16	5.77
Research Ethics	252, 253, 254, 255, 256, 257, 258, 259, 260, 261	10	3.61
Research Productivity	262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277	16	5.77
Total		277	100

CONCLUSIONS

- The emerging dimensions of research productivity includes Personal Agency which highlights the individuals capacity to assert action in conquering research endeavors; Motivation which underlines the importance of internal drives and external incentives; One's Affordances which puts premium the resources and opportunities within reach of the researchers and Professional backgrounds capturing ones prior experience and expertise.

- The CVI Index = 1.00 after three rounds of revision, supports the quality of the ATRP in terms of content relevance and appropriateness.

- The ATRP posits a level of reliability, *Cronbach Alpha = 0.917* (CTT-based)) to a *1-PL reliability of 0.99 (IRT-based)* in its final form.

- The ATRP is unidimensional. This affirms that the test scale is accurately measuring the construct of research productivity ensuring its validity.

- The ATRP significantly fits with the Rasch Measurement Model highlighting the relative difficulty of items and the abilities of respondents, indicating a well-calibrated scale for measuring research productivity aptitude.

- The ATRP significantly predicts research productivity but with a modest effect size.

- The emerging dimensions of research productivity includes Personal Agency which highlights the individuals capacity to assert action in conquering research endeavors; Motivation which underlines the importance of internal drives and external incentives; One's Affordances which puts premium the resources and opportunities within reach of the researchers and Professional backgrounds capturing ones prior experience and expertise.

- The CVI Index = 1.00 after three rounds of revision, supports the quality of the ATRP in terms of content relevance and appropriateness.

- The ATRP posits a level of reliability, *Cronbach Alpha = 0.917* (CTT-based)) to a *1-PL reliability of 0.99 (IRT-based)* in its final form.

- The ATRP is unidimensional. This affirms that the test scale is accurately measuring the construct of research productivity ensuring its validity.

- The ATRP significantly fits with the Rasch Measurement Model highlighting the relative difficulty of items and the abilities of respondents, indicating a well-calibrated scale for measuring research productivity aptitude.

- The ATRP significantly predicts research productivity but with a modest effect size.

RECOMMENDATIONS

- Broaden application ATRP in forms like, human resource assessment, institutional assessment, policy development, and curriculum development.

- Continue the improvement cycle of the ATRP by subjecting it to validation particularly in improving its *effect size* in the predictive capacity though incorporation of other factors of the researcher in the model; as well as through further reliability testing across varying populations.

- Utilize rash model insights in re-evaluating the items in the scale to balance the item difficulties across ability levels.

Strengthen ethical considerations in the norm process by subjecting the descriptions to a consequential validity review.

REFERENCES

[1] Abramo, G., D'Angelo, C. A., & Costa, F. D. (2010). Citations Versus Journal Impact Factor as Proxy of Quality: Could the Latter ever be Preferable? Scientometrics, 84(3), 821–833.

[2] Abram, G. and Ciriaco & D'Angelo C.A. (2018). How do you define and measure research productivity? CoRR abs/1810.12830 (2018). arXiv:1810.12830 http://arxiv.org/abs/1810.12830

[3] Ackerman, T. A. (1992). A Didactic Explanation of Item Bias, Item Impact, and Item Validity from a Multidimensional Perspective. Journal of Educational Measurement, 29, 67-91. https://doi.org/10.1111/j.1745-3984.1992.tb00368.x

[4] Aiken, L.R. (1985). Psychological Testing and Assessment. 5th ed. Boston: Allyn and Bacon, Inc.

[5] Aiken. L. R. (1988). Psychological Testing and Assessment (6th Ed.). Boston: Allyn and Bacon. Inc.

[6] Aiken, L. R. (2002). Attitudes and Related Psychosocial Constructs: Theories, Assessment and Research. Thousand Oaks, CA: Sage.

[7]   Ajzen, I. (2001). Nature and operation of attitudes. Annual Review of Psychology, 52, 5–27. https://doi.org/ 10.1146/annurev.psych.52.1.27.

[8]   Åkerlind, G. S. (2008). An academic perspective on research and being a researcher: an integration of the literature. Studies in Higher Education, 33(1), 17-31.

[9]   Akinboye, J.O. (2001). Concepts in Formal Education. Ibadan: University Press Limited.

[10]   Althouse, L. Test Development: Ten Steps to a Valid and Reliable Certification Exam. Training Sales and Marketing SAS Institute, Inc. SAS Campus Drive Cary, NC 27513 919-677-8000 919-677-4444, linda.althouse@sas.com

[11]   Alhija, F. & Majdob, A. (2017). Predictors of Teacher Educators' Research Productivity. Australian Journal of Teacher Education. Tel Aviv University, Israel.

[12]   Aliyu, R.T. (2015). Construct Validity of Mathematics Test Items using the Rasch Model. An International Journal of Social Science and Humanities Research. 3(2), 22-28.

[13]   Aliyu, Taiwo R & Akinoso, Sabainah. (2016). Development and Validation of Mathematics Aptitude Test (Mat) Using the Rasch and 2-Pl Model of IRT. 16. 1-15.

[14]   American Psychological Association, American Educational Research Association, and National Council on Measurement in Education. (1974). Standards for Educational and Psychological Tests. Washington, DC: American Psychological Association.

[15]   American Psychological Association, American Educational Research Association, and National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

[16]   American Psychological Association, American Educational Research Association, and National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

[17]   American Psychological Association. (1954). Technical Recommendations for Psychological Tests and Diagnostic Techniques. Psychological Bulletin 51: 201–38.

[18]   American Psychological Association. (1966). Standards for Educational and Psychological Tests and Manuals. Washington, DC: American Psychological Association.

[19]   Anastasi. A. (1997). Psychological testing (7th Ed). New Jersey: Prentice-Hall Inc.

[20]   Andrich, D. (1992). The Application of an Unfolding Model of the PIRT Type to Measurement of Attitude. Applied Psychological Measurement, vol. 12, pp. 33-5.

[21]   Ann, P.S. (2004). Measurement, Assessment and Evaluation. Lagos: Concepts Publications Limited

[22]   Badenes-Ribera L, Silver N.C and Pedroli E. (2020) Editorial: Scale Development and Score Validation. Front. Psychol. 11:799. doi: 10.3389/fpsyg.2020.00799

[23]   Baker, F. B. (2001). The basics of item response theory. ERIC Clearinghouse on Assessment and Evaluation

[24]   Ballado, R., Morales, R. and Ortiz, R. (2014). Development and Validation of a Teacher Education Aptitude Test. International Journal of Interdisciplinary Research and Innovations. College of Education, University of Eastern Philippines, Catarman, Northern Samar.

[25]   Barley, S., & Kunda, G. (2004). Gurus, hired guns, and warm bodies. Princeton, NJ: Princeton University Press.

[26]   Beck, R. C. (2004). Motivation Theories and Principles (5th Ed.). New Jersey: Upper Saddle River.

[27]   Bennett. G. K. Seashore. H. G. & Wesman. A. G. (1948). Differential Aptitude Tests. Journal of Consulting Psychology. 12. 62.

[28]   Bland, C. J., Center, B. A., Finstad, D. A., Risbey, K. R., & Staples, J. (2006). The Impact of Appointment Type on the Productivity and Commitment of Full-Time Faculty in Research and Doctoral Institutions. The Journal of Higher Education, 77(1), 89–123. https://doi.org/10.1080/00221546.2006.11778920.

[29]   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 Health 6:149. doi: 10.3389/fpubh.2018.00149

[30]   Bond TG, Fox C. (2013) Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Mahwah, NJ: Erlbaum.

[31]   Bond, T. G., & Fox, C. (2015). Applying the Rasch model fundamental measurement in the human sciences (3rd ed.). New York, NY: Routledge.

[32]   Boud, D., & Lee, A. (Eds). (2009). Changing Practices in Doctoral Education. London: Routledge.

[33] Brew, A. (2001). Conceptions of Research: A Phenomenographic Study. Studies in Higher Education, 26(2), 271-285.

[34] Brew, A. and Boud, D. (2009). Understanding Academics' Engagement with Research. In Brew, A. and Lucas, L. (Eds.). Academic Research and Researchers. London: SRHE and the Open University Press, 189-203.

[35] Brew, A., & Lucas, L. (2009). Academic Research and Researchers. Maidenhead, UK: McGraw Hill, Society for Research into Higher Education and the Open University Press.

[36] Brew, A., Boud, D., Namgung, S., Lucas, L. & Crawford, K. (2015). Research Productivity and Academics' Conceptions of Research. Higher Education. 71. 10.1007/s10734-015-9930-6.

[37] Bryman, A., & Bell, E. (2011). Business research methods (3rd ed.). New York: Oxford University Press.

[38] Bryans, P., & Mavin, S. (2006). Visual Images: A Technique to Surface Conceptions of Research and Researchers. Qualitative Research in Organizations and Management, 1(2), 113-128.

[39] Bushra Mushtaq (2018). Validity of Tests. JOJ Nurse Health Care. Skims Nursing College, Srinagar, Juniper Publishers, India.

[40] Chatterjee, A. (2007). A Controlled and Effective Education System for a country. Retrieved on February 24, 2008 from http://www.indianstudentresearch.blogspot.com/ 2007/11/controlled-andeffectiveeducation. html.

[41] Chen, Y., Gupta, A., & Hoshower, L. (2006). Factors That Motivate Business Faculty To Conduct Research: An Expectancy Theory Analysis. Journal of education for business, 81, 179-189. doi:10.3200/JOEB.81.4.179-189.

[42] Chen, Y., Nixon, M. R., Gupta, A., & Hoshower, L. (2010). Research Productivity of Accounting Faculty: An Exploratory Study. American Journal of Business Education, 3, 101-113. Retrieved on July 12, 2021 from http://search.proquest.com.libraryproxy.griffith.edu.au/.

[43] Cochran-Smith, M. (2005). Teacher Educators as Researchers: Multiple Perspectives. Teaching and Teacher Education, 21(2), 219–225. https://doi.org/10.1016/j.tate.2004.12.003.

[44] Cole, J. R., & Zuckerman, H. (1984). The Productivity Puzzle: Persistence and Change in Patterns of Publication of Men and Women Scientists. In P. Maier, M. W. Steinkamp & M. L. Maehr (Eds.), Advances in Motivation and Achievement (Vol. 2) (pp. 217–258). Greenwich, CT: JAI Press.

[45] Creswell, J.W., Fetter, M. D., & Ivankova, N. V. (2004). Designing a mixed- method study in primary care. The Annals of Family Medicine, 2(1), 7-12.

[46] Cronbach, L.J. (1971). Test validation. In Educational measurement, vol. 2, ed. R. Thorndike, 443–507. Washington, DC: American Council on Education.

[47] Cronbach, L.J. (1988). Five Perspectives on the Validity Argument. In Test Validity, ed. H. Wainer and H. Braun, 3–17. Hillsdale, NJ: Lawrence Erlbaum.

[48] Cronbach, Lee Joseph. (1949). Essentials of Psychological Testing. New York: Harper and Brothers Publishing.

[49] Crooks, T.J., K.T. Kane, and A.S. Cohen. (1996). Threats to the Valid Use of Assessments. Assessment in Education: Principles, Policy & Practice 3, no. 3: 265–85.

[50] D'Amico, R., Vermigli, P., & Canetto, S. S. (2011). Publication Productivity and Career Advancement by Female and Male Psychology Faculty: The Case of Italy. Journal of Diversity in Higher Education, 4(3), 175–184. https://doi.org/10.1037/a0022570.

[51] Deci, E. L., & Ryan, R. M. (1985). The General Causality Orientations Scale: Self-Determination in Personality. Journal of Research in Personality, 19(2),109–134. https://doi.org/10.1016/0092-6566(85)90023-6.

[52] Demaio, T., & Landreth, A. (2004). Do Different Cognitive Interview Methods Produce   Different Results. In S. Presser, J. Rothgeb, M. Couper, J. Lessler, E. Martin, J.   Martin, & E. Singer (Eds.), Questionnaire Development and Testing Methods.   Hoboken, NJ: Wiley.

[53] Dever, M., & Morrison, Z. (2009). Women, Research Performance and Work Context. Tertiary Education and Management, 15(1), 49-62.

[54] DeVellis, R. F. (2017). Scale Development: Theory and Applications (4th ed.). Thousand Oaks, CA: Sage.

[55] Dornyei, Z. & P.Skehan. (2003). Individual Differences in Second Language Learning. In Doughty C. and M. Long (Eds), Handbook of Second Language Acquisition. Oxford: Blackwell.

[56] Ducharme, M. (1996). A Study af Teacher Educators: Research from the USA. Journal of Education for Teaching: International Research and Pedagogy, 22(1), 57-70. https://doi.org/10.1080/02607479650038427

[57] Edgar, F., & Geare, A. (2013). Factors Influencing University Research Performance. Studies in Higher Education, 38(5), 774-792.

[58]   Eisenberger, R., & Cameron, J. (1998). Reward, Intrinsic Interest, and Creativity: New Findings. American Psychologist, 53(6), 676-679. https://doi.org/10.1037/0003-066X.53.6.676.

[59]   Eloy, J. A., Svider, P. F., Cherla, D. V., Diaz, L., Kovalerchik, O., Mauro, K. M., Baredes, S., & Chandrasekhar, S. S. (2013). Gender Disparities in Research Productivity among 9952 Academic Physicians. The Laryngoscope, 123, 1865–1875. https://doi.org/10.1002/lary.24039.

[60]   Fox, K. J., & Milbourne, R. (1999). What Determines Research Output of Academic Economists? The Economic Record, 75(230), 256-267.

[61]   Fraenkel, J. R. & Wallen, N. E. (2009). How to Design Evaluate Research in Education. New York: McGraw-Hill Companies, 7th edition.

[62]   Furr, R. M. (2011). Scale Construction and Psychometrics for Social and Personality Psy

[63]   chology. New Delhi, IN: Sage Publications. https://doi.org/10.4135/9781446287866

[64]   Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and Why Incentives (Do Not) Work to Modify Behavior. Journal of Economic Perspectives, 25(4), 191-210. https://doi.org/10.1257/jep.25.4.191.

[65]   Gordon, B. 1992. Are Canadian firms under investing in training? Canadian Business Economics 1,1, 25–33.

[66]   Grapin, S. L., Kranzler, J. H., & Daley, M. L. (2013). Psychology in the Schools: Scholarly Productivity and Impact of School Psychology Faculty in APA-Accredited Programs. Psychology in the Schools, 50(1), 87-101.

[67]   Greenberg, J., & Baron, R. A. (2010). Behavior in Organizations. Upper Saddle River, NJ: Pearson Education.

[68]   Gronlund, N. E. & Linn, R. L. (1990). Measurement and Evaluation in Teaching. New York: Macmillan Publishing Company, 6th edition.

[69]   Gronlund. N. E. (1981). Measurement and Evaluation In Teaching. New York Macmillan Publishing Co.

[70]   Guilford, J.P. (1946). New Standards for Test Evaluation. Educational and Psychological Measurement 6, no. 5: 427–39.

[71]   Hancock, K. J., Baum, M. A., & Breuning, M. (2013). Women and Pre-Tenure Scholarly Productivity in International Studies: An Investigation into the Leaky Career Pipeline. International Studies Perspective, 14(4), 507-527. DOI: 10.1111/insp.12002.

[72]   Hardré, P. L. (2012). Community College Faculty Motivation for Basic Research, Teaching Research, and Professional Development. Community College Journal of Research and Practice, 36, 539-561. doi:10.1080/10668920902973362.

[73]   Hill, C.E. Thompson, B.J., & William, E. N. (1997). A guide to conducting consensual qualitative research. The counselling psychologist, 25(4), 517 – 572).

[74]   Hinkin T.R. (1995). A review of scale development practices in the study of organizations. J Manag. (1995) 21:967–88. doi: 10.1016/0149- 2063(95)90050-0

[75]   Itsuokor, D. E. (1995). Essentials of Test and Measurement. Illorin: Woye and Sons (Nig)   Ltd.

[76]   Irwing, P., & Hughes, D. J. (2018). Test Development. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development, V.I (pp. 4-47). Hoboken, NJ: Wiley

[77]   Kline P. (1993). A Handbook of Psychological Testing. 2nd Edn. London: Routledge; Taylor & Francis Group (1993).

[78]   Kline, R. B. (2009). Becoming a Behavioral Science Researcher: A Guide to Producing Research That Matters. New York: Guilford Publications.

[79]   Hambleton, R. K., & Jones, R. W. (1993). "An NCME instructional module on comparison of classical test theory and item response theory and their applications to test development." Educational Measurement: Issues and Practice, 12(3), 38-47.

[80]   Hazelkorn, E. (2004). Growing Research: Challenges for Late Developers and Newcomers. Higher Education Management and Policy, 16(1), 19-140.

[81]   Hazelkorn, E. (2008). Learning to Live with League Tables and Ranking The Experience of Institutional Leaders. Higher Education Policy, 21, 193–215. doi:10.1057/hep.2008.1.

[82]   Hedjazi, Y., & Behaves, J. (2011). Study of Factors Influencing Research Productivity of Agriculture Faculty Members in Iran. Higher Education, 62(5), 635-647. https://doi.org/10.1007/s10734-011-9410-6

[83]   Hesli, V. L., & Lee, J. M. (2011). Faculty Research Productivity: Why do Some of Our Colleagues Publish more than others? Political Science & Politics, 44(2), 393-408.

[84]   Hinkin T.R. (1995). A review of scale development practices in the study of organizations. J Manag. 21:967–88. doi: 10.1016/0149 - 2063(95)90050-0

[85]   Hinton, Martin (2011). The Assessment of Aptitude: Past and Future. (39-50) In P. Krakowian (ed). Łódź Papers in Language Testing. Warsaw, Poland.

[86]   HLS/APT Committee. (2000). Guidelines for APT decisions: Renewal, Continuing Appointment, and DSI. Revised draft # 5, March 6. Department of Health Science, State University of New York College at Brockport. Retrieved from http://www.brockport.edu/acadaff/apt/hls.doc.

[87]   Hunter, D. E., & Kuh, G. D. (1987). The "Writing Wing": Characteristics of prolific contributors to the higher education literature. The Journal of Higher Education, 58, 443-462. doi:10.2307/1981317.

[88]   Kara, F. & Çelikler, D. (2015). Development of Achievement Test: Validity and Reliability Study for Achievement Test on Matter Changing. Journal of Education and Practice. Department of Elementary Science Education, Ondokuz Mayıs University, Samsun, Turkey.

[89]   Kempa, R. (1986). Assessment in Science. London, Cambridge: London Cambridge University Press.

[90]   Kesmodel US. Cross-sectional studies – what are they good for? Acta Obstet Gynecol Scand 2018; 97:388–393.

[91]   Kiley, M., & Mullins, G. (2005). Supervisors' Conceptions of Research: What Are They? Scandinavian Journal of Educational Research, 49(3), 245-262.

[92]   Kim, D., Wolf-Wendel, L., & Twombly, S. (2011). International Faculty: Experiences of Academic Life and Productivity in U.S. Universities. Journal of Higher Education, 82(6), 720-747.

[93]   Kline, T. J. B. (2005). Psychological testing. Thousand Oaks, CA: Sage Publishing.Google Scholar

[94]   Kline P. A. (1993) A Handbook of Psychological Testing. 2nd Edn. London: Routledge; Taylor & Francis Group.

[95]   Kohn, A. (1993). Why Incentive Plans Cannot Work? Harvard Business Review, September-October, 54-63.

[96]   Kotrlik, J., Bartlett, J., Higgins, C., & Williams, H. (2002). Factors Associated aith Research Productivity of Agricultural Education Faculty. Journal of Agricultural Education, 43(3), 1-10. https://doi.org/10.5032/jae.2002.03001.

[97]   Kpolovie, P. J. (2010). Advanced research methods. Owerri: Spring Field Publisher Ltd.

[98]   Kyriazos, T. A., & Stalikas, A. (2018). Applied Psychometrics: The Steps of Scale Development and Standardization Process. Psychology, 9, 2531-2560. https://doi.org/10.4236/psych.2018.911145

[99]   Krokfors, L, Heikki, K., Katariina, S., Auli, T., Katriina, M., Riitta. J., Reijo B., & Pertti, K. (2011). Investigating Finnish teacher educators' views on research-based teacher education. Teaching Education, 22(1), 1–13. https://doi.org/10.1080/10476210.2010.542559.

[100]   Kyvik, S. (2013). The academic researcher role: enhancing expectations and improved performance. Higher Education, 65, 525-538.

[101]   Kyvik, S., & Teigen, M. (1996). Childcare, research, collaboration, and gender differences in scientific productivity. Science Technology Human Values, 21(54), 54-71. https://doi.org/10.1177/016224399602100103.

[102]   Lawshe, C. H. (1975). A Quantitative Approach to Content Validity. Personnel Psychology, 28, 563-575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

[103]   Lehmann, D. R., & Hulbert, J. (1972). Are Three-Point Scales Always

[104]   Leahey, E. (2006). Research Specialization as a Missing Link. Gender & Society, 20(6), 754-780. https://doi.org/10.1177/0891243206293030.

[105]   Lee, S., & Bozeman, B. (2005). The Impact of Research Collaboration on Scientific Productivity. Social Studies of Science, 35(5), 673-702.

[106]   Levin, S. G., & Stephan, P. E. (1991). Research Productivity over the Life Cycle: Evidence For Academic Scientists. The American Economic Review, 81, 114-132. Retrieved on July 12, 2021 from http://www.jstor.org.libraryproxy.griffith.edu.au/.

[107]   Linacre, J. M. (2022). Winsteps® [Version 5.3.2] [Computer Software]. Portland, Oregon: Winsteps.com.

[108]   Linn, R. L., & Gronlund, N.E. (2000). Measurement and assessment in teaching (8thed.).Delhi:

[109]   Lissoni, F., Mairesse, J., Montobbio, F., & Pezzoni, M. (2011). Scientific productivity and academic promotion: a study on French and Italian physicists. Industrial and Corporate Change, 20(1), 253–294. https://doi.org/10.1093/icc/dtq073.

[110]   Lynn, M. R. (1986). Determination and Quantification of Content Validity. Nursing Re

[111]   search, 35, 382-386. https://doi.org/10.1097/00006199-198611000-00017

[112]   Macklem, Gayle L. (1990). Measuring Aptitude: Practical Assessment, Research, and Evaluation: Vol. 2, Article 5. Scholar Works UMass Amherst Publishing.

[113] Maggay, Jake (2017). College Aptitude Test Simple Checker. Asia Pacific Journal of Multidisciplinary Research, Vol. 5, No. 3.

[114] Mankar. J. & Chavan. D. (2013). Differential aptitude testing of youth. International Journal of Scientific and Research Public. 3 (7). 1-6.

[115] Marais. A. C. (2007). Using the Differential Aptitude Test to Estimate Intelligence and Scholastic Achievement at Grade Nine Level. Thesis. Pretoria: University South Africa.

[116] Mawoki, M. A., & Babandako, A. Y. (2011). An Evaluation of Staff Motivation, Dissatisfaction and Job Performance in an Academic Setting. Australian Journal of Business and Management Research, 1(9), 1-13. Retrieved on July 12, 2021 from http://www.ajbmr.com/articlepdf/AJBMR_17_25i1n9a1.pdf

[117] Mehrens, W. (1997). The Consequences of Consequential Validity. Educational Measurement: Issues and Practice 16, no. 2: 16–8.

[118] Messick, S. 1989. Validity. In Educational Measurement, vol. 3, ed. R.L. Linn, 13–103. New York: American Council on Education, MacMillan.

[119] Meyer, J. P., & Allen, N. J. (1991). A Three-Component Conceptualization of Organizational Commitment. Human Resource Management Review, 1, 61-89. doi:10.1016/1053-4822(91)90011-Z.

[120] Meyer, J.H.F., M.P. Shanahan, and R.C. Laugksch. (2005). Students' Conceptions of Research: I – A Qualitative and Quantitative Analysis, Scandinavian Journal of Educational Research 49: 225–44.

[121] Milfont, T. L., & Fischer, R. (2010). Testing Measurement Invariance across Groups: Ap

[122] plications in Cross-Cultural Research. International Journal of Psychological Research, 3, 111-121. https://doi.org/10.21500/20112084.857

[123] Morrison, K. M., & Embretson, S. (2018). Item Generation. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development, V.I (pp. 46-96). Hoboken, NJ: Wiley. https://doi.org/10.1002/9781118489772.ch3

[124] Moss, P.A. (1998). The Role of Consequences in Validity Theory. Educational Measurement: Issues and Practice 17, no. 2: 6–12.

[125] Nakhaie, M. R. (2002). Gender Differences in Publication among University Professors In Canada. Canadian Review of Sociology and Anthropology – La Revue Canadienne de Sociologie Et d'Anthropologie, 39(2), 151–179. https://doi.org/10.1111/j.1755-618X.2002.tb00615.

[126] Nenty, H. J. (2005). The Application of Item Response Theory in Strengthening Assessment Role in the Implementation of National Education Policy.

[127] Nguyen, Quy Huu. (2015). Factors Influencing the Research Productivity of Academics at the Research-Oriented University in Vietnam. School of Education and Professional Studies, Vietnam.

[128] Nunnally, J.O. (2004). Memory, New York: Macmillan Press.

[129] Nunnally, J.C. and Bernstein, I.H. (1994) Psychometric theory. McGraw-Hill, Inc., New York.

[130] Nuqui, A., & Cruz, R. (2012). Determinants of Faculty Research Productivity in Augustinian Higher Education Institutions in Luzon. IAMURE International Journal of Education, 3, 56-74. Retrieved from http://dx.doi.org/10.7718/iamure.ije.v3i1.191.

[131] Nygaard, L. P. (2015). Publishing and Perishing: An Academic Literacies Framework for Investigating Research Productivity. Studies in Higher Education. DOI: 10.1080/03075079.2015.1058351.

[132] Odili, J. N., Osadebe, P. U. Aliyu, R. T. (2015). Assessment of Stability of Item Parameter in a Mathematics Achievement Test under The Rasch Model. A paper published in Journal of Association of Educational Researcher and Evaluators of Nigeria (ASSEREN), 1(1), 1-8.

[133] Padilla-Gonzalez, L., Metcalfe, A. S., Galaz-Fontes, J. F., Fisher, D., & Snee, I. (2011). Gender Gaps in North American Research Productivity: Examining Faculty Publication Rates in Mexico, Canada, and the U.S. Compare: A Journal of Comparative and International Education, 41(5), 649-668.

[134] Parasuraman, A., Grewal, D., & Krishnan, R. (2006). Marketing Research (pp. 68-69).

[135] Boston, New York: Houghton Mifflin Company. Presser, S., & Blair, J. (1994). Survey Pretesting: Do Different Methods Produce Different Results? In P. Marsden (Ed.), Sociology Methodology (Vol. 24, pp. 73-104). Washington DC: American Sociological Association.

[136] Price, L. R. (2017). Psychometric Methods: Theory into Practice. New York: The Guilford Press

[137] Popham, W.J. (1997). Consequential Validity: Right Concern – Wrong Concept. Educational Measurement: Issues and Practice 16, no. 2: 9–13.

[138] Prozesky, H. (2008). A Career-History Analysis Of Gender Differences In Publication Productivity Among South African Academics. Science Studies, 21(2), 47–67.

[139] Quimbo, M.-A. T., & Sulabo E. C. (2014). Research Productivity and its Policy Implications in Higher Education Institutions. Studies in Higher Education, 39(10), 1955-1971.

[140] Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Danish Institute for Educational Research.

[141] Rauber, M., & Ursprung, H. W. (2007). Life Cycle and Cohort Productivity in Economic Research: The Case of Germany. CESifo working paper, No. 2093.

[142] Reckase, M. D. (2009). Multidimensional item response theory. Springer.

[143] Rehman, A. (2007). Development and validation of objective test items analysis in the subject physics for class IX in Rawalpindi city. Retrieved on July 12, 2021 from the International Islamic University, Department of Education Web site: http://eprints.hec.gov.pk/2518/1/2455.htm.

[144] Robinson, G., & Gould, M. (2000). What Are The Attitudes Of General Practitioners Towards Research? British Journal of General Practice, 50(454), 390-392.

[145] Ruth. A. (1971). A Study of the Predictability of High School Grades and the Differential Aptitude Tests for Success in Vocational Programs in Health Careers. Thesis. Wisconsin: Wisconsin University.

[146] Sax, L. J., Hagedorn, L. S., Arredondo, M., & DiCrisi, F. A. (2002). Faculty Research Productivity: Exploring the Role of Gender and Family Related Factors. Research in Higher Education, 43(4), 423-446.

[147] Sax. G. (1980). Principles of Educational and Psychological Measurement and Evaluation. Belmont: Wardsworth Publication Company. Inc.

[148] Sawilowsky, S. S. (2007). Construct Validity. In N. J. Salkind (Ed.), Encyclopedia of Measurement and Statistics (pp. 178-180). Thousand Oaks, CA: Sage.

[149] Serenko, & Bontis. (2004). Meta-Review of Knowledge Management and Intellectual Capital Literature: Citation Impact and Research Productivity Rankings. Knowledge and Process Management, 11(3), 185-198.

[150] Setiawati, F.A. (2020). Aptitude Test's Predictive Ability for Academic Success in Psychology Student. Psychological Research and Intervention, Department of Psychology, Universitas Negeri Yogyakarta, Indonesia.

[151] Shariatmadari, M., & Mahdi, S. (2012). Barriers to Research Productivity in Islamic Azad University: Exploring Faculty Members' Perception. Indian Journal of Science and Technology, 5(5), 2765- 2769.

[152] Shepard, L.A. (1993). Evaluating Test Validity. Review of Research in Education 19: 405–50.

[153] Silver, E. A. (2009). Some Ideas on Enhancing Research Productivity. International Journal of Production Economics, 118, 352–360.

[154] Skehan, P. (1986). Where Does Language Aptitude Come From? In P. Meara (ed.) Spoken language. Papers from the annual meeting of the British Association for Applied Linguistics (Edinburgh, September 1985).

[155] Skoie, H. (2000). Faculty Involvement in Research in Mass Higher Education: Current Practice and Future Perspectives in the Scandinavian Countries. Science and Public Policy, 27, 409- 419.

[156] Smeby, J. C., & Try, S. (2005). Departmental Contexts and Faculty Research Activity in Norway. Research in Higher Education, 46(6), 593-619.

[157] Sridhar S., Dias, B., & Sequeira, A. H. (2010). Measuring Faculty Productivity – A Conceptual Review. St Aloysius College-AIMIT Working Paper Series 1(1), 2-25.

[158] Stack, S. (2004). Gender, Children and Research Productivity. Research in Higher Education, 45(8), 891-920.

[159] Stake, R. (2006). Multiple case study analysis. New York: Guilford.

[160] Stansfield, C.W. & D.J. Reid (2004). The Story Behind the Modern Language Aptitude Test: An Interview with John B. Carroll. In Language Assessment Quaterly 1(1), 43-56.

[161] Stansfield, C.W. (2009). STLI Language Aptitude Testing. Retrieved on July 21, 2021 from www.2lti.com.

[162] Streiner, D. L., Norman, G. R., & Cairney, J. (2015). Health Measurement Scales: A Practical Guide to Their Development and Use (5th ed.). Oxford, UK: Oxford University Press. https://doi.org/10.1093/med/9780199685219.001.0001

[163] Stobart, G. (2001). The Validity of National Curriculum Assessment. British Journal of Educational Studies 49, no. 1: 26–39.

[164] Stubb, J., Pyhältö, K., & Lonka, K. (2014). Conceptions of Research: The Doctoral Student Experience in Three Domains. Studies in Higher Education, 39(2), 251-264.

[165] Suitor, J. J, Mecom, D., & Feld, I. S. (2001). Gender, Household Labor, and Scholarly Productivity among University Professors. Gender Issues, 19, 50-67.

[166] Sumintono, Bambang & Widhiarso, Wahyu (2015), Aplikasi Pemodelan Rasch Pada Asesment Pendidikan, Bandung: Trim Komunikasi, Cet I.

[167] Sumintono, Bambang & Widhiarso, Wahyu (2015), Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial, Bandung: Trim Komunikasi, Cet II.

[168] Swain, S. K., Pradhan, C., &Khotoi, S. P. K. (2000). Educational measurement: Statistics and guidance. Ludhiana: Kalyani.

[169] Talib, R. N.F. O. , Hassan, S. A. & Umi Kalthom Abdul Mana, U.K. (2020). "Development and Validation of an Instrument to Measure STEM Teachers' Instructional Preparedness". Asian Journal of University Education (AJUE)Volume 16, Number 3, pp. 193 - 205

[170] Taiwo, A. R & Oyebola, A. S.(2016)."Development and Validation of Mathematics Aptitude Test (Mat) Using the Rasch and 2-Pl Model of IRT" IJES, Volume 16 No. 2.

[171] Thurstone, L. L. (1947). Multiple-Factor Analysis. University of Chicago Press, Chicago, USA.

[172] Tien, F. F. (2000). To What Degree does the Desire for Promotion Motivate Faculty to Perform Research? Testing the expectancy theory. Research in Higher Education, 41, 723-752. doi:10.1023/a:1007020721531.

[173] Tien, F. F., & Blackburn, R. T. (1996). Faculty Rank System, Research Motivation, and Faculty Research Productivity: Measure Refinement and Theory Testing. The Journal of Higher Education, 67, 2-22. doi:10.2307/2943901.

[174] Toplis, K. (1991). Behavioral Modification Techniques, New York: Prentice Hall Publishers.

[175] Vermunt, J. (2005). Conceptions of Research and Methodology Learning: A Commentary on the Special Issue. Scandinavian Journal of Educational Research, 49(3), 329–334.

[176] Waltz, C. W., & Bausell, R. B. (1981). Nursing Research: Design, Statistics and Computer Analysis. Philadelphia, PA: F.A. Davis.

[177] Wechsler, D. (1958). The Measurement and Appraisal of Adult Intelligence (4th ed.). Baltimore, MD: Williams and Wilkins. https://doi.org/10.1037/11167-000

[178] Wichian, S. Na., Wongwanich, S., & Bowarnkitiwong, S. (2009). Factors Affecting Research Productivity of Faculty Members in Government Universities: Lisrel and Neural Network Analysis. Journal of Social Science, 30, 67-78.

[179] Wiersma, W. (1986). Research Methods in Education: An Introduction. (4th ed) Boston: Allyn and Bacon Inc.

[180] Williams, W. W., & Ceci, S. J. (2012). When Scientists Choose Motherhood: A Single Factor goes a Long Way in Explaining the Dearth of Women In Math-Intensive Fields. How can we address it? American Scientist, 100(2), 138.

[181] Willis, G., Schechter, S., & Whitaker, K. (2000). A Comparison of Cognitive Interviewing, Expert Review and Behavior Coding: What Do They Tell Us? In Proceedings of the Section on Survey Methods (pp. 28-37). Alexandria, VA: American Statistical Association.

[182] Wolming, Simon & Wikström, Christina (2010). The Concept of Validity in Theory and Practice. Department of Educational Measurement, Umeå University, Umeå. Routledge Taylor and Francis Group, Sweden.

[183] Woodworth, R. S. and Marquis, D. G. (1963). Psychology.20th Edition, Strand: Methuen and Co. Ltd.

[184] Worthen, B. R., Borg, W. R. & White, K. R. (1993). Measurement and Evaluation in the Schools. New York: Longman.

[185] Wright, B.D. (1994). Unidimensionality coefficient. Rasch Measurement Transactions, 8(3), p.385.

[186] Zangirolami-Raimundo J, Echeimberg JO, Leone C. Research methodology topics: Cross-sectional studies. Journal of Human Growth and Development. 2018; 28(3):356360. DOI: http://dx.doi.org/10.7322/ jhgd.152198