IMPROVING EDUCATIONAL STATISTICS ON HOTS THROUGH RELIGIOUS MODERATION AND COGNITIVE TEST

Enung Nugraha[1*)], Encep Syarifudin[2)], Aan Ansori[3)], Lisda Cahya Velia[4)], Vania Nurachma[5)]

[1,2,3,4,5] Universitas Islam Negeri Sultan Maulana Hasanuddin Banten, Indonesia

e-mail: enung.nugraha@uinbanten.ac.id[*1)]; encep.syarifudin@uinbanten.ac.id[2)]; aan.ansori@uinbanten.ac.id[3)]; lisda_cahyavelia@gmail.com[4)]; vanianurachma138@gmail.com[5)]

Abstract

This study aims to determine the development and effectiveness of cognitive test instruments to train students' higher-order thinking skills (HOTS) in educational statistics courses. The method employed is research and development (R&D) based on the ADDIE model, which includes analysis, design, development, implementation, and evaluation. The research sample was purposively selected, consisting of PGMI students at Universitas Islam Negeri Sultan Maulana Hasanuddin Banten who were enrolled in educational statistics courses. The research findings indicate that lecturers need to refine syllabi according to the specified curriculum, develop cognitive test instrument grids in the form of questions with indicators, and design statistical questions based on religious moderation using the ADDIE model. The development of cognitive test instruments based on religious moderation is effectively used to measure higher-order thinking skills (HOTS) in educational statistics for students. The implications of this research highlight efforts to train students' higher-order thinking skills through educational statistics based on religious moderation using cognitive test instruments.

Keywords: *Religious Moderation; Educational Statistics; HOTS; Cognitive Test; Students*

Abstrak

Penelitian ini bertujuan untuk menentukan pengembangan dan efektivitas instrument tes kognitif untuk melatih keterampilan berpikir tingkat tinggi (HOTS) mahasiswa dalam mata kuliah statistic pendidikan. Metode yang digunakan adalah penelitian dan pengembangan (R&D) berdasarkan model analisis, desain, pengembangan, implementasi, dan evaluasi (ADDIE). Sampel penelitian dipilih secara purposive, yaitu mahasiswa PGMI universitas Islam Negeri Sultan Maulana Hasanuddi Banten yang mengikuti mata kuliah statistic pendidikan. Hasil penelitian menunjukkan bahwa dosen perlu mematangkan silabus sesuai dengan kurikulum yang ditentuan, membuat kisi-kisi instrument tes kognitif dalam bentuk soal dengan indicator, serta menyusun soal-soal statistic yang berbasis pada moderasi beragama sesuai dengan model ADDIE. Pengembangan instrument tes berbasis moderasi beragama efektif digunakan untuk mengukur keterampilan berpikir tingkat tinggi (HOTS) dalam statistic pendidikan mahasiswa. Implikasi penelitian ini menunjukkan upaya pelatihan keterampilan berpikir tingkat tinggi mahasiswa melalui statistic pendidikan yang berbasis moderasi beragama menggunakan instrument tes kognitif.

Kata Kunci: *Moderasi Beragama; Statistik Pendidikan; HOTS; Tes Kognitif; Mahasiswa*

* Corresponding author: enung.nugraha@uinbanten.ac.id

ENUNG NUGRAHA, ENCEP SYARIFUDIN, AAN ANSORI, LISDA CAHYA VELIA,
VANIA NURACHMA

مستخلص

يهدف هذا البحث إلى تعيين تطوير وفعالية أدوات الاختبار المعرفي لتدريب الطلاب على مهارات التفكير العليا (HOTS) في دراسة الإحصاء التربوي. الطريقة المستخدمة هي البحث والتطوير (R&D) بناءً على نموذج التحليل والتصميم والتطوير والتنفيذ والتقييم(ADDIE). عينة البحث مختار بشكل قصدي، وهم طلاب PGMI في جامعة السلطان مولانا حسن الإسلامية الحكومية بنتن الذين شاركوا دراسة الإحصاء التربوي. وأظهرت نتائج البحث أن المحاضرين يحتاجون إلى استكمال إعداد التعليم مناسب بالمنهج الدراسي المحدد، وإنشاء شبكة من أدوات الاختبار المعرفي على شكل أسئلة ذات مؤشرات، وتأليف أسئلة إحصائية على أساس الاعتدال الديني وافق بنموذج ADDIE. تطوير أداة اختبار معتمد على الاعتدال الديني فعالاً لقياس مهارات التفكير العليا (HOTS) في إحصاءات تعليم الطلاب. وتظهر مضامين هذا البحث الجهود المبذولة لتدريب مهارات التفكير العليا لدى الطلاب من خلال الإحصاءات التعليمية القائمة على الاعتدال الديني باستخدام أدوات الاختبار المعرفي.

الكلمات الرئيسية: الاعتدال الديني، إحصائيات التعليم، HOTS، الاختبارات المعرفية، الطلاب

A. Introduction

To build quality tertiary institutions, it is necessary to improve the competence of lecturers as professional educators, one of which is the skill component of educational resources to support other parts, including learning systems, facilities and infrastructure, learning media, evaluation and evaluation, and assessment. One of the essential components in the review is competence, which determines how well the lecturer performs in carrying out its functions during the learning process.[1]

Based on the competencies that lecturers must have, pedagogic competence is vital to the assessment of learning results and the development of students.[2] Evaluating learning outcomes is an effort to evaluate whether the educational objectives have been accomplished by carrying out tests or assessments of learning outcomes.[3]

What is included in the pedagogic competency component includes lecturer understanding, instructional design and implementation, evaluation of educational outcomes, and student development in actualising their various potentials. For that reason, [4] we argue that

---

[1] Suyadi et al., "Academic Reform and Sustainability of Islamic Higher Education in Indonesia," *International Journal of Educational Development* 89, no. 102534 (2022): 1–11, https://doi.org/10.1016/j.ijedudev.2021.102534.

[2] Enung Nugraha and Muhamad Fauzi, "Digital Learning Education Development Towards Modern Islamic Culture: A Strengthening 'Merdeka Belajar' Strategy," *Al-Qalam: Jurnal Kajian Keislaman* 37, no. 2 (December 2020): 1–20, https://doi.org/http://dx.doi.org/10.32678/alqalam.v37i2.3667.

[3] Rina Harjanti, Yetty Supriyati, and Wardani Rahayu, "Evaluation of Learning Programs at Elementary School Level of 'Sekolah Alam Indonesia (SAI)'. (Evaluative Research Using Countenance Stake's Model)," *American Journal of Educational Research* 7, no. 2 (2019): 125–32, https://doi.org/10.12691/education-7-2-2.

[4] Carter et al. (2022)

the way to provide a sense of personality competence is by knowing the abilities of more stable students, being wise and prudent, having a character or noble character, being authoritative, and being a role model for many people sustainably accompanied by self-evaluation and self-development.

Lecturer performance greatly determines the success of teaching and learning activities on campus.[5] The lecturer's performance has provided material to students that these students will understand. When students understand the material presented by the lecturer, the lecturer's performance can determine the success of teaching and learning activities on campus. Lecturer performance can be assessed by asking validity and reliability questions.

The skill to engage in higher-order thinking among students, especially students enrolled in the Madrasah Ibtidaiyah Teacher Training Program, Sultan Maulana Hasanudin Banten State Islamic University, is expected to engage in advanced cognitive processes. This is a concern that must be resolved. Higher-level reasoning abilities are essential to solve a problem in the scientific field. In addition, this ability will form a way of thinking after completing their studies at university but in the world of industry or work.[6] However, unfortunately in the university assessment process, they are still trying to maximise their curriculum development according to work demands.

In achieving a good level of knowledge and assessment, [7] said that a test or assessment instrument is needed to provide feedback for lecturers or students because the results of these tests can be a reference for improvement and development in the future and serve as motivation for students to explore academic abilities to achieve more.

The phenomenon that occurs in the field is that many students believe that statistics are difficult to learn because it is related to numbers. Even though the discussion in statistics has a gradual sequence of material starting from introductory material to advanced material, to study the following material, of course, a student has learned and mastered the previous material because the material is interrelated and continuous. Thus, looking at statistics is relatively easy as long as it is by order of the material that has been determined.

Based on the existing phenomena and opinions, a solution is needed to solve this

---

[5] Esther E. van Dijk et al., "What Makes an Expert University Teacher? A Systematic Review and Synthesis of Frameworks for Teacher Expertise in Higher Education," *Educational Research Review* 31, no. 100365 (2020): 1–16, https://doi.org/10.1016/j.edurev.2020.100365.

[6] Christine Velez et al., "Application of the Indigenous Evaluation Framework to a University Certificate Program for Building Cultural Awareness in Science, Technology, Engineering, and Mathematics," *Evaluation and Program Planning* 92, no. 102066 (2022): 1–9, https://doi.org/10.1016/j.evalprogplan.2022.102066.

[7] Nugraha et al. (2019)

problem so that students, especially PGMI students at UIN SMH Banten, have increased higher-order thinking skills. Lecturers can do this by developing a cognitive test instrument for educational statistics courses that includes High-Order Thinking Skill (HOTS) questions.

The test is a measuring tool to determine student abilities. Let us look at the use of tests in daily educational practice. Two goals are to be achieved: knowing the achievement status of students, which is then compared with internal or external criteria, and understanding students' potential based on the information about the situation mentioned above[8]. Students can be used as a basis to guide judgments related to the placement and distribution of these students in the future.

The test questions made by the teacher are generally prepared in a hurry and are only tried out after being used. As a result, many of the items used in the exam need to produce correct or accurate data about student learning outcomes.[9] If the decision taken is right, which is caused by the instrument used to collect data that needs to be adequately prepared, then, of course, such a decision is a decision that cannot be accounted for.

Educational tests are structured using several items. The score on each item is added to the score l others to achieve a total score. Tests used for evaluation purposes should meet two essential requirements: validity and reliability. The importance of evaluating [10] states that analysis items are an integral part of the validity and reliability of a test. This item analysis is carried out by an evaluator who usually evaluates after all the items given to students are returned and the score has been determined. Multiple choice items, in principle, consist of a subject matter or problem and a list of recommended options to be filled out by students who want to be evaluated.[11]

In addition, each test item is also distinguished into two essential parts, namely the subject matter and alternative answers.[12] It is called the subject matter because the central part of the sentence contains problematic learning outcomes that students want to ask about. If you go deeper, the subject matter of the type of choice objective test can also be categorized into

---

[8] Meyda Mustika Nugraheni, Anam Sutopo, and Djalal Fuadi, "Penilaian Afektif Dalam Pembelajaran Tematik Masa Pandemi Covid-19 Di Sekolah Dasar," *ELSE (Elementary School Education Journal) : Jurnal Pendidikan Dan Pembelajaran Sekolah Dasar* 5, no. 2 (2021): 210, https://doi.org/10.30651/else.v5i2.8233.

[9] Mayke W.C. Vereijken et al., "Student Learning Outcomes, Perceptions and Beliefs in the Context of Strengthening Research Integration into the First Year of Medical School," *Advances in Health Sciences Education* 23, no. 2 (2018): 371–85, https://doi.org/10.1007/s10459-017-9803-0.

[10] Owston (2008)

[11] Eko Triatmojo, Wibowo -, and Yuli Agustina, "Improving The Motivation and Learning Outcomes Students with Applying 'Everyone Is Teacher Here' Learning Method," *Jurnal Pendidikan Bisnis Dan Manajemen* 3, no. 2 (2017): 122–30, https://doi.org/10.17977/um003v3i22017p122.

[12] Prakash Srinivasan, *Evaluation in Education* (Raleigh: Laxmi Book Publication, 2020).

two forms: direct and indirect. Usually, evaluators plan the answer section systematically and provide careful responses, while the rest may contain errors. The wrong answer is, by its function, to confuse students who need to study better. Because of this function, some wrong answers are often called distractor answers. This trap answer is also often referred to by some teachers as an alternative answer or optional answer and ser. It serves students' attention in choosing the correct answer if they study poorly. Multiple-choice questions are questions whose answers must be selected from several possible solutions that have been provided.[13]

Each multiple-choice question comprises a subject matter (stem) and answers choices (options). The response options include a correct answer and a distractor. The correct answer is the accurate or most appropriate option. A distractor is an incorrect or misleading answer designed to divert the attention of someone to be tricked into choosing it if he does not master the material well.[14]

Religious moderation is a perspective, an attitude.[15] Opting to emphasize balance in moral values and character as a reflection of individual or collective religious views is undoubtedly influenced by the surrounding diversity and social context.

Assessment of multiple-choice test items can generally be divided into two types, namely evaluations that take into account wrong answers and do not take into account wrong answers. Some teachers assess by considering the wrong item answers to think the wrong answers counted and used as a fine to deduct the correct answers.[16] Assessment without taking into account wrong answers. That is, the wrong answer does not affect the value of the correct answer. The final score of the multiple-choice test items equals the number of correct answers. If almost all students get poor scores, the test is too difficult. It is better if all students get good scores; it can be interpreted that the test is too easy. Of course, the interpretation of the test questions would be different if the trial had been prepared and possible to meet the test requirements.[17]

The problem identification included the pedagogic competency component, including

---

[13] Widana I Wayan, "HIGHER ORDER THINKING SKILLS ASSESSMENT (HOTS)," *JISAE: Journal of Indonesian Student Assessment and Evaluation* 3, no. 1 (2017): 32–44, https://doi.org/https://doi.org/10.21009/jisae.v3i1.4859.

[14] Sitti Saenab et al., "Recode to Re-Code: An Instructional Model to Accelerate Students' Critical Thinking Skills," *Education Sciences* 11, no. 1 (2021): 1–14, https://doi.org/10.3390/EDUCSCI11010002.

[15] Enung Nugraha et al., *Moderasi Keislaman Dalam Bingkai Kebhinekaan* (Serang: Media Madani, 2021).

[16] Victoria Rolfe, Kajsa Yang Hansen, and Rolf Strietholt, "Integrating Educational Quality and Educational Equality into a Model of Mathematics Performance," *Studies in Educational Evaluation* 74, no. May (2022): 101171, https://doi.org/10.1016/j.stueduc.2022.101171.

[17] Nadia Parsazadeh et al., "The Construction and Validation of a Usability Evaluation Survey for Mobile Learning Environments," *Studies in Educational Evaluation* 58, no. May (2018): 97–111, https://doi.org/10.1016/j.stueduc.2018.06.002.

the lecturer's understanding of the lecturer, the planning and execution of instruction, assessment of learning results, and fostering student development to realize their full potential. The way to provide an understanding of personality competence is by knowing the abilities of more stable students, being wise and prudent, having the character or having noble character, being authoritative, and being a role model for many people sustainably accompanied by self-evaluation and self-development. Many lecturers use today's media in their performance of lecturers when teaching and learning activities on campus. With the aim that students do not get bored when learning in class and add insight to students. This is one of the lecturers' performance, quality, and learning assessments for students. With the background of the problem above, there are several problem formulations: What is the process for developing a cognitive test instrument to measure advanced cognitive skills in educational statistics courses at PGMI of UIN SMH Banten? What is the effectiveness of cognitive test instruments in improving higher-order thinking skills for PGMI UIN SMH Banten students in educational statistics courses?

This study aimed to determine the process of developing and the effectiveness of cognitive test instruments in improving higher-order thinking skills for PGMI UIN SMH Banten students in educational statistics courses.

B. Methods

This research is classified as Research and Development, a process aimed at creating a new product or enhancing an existing one, with clear accountability. The research methodology employed is the Research and Development approach using the ADDIE model. The ADDIE model consists of five phases: Analysis, Design, Development, Implementation, and Evaluation.[18] The development process involves multiple evaluations by a team of experts, individual research participants, and tests conducted on both small and large scales (field trials), along with revisions to refine the final product. Although the development procedure may be expedited, it still incorporates testing and revisions to ensure that the developed product meets the established criteria, is empirically validated, and is free from errors.

In [19] our opinion, The ADDIE model is regarded as a more logical and comprehensive framework compared to other models. As such, it is applicable to the development of various

---

[18] Hadi Erfani, *Research and Development* (Basin: LAP Lambert Academic Publishing, 2019), https://doi.org/10.12968/npre.2007.5.5.23745.

[19] Stigler et al. (2020)

types of products, including instructional models, learning strategies, teaching methods, media, and educational materials.

The ADDIE model offers the advantage of being simpler, well-structured, and extensively utilized in the creation of effective learning programs and products, all while being validated by experts. However, the ADDIE model has drawbacks, such as the long time and too much emphasis on content rather than experiential learning. This weakness is overcome by careful preparation and the addition of components that are lacking, such as learning experiences. ADDIE is also an instructional design model that builds performance-based materials and products.[20]

The research location was UIN Sultan Maulana Hasanuddin Banten, Jalan Jenderal Sudirman No. 30 Panancangan Cipocok Jaya, Sumurpecung, Kec. Serang, City of Serang, Banten 42118. This research was conducted on Monday - 07 - November - 2022. There were 70 PGMI students taking semester five statistics courses.

The population in this study consists of 70 fifth-semester students from the PGMI Department enrolled in the Educational Statistics course. A population refers to a group of individuals or subjects within a specific area and time frame who possess particular characteristics relevant to the research. This study employs purposive sampling, a non-random sampling technique in which researchers deliberately select participants based on specific criteria. The selection criteria include students who have completed prerequisite courses related to statistics, demonstrated active participation in academic discussions, and are willing to engage in the research process. This approach ensures that the selected participants can effectively address the research questions.

The research procedures, in line with the stages of the ADDIE Model, is carried out as follows. [21][22]:

1. Analysis

   At this stage, the primary task is to assess the need for developing instructional materials in relation to the learning objectives.

2. Design

   The design phase involves various plans for the creation of instructional materials.

---

[20] Hiroyuki Yoshikawa, *Design Methodology for Research and Development Strategy (Realizing a Sustainable Society)*, *Japan Science and Technology Agency (JST)* (Tokyo: Center for Research and Development Strategy Japan Science and Technology Agency, 2012).

[21] (Tegeh et al., 2014:

[22] van den Akker, 1999)

3. Development

The development phase in the ADDIE Model involves turning the product design into a tangible result, such as instructional materials. In this research, the development step includes activities to create and revise these materials. A conceptual framework for developing instructional materials was outlined during the design phase. During development, this framework is translated into instructional materials that are ready to be implemented to meet the objectives. Two main goals need to be accomplished during the development of teaching materials: 1) Creating or revising teaching materials that align with the formulated learning objectives, and 2) Selecting the most effective teaching materials to achieve the learning outcomes.

4. Implementation

The implementation phase in this study involves putting the developed teaching materials into practice in real classroom settings. During this phase, the designed teaching materials are applied to actual learning situations. The developed materials are delivered as part of the learning process. Following the implementation of the materials in learning activities, an initial evaluation is conducted to gather feedback to improve the development of future teaching materials. The main objectives of the implementation phase are: 1) Guiding students to meet the learning objectives, 2) Facilitating problem-solving to address challenges students previously encountered in the learning process, and 3) Ensuring that students' skills improve by the end of the learning process.

5. Evaluation

Evaluation is the final phase of the ADDIE learning system design model. It is a process conducted to assess the value of the developed teaching materials in the learning process. Evaluation occurs in two forms: formative and summative. Formative evaluations are conducted at the end of each in-person session (weekly), while summative evaluations are done at the conclusion of the entire course (semester). Summative evaluation assesses the final competencies or learning objectives intended to be achieved. The results of the evaluation are used to provide feedback on the development of the teaching materials. Based on the evaluation findings, revisions are made to address any gaps where the objectives of the teaching material development were not fully met.

IMPROVING EDUCATIONAL STATISTICS ON HOTS THROUGH RELIGIOUS MODERATION AND COGNITIVE TEST

Table 1: Cognitive Test Instrument Grid

Material	Indicator	Question Form/Number	Number of questions
Able to recognise the scope and use of statistics in the field of education and the distribution of statistical data	1. Explain some statistical terms 2. Describe the distribution of statistics 3. Identify statistical characteristics 4. Give examples of functions, roles and uses of statistics 5. Explain the meaning of data 6. Explain the requirements so that the data can be analysed 7. Classifying data classification 8. Describe ways of collecting data 9. Describe the instruments used in collecting data. Give examples of the presentation of statistical data	MC/1, 2 MC/3, 4 MC/5, 6 MC/7, 8 MC/9, 10 MC/11, 12 MC/13, 14 MC/15, 16 MC/17, 18	18
Create a Frequency Distribution Table	1. Explain terms related to the frequency distribution 2. Describe the steps for compiling a frequency distribution table 3. Create a frequency distribution table 4. Presenting a single data into a frequency distribution table 5. Presenting group data into a frequency distribution table 6. Make a cumulative frequency distribution table 7. Make a relative frequency distribution table 8. Presenting data from the frequency distribution table into a histogram 9. Presenting data from the frequency distribution table into polygons	MC/19, 20 MC/21, 22 MC/23 MC/24 MC/25 MC/26 MC/27 MC/28 MC/29	11
Central Value Measure	1. Explain the meaning of the centre value 2. Classify the various central values 3. Calculate the mean 4. Calculate the median 5. Counting mode 6. Count, quartiles 7. Calculate deciles 8. Calculate percentile	MC/30, 31 MC/32, 33 MC/34, 35 MC/36, 37 MC/38, 39 MC/40, 41 MC/42, 43 MC/44, 45	16
Dispersion Size	1. Explaining dispersion 2. Describing the types of data distribution 3. Calculating the standard deviation 4. Calculating the variance 5. Calculating the coefficient of variation	MC/46, 47 MC/48, 49 MC/50, 51 MC/52, 53 MC/54, 55	10
Introduction to Inferential Statistics	1. Explaining the meaning of inferential statistics 2. Formulate the null hypothesis and alternative hypotheses 3. Establish criteria for testing the hypothesis	MC/56, 57 MC/58, 59 MC/60 MC/86, 87	10

	4. Determine the Level of Significance		
	5. Determine the degrees of freedom	MC/63, 66	
	6. Test the hypothesis	MC/65	
Univariate Statistical Analysis	1. Testing the normality of the data using the Liliefors test	MC/61, 62	7
	2. Testing the normality of the data using the Chi-Square test	MC/64, 67, 68	
	3. Calculating the Z-Score test	MC/69, 70	
Correlation Analysis	1. explain the meaning of correlation	MC/71, 72	10
	2. Distinguish between numbers, symbols and correlation signs	MC/73, 74	
	3. Calculate the correlation coefficient	MC/75, 78	
	4. Test the significance of the correlation	MC/76, 79	
	5. Calculate the coefficient of determination	MC/77, 80	
Regression Analysis	1. Explain the meaning of regression	MC/81, 82	9
	2. Explain the use of regression analysis	MC/83, 84	
	3. Calculating the regression equation model and its conclusions	MC/85, 88	
	4. Count the number of squared		
	5. Test the linearity of the regression and its conclusions	MC/89	
	6. Test the significance of the regression and its conclusions	MC/90	
		MC/91	
Bivariate Comparative Analysis	1. Describe comparative research	MC/92, 93	4
	2. Using the "t" test for two small samples that are not related to each other	MC/94	
	3. Using the "t" test for two small samples that are related to each other	MC/95	
Analisis of Varians (ANOVA)	1. Explaining the Analysis Of Variance	MC/96, 97	5
	2. Use of ANOVA in quantitative research		
	3. Practice Computing one-way ANOVA	MC/98, 99	
	4. Practice Computing two way ANOVA	MC/100	
Able to practice the SPSS program	1. Unable to process data with SPSS	MC/101, 102, 103, 104, 105, 106	6
	2. Able to analyse data processing with SPSS		

C. Results and Discussion

The findings of the study consist of crucial data collected through tests, questionnaires, interviews, documents, and other methods. To enhance the clarity of the results, tables, figures, or charts can be included. It is important to avoid presenting the same data in multiple separate tables. Each table, figure, and chart should be accompanied by a discussion or interpretation. All tables, figures, and charts must be centered and numbered in sequence. For qualitative research, the results section should be organized into detailed subsections that directly relate to the research focus and its categories.

IMPROVING EDUCATIONAL STATISTICS ON HOTS THROUGH RELIGIOUS MODERATION AND COGNITIVE TEST

This study used to research and development approaches utilizing the ADDIE model (*Analysis, Design, Development, Implementation, Evaluation*) to develop a test instrument based on religious moderation to measure high-level thinking skills in statistics courses in FTK UIN Banten students. This model consists of five stages, namely: 1) analysis stage, 2) design stage, 3) development stage, 4) implementation stage, and 5) evaluation stage. The development steps are as follows:

1. *Stage of Analysis (Analysis)*

In this phase, an analysis of needs and competencies is conducted. As for the needs analysis results, the researchers observed the need for statistics courses at the Tarbiyah and Teacher Training Faculty of UIN SMH Banten. The result is that the Tarbiyah and Teacher Training Faculty of UIN SMH Banten requires a question bank for statistics courses. Even with different lecturers, students get the same knowledge, evaluation, and understanding. This makes it easier for lecturers who teach statistics because of the same guidelines. Apart from that, this FTK requires more varied questions related to daily life, especially moderation in religion, so students find statistics fun. Next is the competency analysis stage, which is related to the competencies required to be achieved by students. The material chosen to develop the instrument is because it covers all the material in the statistics course. The results of the instructional analysis in the development of this question instrument are attached.

2. The design stage (Design)

After conducting the analysis, the researcher designed the instrument questions. The design of this statistical question instrument is based on religious moderation, in which each question is packaged using stories related to religious moderation. The product design stages are as follows:

a. Designing a syllabus

b. Designing the instrument grille

c. Development stage

This stage is the production stage in developing a product as a test instrument based on religious moderation from design to actual product. The manufacturing steps are as follows:

a. Make a statistics course syllabus based on the existing curriculum.

b. Then, after the syllabus is formed, the researcher makes a grid of instrument questions according to the indicators in the syllabus.

c. After that, the researcher made statistical questions based on religious moderation according to the predetermined grid.

d. Instrument questions consist of 106 questions.

3. Implementation stage (Implementation)

At this stage, the activity is to implement the test instrument. Implementation, in this case, is intended to test the validity and feasibility of the product that has been developed. Instrument validation was carried out in the 7th semester of PBA majors with a sample of 20 people.

4. Evaluation stage

The evaluation stage is to validate the products developed through product trials on semester seven students. At each stage of this development, evaluations and revisions are carried out to improve the products produced. The results are 85 valid question items with the level of difficulty and different power as follows:

Table 2: Difficulty Level of Test Instruments

Difficulty Level	Number of questions
Very easy	17
Easy	49
Currently	36
Hard	4
Very Difficult	0

[23] suggests the classification of the difficulty level of the questions, namely if the p-value (percentage of the difficulty level of the questions):
• 0.81 – 1.00 straightforward questions
• 0.61 – 0.80 easy questions
• 0.41 – 0.60 medium questions
• 0.21 – 0.40 difficult questions
• 0.00 – 0.20 items are tough questions.

The p-value is attached in the following table:

Table 3: Power Levels of Different Test Instruments

Difference Power (Biserial Point Correlation)	
Very Good Items	76
Items are Good	0
Items Need to be Revised	19
Discarded Items	11

Then after revision, there were 100 valid question items. To find out the effectiveness of the results of developing a test instrument based on religious moderation can be seen through

---

[23] Riinawati (2021)

the results of product trials before and after revision. The names of the respondents are as follows:

Table 4: Respondent Data

Resp.	Trial results in 1	Trial results 2
1	60	95
2	55	78
3	68	88
4	59	85
5	47	76
6	70	100
7	67	87
8	54	76
9	45	75
10	43	75
11	25	78
12	65	77
13	18	70
14	59	85
15	65	80
16	45	75
17	49	76
18	53	82
19	29	74
20	30	80
Total	1.006	1.612
Average	50,3	80,6

Based on the table above, the average result of trial 1 is 50.3, and the average result of trial 2 is 80.6. These data indicate that developing a test instrument based on religious moderation to measure higher-order thinking skills in statistics courses for PGMI FTK UIN Banten students is effective for learning activities.

The process of developing a test instrument based on religious moderation to measure high-order thinking skills in statistics courses for PGMI FTK UIN Banten students

[24]further explained that HOTS (Higher Order Thinking Skills) is a thinking skill that is tested at a higher level because it does not only test aspects of memory or remembering but even tests aspects of analysis, synthesis, and evaluation.

---

[24] Ibrahim et al. (2020)

ENUNG NUGRAHA, ENCEP SYARIFUDIN, AAN ANSORI, LISDA CAHYA VELIA, VANIA NURACHMA

The data above shows a positive adaptation to using higher-order thinking; according to,[25] "HOT means dealing with situations we have never encountered before". In other ways, [26] state that higher-order thinking can occur if students use their skills and knowledge in new situations. In this case, higher-order thinking means tackling situations that have never been faced before. Higher-order thinking processes occur when individuals encounter new, and not clearly defined, problems.

Successful mastery of a concept can only be obtained if students have high-order thinking skills because the concept understood will stick in the student's memory for a long time. Students from the creativity of the problems solved can also show optimal learning outcomes. Creativity is an essential component that students have in order to be able to solve a problem. This ability is essential because every human being is constantly faced with various problems that must be adequately resolved.[27]

For students, higher-order thinking skills can be familiarised through course activities. How about in-course activities in general? After the course is implemented, an assessment is carried out to determine the level of success in achieving the course objectives. According to,[28] an assessment is carried out to collect information related to the significant progress of students according to the competencies that must be mastered. Referring to the information that students' ability to perform higher-order thinking skills is still low, there is a need to develop a test instrument that can measure higher-order thinking skills.

Statistical ideas can be conveyed in various ways, such as pictures, graphs, diagrams, symbols, or words. Representing ideas in various forms requires deep understanding and skills in connecting one concept to another. Last, or sixth, is using manipulative tools.[29] This manipulative tool is needed to clarify conditions in a problem. In addition, he implicitly argues that in order to train higher-order thinking skills, problems are needed. The problems used are challenging (motivate students to participate in learning), require the cooperation of several people to solve them, and have connections with other disciplines.

---

[25] Suastra et al. (2021)

[26] Yuwono and Pasani (2018)

[27] Ana Gessa, Eyda Marin, and Pilar Sancha, "A Practical Application of Statistical Process Control to Evaluate the Performance Rate of Academic Programmes: Implications and Suggestions," *Quality Assurance in Education* 30, no. 4 (2022): 571–88, https://doi.org/10.1108/QAE-03-2022-0065.

[28] Hewi and Shaleh (2020)

[29] Achmad Badrun Kurnia, Tom Lowrie, and Sitti Maesuri Patahuddin, "The Development of High School Students' Statistical Literacy across Grade Level," *Mathematics Education Research Journal*, no. April (2023): 1–29, https://doi.org/10.1007/s13394-023-00449-x.

Thus, the process of developing statistics courses related to religious moderation will increase critical power so that by utilising higher-order thinking, students can make changes in learning and understanding path analysis which is part of the statistics learning material taught in the PGMI class of UIN SMH Banten in the fourth semester.

The effectiveness of developing a test instrument based on religious moderation to measure high-order thinking skills in statistics courses in PGMI FTK UIN Banten students

Student mistakes in working on questions can be a clue to how far students have mastered the material provided. In statistics courses, learning has only been carried out as usual teaching. In addition, when teaching, many students are still not optimal in doing the assignments given.[30] When working on questions, students tend to focus on one handbook they have.

A statistics course is one of the courses taught in universities. It is hoped that with this course, students can explore quantitative information. It is anticipated that students will be able to utilize a scientific approach, for example: in writing a thesis to solve a problem. Statistics play a role in handling quantitative data obtained from research when solving problems. Thus, statistical analysis will obtain an overview of situations, conditions or facts, providing reasonable conclusions.

Statistics is essential, but generally, students are less interested in learning it. There is a perception that studying statistics requires the mathematical ability to be one reason. Statistics and mathematics use arithmetic principles similarly.[31] The fundamental difference, among other things, is that mathematics deals with something specific, precise, exact, and precise, while statistics deals with something uncertain, emphasising reasoning and decision-making.

Statistics is vital for students because it relates to thesis writing.[32] Student readiness, thoroughness, and analytical thinking are needed to start each meeting in statistics courses. Therefore, based on the majority of students think that statistics courses are complex subjects to understand.

---

[30] Mahmud A. Mansaray and Philliph M. Mutisya, "Applied Statistical Models of Assessment and Evaluation of Instructional Efficiency," *International Journal of Academic Research in Business and Social Sciences* 7, no. 5 (2017): 223–57, https://doi.org/10.6007/ijarbss/v7-i5/2928.

[31] Gessa, Marin, and Sancha, "A Practical Application of Statistical Process Control to Evaluate the Performance Rate of Academic Programmes: Implications and Suggestions."

[32] Dwi Priyono and Ahmad Ahmad, "The Implementation of Higher Education Funding in Indonesia," *OALib* 05, no. 06 (2018): 1–11, https://doi.org/10.4236/oalib.1104049.

Sociologically, the diversity of Indonesian culture is also a consideration in practising religion by its adherents. At a time when the world community, including the Indonesian people, needs religion again, a moderate religious perspective is offered. This concerns extreme attitudes and behaviour in religion.[33] Religious moderation is essential due to conflicts in various regions in the name of religion.

Understanding religious texts tends to polarise religious adherents into two extreme poles.[34] One extreme elevates the text excessively, disregarding the role of reasoning or capability. The scriptures are followed literally without considering the context in which they were revealed. This approach is often associated with conservative groups. On the other hand, the opposite extreme, typically referred to as the liberal group, places too much emphasis on reason, often neglecting the text itself. Therefore, being overly liberal in interpreting religious teachings is just as extreme.

Islamic thought advocates moderation and tolerance of differences. It fosters openness and embraces diversity, both within sects and across religions. These differences do not hinder collaboration based on humanitarian values. While Islam is regarded as the true faith, this belief does not justify disrespecting other religions. [35] As a result, harmony and solidarity among religions will prevail, similar to what was experienced in Medina during the leadership of Prophet Muhammad SAW.

Thus, the effectiveness of using higher-order thinking makes PGMI UIN Banten students better understand path analysis material in statistics lessons more quickly. This suggests that innovations in learning used in tertiary institutions will be more quickly adapted using cognitive tests.

D. Conclusion

Developing a test instrument based on religious moderation to measure high-order thinking skills in the statistics course for FTK UIN Banten students uses The ADDIE framework, comprising five phases: analysis, design, development, implementation, and evaluation. The development steps are to make a statistics course syllabus following the existing curriculum. Then, after the syllabus was formed, the researcher made a grid of questionnaires

---

[33] Nugraha et al., *Moderasi Keislaman Dalam Bingkai Kebhinekaan*.

[34] Elif Karaosmanoglu, Didem Gamze Isiksal, and Nesenur Altinigne, "Corporate Brand Transgression and Punishing the Transgressor: Moderation of Religious Orientation," *Journal of Product & Brand Management* 34, no. 1 (2018): 1–34, https://doi.org/https://doi.org/10.1108/ JPBM-01-2017-1388.

[35] Muhammad Khairan Arif, "Moderasi Islam (Wasathiyah Islam) Perspektif Al-Qur'an, As-Sunnah Serta Pandangan Para Ulama Dan Fuqaha," *Al-Risalah* 11, no. 1 (2020): 22–43, https://doi.org/10.34005/alrisalah.v11i1.592.

based on the indicators in the syllabus. After that, the researcher made statistical questions based on religious moderation according to the predetermined grid. The effectiveness of developing a test instrument based on religious moderation to measure high-order thinking skills in statistics courses for PGMI UIN SM Banten students is effective for use in learning activities. This is based on the average results of the first and second trials. There is an increase, and it can be concluded that the test instrument is effective for use in statistics courses at PGMI UIN SMH Banten. The research implications show that lecturers at the PGMI of Tarbiyah and the Teacher Training Faculty of UIN SMH Banten can use the test instrument to evaluate material for students.

REFERENCES

Akker, Jan van den. "Principles and Methods of Development Research." *Design Approaches and Tools in Education and Training*, 1999, 1–14. https://doi.org/10.1007/978-94-011-4255-7_1.

Arif, Muhammad Khairan. "Moderasi Islam (Wasathiyah Islam) Perspektif Al-Qur'an, As-Sunnah Serta Pandangan Para Ulama Dan Fuqaha." *Al-Risalah* 11, no. 1 (2020): 22–43. https://doi.org/10.34005/alrisalah.v11i1.592.

Carter, Jarvis W., Yamir Salabarría-Peña, Errol L. Fields, and William T. Robinson. "Evaluating for Health Equity among a Cluster of Health Departments Implementing PrEP Services." *Evaluation and Program Planning* 90, no. July 2020 (2022). https://doi.org/10.1016/j.evalprogplan.2021.101981.

Dijk, Esther E. van, Jan van Tartwijk, Marieke F. van der Schaaf, and Manon Kluijtmans. "What Makes an Expert University Teacher? A Systematic Review and Synthesis of Frameworks for Teacher Expertise in Higher Education." *Educational Research Review* 31, no. 100365 (2020): 1–16. https://doi.org/10.1016/j.edurev.2020.100365.

Erfani, Hadi. *Research and Development*. Basin: LAP Lambert Academic Publishing, 2019. https://doi.org/10.12968/npre.2007.5.5.23745.

Gessa, Ana, Eyda Marin, and Pilar Sancha. "A Practical Application of Statistical Process Control to Evaluate the Performance Rate of Academic Programmes: Implications and Suggestions." *Quality Assurance in Education* 30, no. 4 (2022): 571–88. https://doi.org/10.1108/QAE-03-2022-0065.

Harjanti, Rina, Yetty Supriyati, and Wardani Rahayu. "Evaluation of Learning Programs at Elementary School Level of 'Sekolah Alam Indonesia (SAI)'. (Evaluative Research Using Countenance Stake's Model)." *American Journal of Educational Research* 7, no. 2 (2019): 125–32. https://doi.org/10.12691/education-7-2-2.

Hewi, La, and Muh Shaleh. "Refleksi Hasil PISA (The Programme For International Student Assesment): Upaya Perbaikan Bertumpu Pada Pendidikan Anak Usia Dini)." *Jurnal Golden Age* 4, no. 01 (2020): 30–41. https://doi.org/10.29408/jga.v4i01.2018.

I Wayan, Widana. "HIGHER ORDER THINKING SKILLS ASSESSMENT (HOTS)." *JISAE: Journal of Indonesian Student Assessment and Evaluation* 3, no. 1 (2017): 32–44. https://doi.org/https://doi.org/10.21009/jisae.v3i1.4859.

Ibrahim, Noorashikim Noor, Ahmad Fauzi Mohd Ayub, and Aida Suraya Md Yunus. "Impact of Higher Order Thinking Skills (HOTS) Module Based on the Cognitive Apprenticeship Model (CAM) on Student's Performance." *International Journal of Learning, Teaching and Educational Research* 19, no. 7 (2020): 246–62. https://doi.org/10.26803/IJLTER.19.7.14.

Karaosmanoglu, Elif, Didem Gamze Isiksal, and Nesenur Altinigne. "Corporate Brand Transgression and Punishing the Transgressor: Moderation of Religious Orientation." *Journal of Product & Brand Management* 34, no. 1 (2018): 1–34. https://doi.org/https://doi.org/10.1108/ JPBM-01-2017-1388.

Kurnia, Achmad Badrun, Tom Lowrie, and Sitti Maesuri Patahuddin. "The Development of High School Students' Statistical Literacy across Grade Level." *Mathematics Education Research Journal*, no. April (2023): 1–29. https://doi.org/10.1007/s13394-023-00449-x.

Mansaray, Mahmud A., and Philliph M. Mutisya. "Applied Statistical Models of Assessment and Evaluation of Instructional Efficiency." *International Journal of Academic Research in Business and Social Sciences* 7, no. 5 (2017): 223–57. https://doi.org/10.6007/ijarbss/v7-i5/2928.

Nugraha, Enung, Maruf Akbar, and S Supriyati. "Improving Learning Outcomes of Educational Research Methodology: Perceptions of the Assessment System, Learning Motivation and Emotional Intelligence among Indonesian Undergraduates." In *Proceedings of the First International Conference on Technology and Educational Science, ICSTES 2018, November 21-22 2018, Bali, Indonesia*, 1–6. Bali: EUDL, 2019. https://doi.org/10.4108/eai.21-11-2018.2282021.

Nugraha, Enung, and Muhamad Fauzi. "Digital Learning Education Development Towards Modern Islamic Culture: A Strengthening 'Merdeka Belajar' Strategy." *Al-Qalam: Jurnal Kajian Keislaman* 37, no. 2 (December 2020): 1–20. https://doi.org/http://dx.doi.org/10.32678/alqalam.v37i2.3667.

Nugraha, Enung, Novi Wulandari, Muzaki Muhammad Hafidz, Sidik Muhammad Jafar, Muhammad Imammudin, and Muhammad Taufiqul Rohmi. *Moderasi Keislaman Dalam Bingkai Kebhinekaan*. Serang: Media Madani, 2021.

Nugraheni, Meyda Mustika, Anam Sutopo, and Djalal Fuadi. "Penilaian Afektif Dalam Pembelajaran Tematik Masa Pandemi Covid-19 Di Sekolah Dasar." *ELSE (Elementary School Education Journal) : Jurnal Pendidikan Dan Pembelajaran Sekolah Dasar* 5, no. 2 (2021): 210. https://doi.org/10.30651/else.v5i2.8233.

Owston, Ron. "Models and Methods for Evaluation." *Handbook of Research on Educational Communications and Technology*, 2008, 605–17.

Parsazadeh, Nadia, Rosmah Ali, Mehran Rezaei, and Sanaz Zolfaghar Tehrani. "The Construction and Validation of a Usability Evaluation Survey for Mobile Learning Environments." *Studies in Educational Evaluation* 58, no. May (2018): 97–111. https://doi.org/10.1016/j.stueduc.2018.06.002.

Priyono, Dwi, and Ahmad Ahmad. "The Implementation of Higher Education Funding in Indonesia." *OALib* 05, no. 06 (2018): 1–11. https://doi.org/10.4236/oalib.1104049.

Riinawati. *Pengatar Evaluasi Pendidikan*. Yogyakarta: Thema Publishing, 2021.

Rolfe, Victoria, Kajsa Yang Hansen, and Rolf Strietholt. "Integrating Educational Quality and Educational Equality into a Model of Mathematics Performance." *Studies in Educational Evaluation* 74, no. May (2022): 101171. https://doi.org/10.1016/j.stueduc.2022.101171.

Saenab, Sitti, Siti Zubaidah, Susriyati Mahanal, and Sri Rahayu Lestari. "Recode to Re-Code: An Instructional Model to Accelerate Students' Critical Thinking Skills." *Education Sciences* 11, no. 1 (2021): 1–14. https://doi.org/10.3390/EDUCSCI11010002.

Srinivasan, Prakash. *Evaluation in Education*. Raleigh: Laxmi Book Publication, 2020.

STIGLER, JAMES W., JI Y. SON, KAREN B. GIVVIN, ADAM B. BLAKE, LAURA FRIES, STACY T. SHAW, and MARY C. TUCKER. "The Better Book Approach for Education Research and Development." *Teachers College Record* 123, no. 9 (2020): 1–25. https://doi.org/10.1177/016146812012200913.

Suastra, I Wayan, Ni Ketut Rapi, Putu Yasa, and I Gede Arjana. "Elaborating Indigenous Science Content into Science Learning Process: A New Science Instructional Model to Develop Students' Local Wisdom-Based Characters and Higher Order Thinking Skills." *JPI (Jurnal Pendidikan Indonesia)* 10, no. 3 (2021): 516. https://doi.org/10.23887/jpi-undiksha.v10i3.31176.

Suyadi, Zalik Nuryana, Sutrisno, and Baidi. "Academic Reform and Sustainability of Islamic Higher Education in Indonesia." *International Journal of Educational Development* 89, no. 102534 (2022): 1–11. https://doi.org/10.1016/j.ijedudev.2021.102534.

Tegeh, I Made, I Nyoman Jampel, and Ketut Pudjawan. *Model Penelitian Pengembangan*. Yogyakarta: Graha Ilmu, 2014.

Triatmojo, Eko, Wibowo -, and Yuli Agustina. "Improving The Motivation and Learning Outcomes Students with Applying 'Everyone Is Teacher Here' Learning Method." *Jurnal Pendidikan Bisnis*

Dan Manajemen 3, no. 2 (2017): 122–30. https://doi.org/10.17977/um003v3i22017p122.

Velez, Christine, Bridget Nuechterlein, Susan Connors, Grace RedShirt Tyon, Timberley M. Roane, and David C. Mays. "Application of the Indigenous Evaluation Framework to a University Certificate Program for Building Cultural Awareness in Science, Technology, Engineering, and Mathematics." *Evaluation and Program Planning* 92, no. 102066 (2022): 1–9. https://doi.org/10.1016/j.evalprogplan.2022.102066.

Vereijken, Mayke W.C., Roeland M. van der Rijst, Jan H. van Driel, and Friedo W. Dekker. "Student Learning Outcomes, Perceptions and Beliefs in the Context of Strengthening Research Integration into the First Year of Medical School." *Advances in Health Sciences Education* 23, no. 2 (2018): 371–85. https://doi.org/10.1007/s10459-017-9803-0.

Yoshikawa, Hiroyuki. *Design Methodology for Research and Development Strategy (Realizing a Sustainable Society). Japan Science and Technology Agency (JST).* Tokyo: Center for Research and Development Strategy Japan Science and Technology Agency, 2012.

Yuwono, Imam, and Chairil Faif Pasani. "The Evaluation of Higher Order Thinking Skills Assessment of Special Needs Education Students with Guided Inquiry Method." *Journal of ICSAR* 2, no. 1 (2018): 28–31. https://doi.org/10.17977/um005v2i12018p028.