**MINISTRY OF HIGHER EDUCATION,**

**SCIENCE AND INNOVATIONS OF THE REPUBLIC OF UZBEKISTAN**

**Uzbekistan World Languages University**

****

**Examination Guidelines**

**1. Regulatory Documents for Assessing Students’ Knowledge**

Final assessments in university subjects are organized on the basis of the following regulatory documents: the “Regulation on the System for Monitoring and Assessing Students’ Knowledge in Higher Educational Institutions,” approved by Order No. 19-2018 of the Minister of Higher and Secondary Specialized Education of the Republic of Uzbekistan dated August 9, 2018; the “Regulation on the Procedure for Introducing the Credit-Module System into the Educational Process of Higher Educational Institutions,” approved by Resolution No. 824 of the Cabinet of Ministers dated December 31, 2020; and the organizational measures for the arrangement and conduct of final assessments in higher education institutions, approved by Order No. 452 of the Ministry of Higher Education, Science and Innovation dated December 2, 2024.

**2.** **Test Administration Procedure**

According to the organizational measures for arranging and conducting the final assessment in higher education institutions, approved by Order No. 452 of the Ministry of Higher Education, Science, and Innovation dated December 2, 2024, when the final assessment is conducted in the form of a test, the tests may be administered in specially equipped computer classrooms under video surveillance.

The date, starting time, and duration of the test are determined in advance and announced to the students. Information about the test format (written/computer-based) and its content is provided. The questions are prepared in accordance with the course syllabus and specifications. If the test is conducted in paper form, the test papers are copied and sealed. Computer-based tests are prepared through the HEMIS information system and are administered in computer rooms under video surveillance.

**3.** **Procedure and Methodology for Developing Test Questions.**

In order to ensure the quality of test questions, number of requirements must be followed when developing them. These include alignment of test tasks with learning objectives; relevance of the material; scientific accuracy; consistency; completeness and coherence; differentiation according to the level of mastery; effectiveness (cumulativeness); linguistic clarity and accuracy; unambiguity; strict time allocation; conciseness; appropriate level of difficulty; absence of additional marks in the correct answer; variability; and logical as well as substantive coherence.

One of the main rules of test development methodology is to clarify the purpose for which the test is being created. This includes developing and testing methods for constructing test tasks, didactic requirements for test questions, preparing a test plan, creating a test plan by course sections, and clarifying the plan according to the levels of material mastery.

Not all constructed test tasks can be applied equally. In order to ensure the quality of test questions, the following requirements must be met:

Alignment of test tasks with learning objectives;

* Relevance of the material;
* Scientific accuracy;
* Consistency;
* Completeness and coherence;
* Differentiation according to the level of mastery;
* Effectiveness (cumulativeness);
* Linguistic clarity and accuracy;
* Unambiguity;
* Strict time allocation;
* Conciseness;
* Appropriate level of difficulty;
* Absence of additional marks in the correct answer;
* Variability; logical and substantive coherence.

Let us briefly consider these requirements one by one.

1. **Alignment with Learning Objectives.** The content of test tasks must correspond to clearly identified (specified) learning objectives. The set of learning objectives has a hierarchical structure. Therefore, it is not possible to achieve a general (more distant) goal without first accomplishing the more immediate objectives. Each test question should ensure the realization of a specific learning objective.

2. **Relevance of the Material.** The most important and essential parts of the educational material must be included in the test questions. The number of tasks in a test is usually limited. Therefore, it is possible that not all topics will be fully covered by the directly posed questions. The student’s ability to draw general conclusions from specific situations (deductive reasoning) and, conversely, to derive generalizations from specific facts (inductive reasoning), as well as their ability to apply theoretical rules to particular cases and to explain the underlying principles, demonstrates their mastery and understanding of the subject.

3. **Scientific Accuracy.** The test should include only information that is true and can be substantiated through reasoning and knowledge. It is not recommended to include controversial or debatable viewpoints from the field in test questions. The nature of test tasks generally requires definite answers that are known in advance..

4. **Consistency.** The tasks included in the test should pertain to a specific subject and be interconnected in terms of general knowledge content. Consistency is also reflected in the way the answer to one particular task relates to the overall test results (correlation).

5. **Completeness and Coherence.** It is important that the total number of test tasks is properly distributed in relation to the topics and sections of the subject. This requirement should be addressed during the development of the test plan..

6. **Differentiation According to the Level of Mastery.** Test tasks should be developed to reflect varying levels of mastery for specific components of the educational material (such as recall, description, application of effective knowledge, and creative thinking). This requirement should also be addressed during the development of the test plan.

Tasks designed for the recall level are based on reproducing previously learned information and rely heavily on memory.

 In reproductive-level tests, the student independently reasons based on previously acquired information, and when completing a task, relies on established rules and algorithms in the discipline. In this case, the student describes or reconstructs information from memory.

Tasks aimed at effective knowledge and creative thinking assess the individual’s ability to independently apply general methods to specific conditions described in the task. Activities related to effective knowledge and creative thinking are not based on ready-made rules and algorithms, but are performed according to newly developed or adapted rules for new conditions.

In tasks aimed at the level of creative thinking, the problem is presented in general terms, and the student identifies the necessary actions and circumstances to achieve the goal through creative thinking. In solving such tasks, entirely new information (a rule, conclusion, form, etc.) is created.

7. **Effectiveness (Cumulativeness).** Effectiveness assumes that test tasks become progressively more difficult. However, this is not an absolute requirement. In modern computer-based testing systems, each subsequent test task may be selected based on the student’s response to the previous one. That is, after an incorrect answer, the system can provide a less difficult task.

8. **Strict Time Allocation.** The essence of this requirement is that there should not be test questions that take more than 2 minutes to answer. Analysis of various international test tasks shows the following approximate time allocations: 65 seconds for mathematics, 32 seconds for language and literature, 41 seconds for social and natural sciences, 20–25 seconds for facts and statements, 90 seconds for diagrams, and 90–120 seconds for identifying shapes.

9. **Conciseness.** The wording of test tasks should not be overly lengthy or complicated. When it is necessary to assess the mastery of a complex educational issue, it is advisable to divide it into several concise test tasks.

10. **Linguistic Fluency, Accuracy, and Unambiguity.** Test tasks must not contain ambiguity or allow for multiple interpretations. In the United States, it is customary for an experienced language editor, in addition to a subject teacher and psychologist, to participate in the development of test tasks for each academic subject.

11. **Level of Difficulty (Validity).** To be pedagogically effective, test tasks should not simply present information or have answers that are obviously apparent. Test items with a result of approximately 50% (Pq0.5) are the most effective. Such a result makes it easier to distinguish between well-prepared students and those with less thorough preparation. Therefore, for important assessments such as entrance exams or final examinations, test tasks should be designed so that about 50% of students who have fully mastered the curriculum can answer them correctly. For classroom tests prepared by teachers for their own classes, a target result of 85% is considered optimal. This was explained in detail in the previous section.

12. **Absence of Additional Markers in the Correct Answer.** Inexperienced test designers may, aiming for clarity, present the correct answer in a broader or more detailed manner than the other options. Students quickly notice these differences and select the correct answer based on such cues. This should be strictly avoided.

13. **Variability.** To assess mastery of a particular topic, it is possible to create test questions that differ in logical structure, syntactic construction, morphological features, or degree of complexity, but are related to the same subject matter. This makes it possible to form sets of tasks (invariant tests) aimed at achieving the same learning objective. Computer-based testing systems are built upon this principle of variability, selecting subsequent tasks based on the answers to previous questions.

14. **Formal and Substantive Coherence.** This requirement, borrowed from philosophy and the study of works of art, states that a genuine work of art is unified both in form and content. Similarly, test tasks should also be coherent and harmonious both in form and substance. Creating test questions based on the above requirements necessitates that teachers possess theoretical knowledge of test construction. Therefore, it is important to study testology courses and independently master relevant literature.

Compliance with the above requirements for developing test tasks ensures that students’ knowledge is measured accurately and objectively. To achieve this, it is necessary to organize collaborative work among subject matter experts, teachers, language specialists, pedagogues, and psychologists when creating test tasks.

Development of the Test Plan. It is recommended to begin constructing test tasks for an academic subject with the development of a test plan. First, the total number of test tasks for the subject is determined. The relevant department makes a decision on this matter, taking into account the total number of hours allocated to studying the subject and the requirements for test development. For courses with a large number of classroom hours, it is advisable to select a greater number of test tasks in accordance with the curriculum. This makes it possible to include 1–2 questions in the test for each classroom hour. For example, for a subject with a total of 144 hours, it is possible to prepare 150–160 tasks for one test version. If the subject has fewer hours (e.g., 36 hours), the number of test tasks is determined based on considerations related to test reliability. Generally, a reliability coefficient of approximately 0.7 is achieved with 50 tasks, and to reach a reliability of 0.9, the number of tasks should be 130. Test tasks with a reliability below 0.7 are considered unsuitable for final assessment purposes. Therefore, the minimum number of test tasks for final evaluation has been set at 50.

After the test plan has been developed, requirements for the knowledge and skills that students should acquire by studying the relevant subject are formulated. These requirements are usually written in the “Goals and Objectives of the Subject” section of the subject’s syllabus. Test tasks must fully and accurately determine the level of mastery of the required knowledge and skills. For this purpose, a list of specific issues that need to be reflected in the test tasks (test specifications) is drawn up.

For example:

A – ability to define;

B – knowledge of laws, formulas, and principles;

C – ability to apply laws and formulas in problem-solving;

D – ability to identify similarities and differences;

E – ability to interpret diagrams and charts;

|F – knowledge of the structure and operating rules of devices (technical equipment);

Naturally, the list provided above is not universal and should be developed according to the specifics of each academic subject. For example, for foreign languages, elements such as sentence construction may be included; for history, significant events, notable figures, and the sequence of societal development may be covered.

A certain percentage of test tasks should be allocated to assess each type of knowledge and skill. This is a very important aspect when developing the test plan, as the validity of the test—its alignment with the curriculum—depends largely on this. To address this, test developers analyze the subject’s syllabus and reach a consensus. The accepted number of tasks is then distributed across the sections of the subject and by levels of mastery. For example, Tables 8–9 show how 120 tasks can be distributed by the four sections of a subject and by levels of mastery.

Developing such a plan ensures the validity of the test, its compliance with the content of the curriculum, and guarantees the thoroughness of students’ acquired knowledge, skills, and competencies. Once the general test plan is prepared, the selection and construction of tasks becomes the most responsible stage. As previously mentioned, this is typically carried out by the most experienced instructors who have taught the course multiple times and have already assessed students’ knowledge through testing.

Tests should be prepared in at least two variants, with one group of examinees taking Variant I (or Variant A) and another group taking Variant II (or Variant B). The number of test tasks included depends on when the test is conducted and on the volume of the material covered in the program. During the academic year, tests are usually designed for a duration of 1 hour, while tests for class transfer or final graduation are typically designed for 2 hours.The time allotted to complete each task can range from 1 to 3 minutes, with an average of about 2 minutes per task. According to this, each test variant should contain: about 25 tasks for a 1-hour test and about 50 tasks for a 2-hour test. Tasks should be numbered (1, 2, 3, … h), and their answer options should preferably be labeled alphabetically (A, B, C, D, Ye).

In test sessions for Uzbek literature, students are assessed on their knowledge of literary facts: characteristics of oral and written literature, various genres, the content of works studied at school, key features of main characters, artistic techniques and devices, and forms of creative collaboration.

Test tasks must be prepared confidentially, and the answer keys must be kept solely by the responsible person; after the exam, the answer key should be personally handed to and then retrieved from the examiner.The grading criteria for the test are as follows: Since the grading system is based on 5 points, the total number of test tasks is divided by 5, and the number of correct answers required for each grade is determined. For example, if there are 25 tasks, and each is worth the same amount, dividing by 5 yields 5. Thus, for every 5 correct answers, the student receives 1 grade point: 1–5 correct answers = grade “1”; 6–10 correct answers = grade “2”; 11–15 correct answers = grade “3”; 16–20 correct answers = grade “4”; 21–25 correct answers = grade “5”.

When developing test tasks, the following principles must be observed:

1. Scientific validity;
2. Consistency;
3. Conciseness and accuracy;
4. Non-formality;
5. Logical sequence;
6. Systematic approach;
7. Generality and specificity;
8. Simplicity or complexity;
9. Visualization;
10. Connection between theory and practice;
11. Development of thinking;
12. Creativity;
13. Independence and awareness;
14. Differentiation, i.e., taking into account the individual characteristics of students;
15. Educational value;
16. Compliance with the curriculum;
17. Objective timing;
18. Applicability.

Tests developed in accordance with these principles are of high quality and effective.

**The type of test** is selected according to the subject specifications and the purpose of assessment:

* Multiple Choice Test (Multiple Choice Questions – MCQ)
* Fill-in-the-blanks Test
* Matching Test
* Short Answer Test
* Essay-type Test

Examples of test questions in foreign languages may include the following.

#### Multiple Choice Question:

**Which of the following is a synonym of “rapid”?**
A) Slow
B) Fast ✅
C) Late
D) Lazy

#### Fill-in-the-Blanks Question:

The capital city of France is \_\_\_\_\_\_\_\_.
✅ Answer: Paris

#### Open-ended Question:

**Explain the difference between “affect” and “effect” with examples.**