

## ПЕДАГОГИЧЕСКИЕ НАУКИ

Научная статья

УДК 378.146

ББК 74.480.281

DOI: 10.54348/SciS.2024.2.1

### **Технология многостадийного оценивания в обеспечении качества итоговой аттестации в высшем образовании**

**Алексей Александрович Малыгин**

Ивановский государственный университет, Иваново, Россия, malygin@ivanovo.ac.ru, ORCID ID: 0000-0002-7812-4439

**Аннотация.** Цель статьи – представить технологию многостадийного оценивания при итоговой аттестации в высшем образовании. Эта технология призвана обеспечить соблюдение принципов объективности и независимости итоговой аттестации, которые нормативно установлены в образовательном законодательстве. Научная проблема обоснования выбора технологии оценивания при итоговой аттестации актуальна по нескольким причинам. Во-первых, компетентностная трактовка результатов освоения обучающимися образовательных программ достаточно широко варьируется при декомпозиции компетенций. Во-вторых, отсутствие общепринятых подходов к формированию и применению оценочного инструментария как для внутренней системы оценки качества образования, так и для внешних аттестационных процедур. В-третьих, реалии второго десятилетия XXI века, связанные с широким внедрением дистанционных образовательных технологий, возможностями компьютеров с увеличенными вычислительными мощностями обработки больших данных заставляют переосмысливать подходы в организации всей контрольно-оценочной деятельности как с позиций ее эффективности, так и с позиций доказательной аргументации оценки заданных стандартами результатов и принимаемых на их основе управленческих решений. Эти основания приводят к необходимости поиска релевантной технологии, позволяющей преодолеть существующие сложности и при этом обеспечить получение надежных, валидных и аутентичных результатов, снижая невербальное давление на аттестуемых и обеспечивая качество самой аттестационной процедуры. Для реализации такой аттестационной процедуры и вводится понятие *многостадийное оценивание*, которое представляет собой определенную последовательность взаимосвязанных и взаимообусловленных научно обоснованных этапов и процедур создания, предъявления и оценки результатов выполнения оценочных заданий аттестационного инструментария. Представленная в данной статье технология многостадийного оценивания служит обеспечению качества в аттестационных процедурах высокой значимости, которые имеют существенное значение для принятия различного рода управленческих решений. За счет нескольких стадий оценивания, предусматривающих применение разнообразных оценочных средств на каждом этапе сообразно выделяемому уровню сформированности компетентности (минимальному – «удовлетворительно», базовому – «хорошо», высокому – «отлично»), обеспечивается повышение надежности, валидности и аутентичности результатов аттестуемых выпускников. Реализация такой технологии основывается на бипарадигмальном подходе к образовательным измерениям и теории тестов. Разработка инструментария и его использование в многостадийном оценивании требует учета специальных положений, обеспечивающих научное обоснование качества всех компонентов процесса, средств и результатов образовательных измерений. Обработка, анализ и интерпретация результатов измерений различного уровня должна проводиться в строгом соответствии с правилами и группами допустимых операций, выделенных для каждого уровня измерения, и с учетом специфики оцениваемых переменных. Невыполнение правил обработки данных может привести к несправедливой оценке выпускников и некорректным аттестационным выводам.

**Ключевые слова:** аттестация, технология, многостадийное оценивание, образовательные изме-

рения, результаты обучения, компетенции, компетентность, высшее образование, обеспечение качества.

**Для цитирования:** Малыгин А. А. Технология многостадийного оценивания в обеспечении качества аттестационных процедур в высшем образовании // Научный поиск: личность, образование, культура. 2024. №. 2. С. 2–12. <https://doi.org/10.54348/SciS.2024.2.1>

PEDAGOGICAL SCIENCES

Original article

**Multistage assessment technology to ensure the quality of final attestation in higher education**

**Aleksei A. Malygin**

Ivanovo State University, Ivanovo, Russia, [malygin@ivanovo.ac.ru](mailto:malygin@ivanovo.ac.ru), ORCID ID: 0000-0002-7812-4439

**Abstract.** The purpose of the article is to present the technology of multistage assessment during final attestation in higher education. This technology is designed to ensure compliance with the principles of objectivity and independence of final attestation, which are normatively established in educational legislation. The scientific problem of justifying the choice of assessment technology during final attestation is relevant for several reasons. Firstly, the competency-based interpretation of students learning outcomes on educational programs varies quite widely when competencies are decomposed. Secondly, the lack of generally accepted approaches to the formation and use of assessment tools both for the internal system for assessing the quality of education and for external attestation procedures. Thirdly, the realities of the second decade of the 21st century, associated with the widespread introduction of distance technologies, the capabilities of computers with increased computing power for processing big data, force us to rethink approaches to the organization of all pedagogical control and assessment activities, both from the standpoint of its effectiveness and from the standpoint of evidence-based arguments for assessment results specified by standards and management decisions made on their basis. These reasons lead to the need to search for relevant technology that can overcome existing difficulties and at the same time ensure the receipt of reliable, valid and authentic results, reducing non-verbal pressure on those being certified and ensuring the quality of the attestation procedure itself. To implement such an attestation procedure, the concept of multistage assessment is introduced, which represents a certain sequence of interrelated and interdependent scientifically based stages and procedures for creating, presenting and estimation the results of performing assessment tasks of attestation tools. The multistage assessment technology presented in this article serves to ensure quality in attestation procedures of high significance, which are essential for making various types of management decisions. Due to several stages of assessment, involving the use of various assessment tools at each stage in accordance with the allocated level of competence development (minimum – “satisfactory”, basic – “good”, high – “excellent”), an increase in the reliability, validity and authenticity of the results of graduates is ensured. The implementation of such technology is based on a biparadigm approach to educational measurements and test theory. The development of tools and its use in multistage assessment requires taking into account special provisions that provide scientific substantiation of the quality of all components of the process, means and results of educational measurements. Processing, analysis and interpretation of measurement results at various levels must be carried out in strict accordance with the rules and groups of permissible operations allocated for each level of measurement, and taking into account the specifics of the variables being assessed. Failure to comply with data processing rules may lead to unfair assessment of graduates and incorrect attestation conclusions.

**Keywords:** attestation, technology, multistage assessment, educational measurements, learning outcomes, competencies, competency, higher education, quality assurance.

**For citation:** Malygin A.A. Multistage assessment technology to ensure the quality of final attestation in higher education. *Nauchnyj poisk: lichnost', obrazovanie, kul'tura = Scientific search: personal-ity, education, culture*. 2024. No. 2. Pp. 2–12. (In Russ). <https://doi.org/10.54348/SciS.2024.2.1>

**Актуальность.** Аттестационные процедуры сованных лиц в получаемой информации и дан- в образовании, как и в целом контрольно- ных [Болотов, 2018]. С одной стороны, это объ- оценочная деятельность, всегда вызывает боль- яняется тем, что по результатам аттестации, ший интерес не только у тех, кто проводит и как правило, стремятся сделать выводы о каче- участвует в аттестации, но и у других заинтере- стве образования. С другой стороны, аттестаци-

онные оценки часто отождествляют с серьезными, порой судьбоносными, решениями, если речь идет о завершении определенного уровня образования, присвоении квалификации, выдаче документа об образовании, приема на работу и т.д. Кроме того, часто и неправомерно, но результаты итоговой аттестации используются в социально-политических и экономических контекстах для обоснования тех или иных преобразований. В этой связи одной из задач для системы образования при организации аттестационных процедур является обеспечение их надлежащего качества, которое предлагается рассматривать с позиций обеспечения надежных, валидных и аутентичных результатов (т.е. оценок). Именно достоверные данные, получаемые в результате аттестации, при их корректном использовании выступают информационной основой для принятия различных решений. Эти решения могут варьироваться от корректировки содержания конкретных дисциплин, выбора методов и средств обучения до изменения учебных планов и образовательных программ, а также условий их реализации.

В данной работе предлагается рассмотреть аспект итоговой аттестации в высшем образовании, связанный с обоснованием технологии многостадийного оценивания. Научная проблема обоснования выбора технологии оценивания при итоговой аттестации актуальна по меньшей мере по четырем причинам. Во-первых, компетентностная трактовка результатов освоения обучающимися образовательных программ достаточно широко варьируется при декомпозиции компетенций в рабочих программах дисциплин по одинаковым или смежным направлениям подготовки, несмотря на наличие единых федеральных государственных образовательных стандартов (ФГОС) [Елина, 2015]. Во-вторых, отсутствие общепринятых подходов к формированию и применению оценочного инструментария как для внутренней системы оценки качества образования, так и для внешних аттестационных процедур [Воронова, 2015]. В-третьих, реалии второго десятилетия XXI века, связанные с широким внедрением дистанционных образовательных технологий, возможностями компьютеров с увеличенными вычислительными мощностями обработки больших данных, заставляют переосмысливать подходы в организации всей контрольно-оценочной деятельности как с позиций ее эффективности, так и с позиций доказательной аргументации оценки заданных стандартами результатов и принимаемых на их основе управленческих решений [Звонников, 2021; Малыгин, 2021]. И, наконец, четвертое обстоятельство

связано с нормативно установленными принципами объективности и независимости, которые должны быть реализованы при проведении итоговой аттестации. Указанные основания приводят к необходимости поиска релевантной технологии, позволяющей преодолеть существующие сложности и при этом обеспечивать получение надежных, валидных и аутентичных результатов, снижая невербальное давление на аттестуемых и обеспечивая качество самой аттестационной процедуры [Малыгин, 2023].

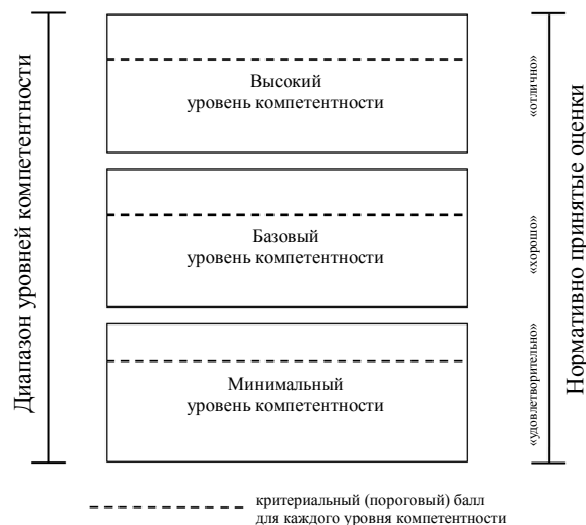
**Методы исследования.** Методологической основой исследования выступают компетентностный подход [Зимняя, 2013], бипарадигмальный подход к образовательным измерениям [Звонников, 2006], классическая и современная теория тестов [Crocker, 2006; Hambleton, 1991], область образовательных измерений [Ефремова, 2006]. К числу нормативных документов относятся федеральный закон от 29.12.2012 г. № 273-ФЗ «Об образовании в Российской Федерации» [Об образовании..., 2012], документы, регламентирующие порядок проведения итоговой аттестации по программам высшего образования, и федеральные государственные образовательные стандарты высшего образования. В число теоретических методов исследования входят: качественный и количественный анализ, классификация, сравнение, систематизация, обобщение и моделирование. Эмпирическими методами исследования выступают тестирование и математическое моделирование на основе вероятностных моделей современной теории тестов (Item Response Theory).

**Результаты исследования и их обсуждение.** На сегодняшний день процедура итоговой аттестации предусматривает две формы аттестационных испытаний – государственный экзамен и защита выпускной квалификационной работы. В данной работе речь идет о проведении государственного экзамена в форме многостадийного оценивания. Исследовательская позиция автора заключается в том, что для оценивания уровня сформированности компетентности выпускников необходимы обе формы аттестационных испытаний. При этом каждая из этих форм позволяет оценить свою совокупность компетенций. *Компетенция* рассматривается нами как способность применять знания, умения, навыки и личностные качества для успешной деятельности в различных профессиональных или жизненных ситуациях, а *компетентность* – как уровень владения выпускником совокупностью компетенций для успешного выполнения трудовых функций в определенной профессиональной области, к которой он

готовился в рамках освоения образовательной программы высшего образования [Звонников, 2012]. Действующие ФГОС высшего образования содержат перечень универсальных и общепрофессиональных компетенций, число которых в среднем превышает 16 наименований. К этому числу добавляется перечень профессиональных компетенций, определяемых вузами самостоятельно на основе профессиональных стандартов или квалификационных справочников в соответствии с выбранным видом деятельности, к которому готовится выпускник при освоении образовательной программы. Очевидно, что оценить уровень сформированности такого числа компетенций, ограничившись только защитой выпускной квалификационной работы, невозможно, как и невозможно это сделать в рамках экзамена по ответам студента на 2-3 вопроса. Поэтому каждое из аттестационных испытаний – и государственный экзамен, и выпускная квалификационная работа выполняют свои задачи, способствуя получению всесторонней и объективной информации о результатах освоения выпускником образовательной программы и уровне его компетентности. Конечно, компетентностный подход привнес много инноваций в привычный образовательный процесс вуза, но практически никак не сказался на оценочной деятельности.

В силу того, что компетенции являются способностями, то их не принято оценивать в виде определенного числа баллов на шкале [Чельшкова, 2012]. Требуется выделение диапазонов компетентности, внутри которых и будут находиться результаты аттестуемых при

выполнении ими оценочных заданий. При выборе диапазонов следует руководствоваться не только дихотомическими правилами (компетентен / некомпетентен), но также учитывать сложившиеся в отечественном образовании практики и нормативные требования. Поэтому предлагается рассматривать три диапазона уровня компетентности, выделяя – минимальный, базовый и высокий уровни. Такая градация согласуется с установленными в нормативных документах по итоговой аттестации пятибалльной шкалой оценивания [Об утверждении Порядка..., 2015]. В реальности эта шкала является четырехбалльной, поскольку предусматривает выставление по результатам аттестационных испытаний выставление одной из четырех оценок – «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно». В последнем случае аттестуемый выпускник считается не прошедшим итоговую аттестацию. Таким образом, предлагаемые три уровня компетентности выпускника (минимальный, базовый и высокий) будут соотноситься с принятой градацией оценок («удовлетворительно», «хорошо», «отлично»). Для этого целесообразно предложить использовать вертикальную шкалу, которая легко интерпретируется, задается в понятных для аттестуемых студентов и выпускников категориях, исключая абстрактные оценочные прилагательные (например, превосходный, хороший, плохой и др.), и дифференцированно соотносится с традиционной системой оценок, принятой в отечественной высшей школе (рисунок 1).



**Рисунок 1.** Вертикальная шкала для диапазона уровней компетентности, соотнесенная с нормативно принятыми оценками

**Figure 1.** Vertical scale for a range of competency levels, correlated with normatively accepted assessments

Продвижение по этой шкале от минимального к высокому уровню компетентности предполагает выполнение аттестуемым оценочных заданий соответствующей трудности: чем выше уровень компетентности, тем труднее совокупность заданий. Для реализации такой аттестационной процедуры и вводится понятие *многостадийное оценивание*, которое представляет собой определенную последовательность взаимосвязанных и взаимообусловленных научно обоснованных этапов и процедур создания, предъявления и оценки результатов выполнения оценочных заданий аттестационного инструментария [Мальгин, 2023]. Выбор многостадийного оценивания для аттестационного испытания обусловлен стремлением принимать ответственные и обоснованные решения в оценочных процедурах высокой значимости. За счет сочетания различных свойств и возможностей оценочных средств и многоэтапности самой процедуры повышается надежность и валидность результатов, появляется возможность оценить различные способы деятельности выпускников и сделать аргументированные выводы об уровне их компетентности. Такое оценивание, проводимое в несколько этапов (стадий), отвечает требованиям компетентностного подхода, а также профессиональных стандартов, в которых трудовые функции специалистов описаны через необходимые знания, умения и трудовые действия.

Понимая под технологией ведущую форму организации деятельности и систему условий, форм, методов, средств и критериев решения поставленной задачи [Новиков, 2014], опишем реализацию многостадийного оценивания. Исходя из уровневости оценивания компетентности, аттестационные процедуры должны предусматривать использование различного оценочного инструментария и соответствующих шкал, на которые размещается результат измерений и корректно интерпретируется. Для этого при разработке и использовании аттестационных процедур необходимы положения бипарадигмального подхода к образовательным измерениям и теории тестов, определяющие допустимость и корректность использования того или иного оценочного инструментария.

Критериально-ориентированный подход, имеющий своей целью интерпретировать результаты аттестуемых по отношению к содержательной области, включенной в измеритель, и разделить их на аттестованных и не аттестованных согласно заранее установленному критериальному баллу (стандарту выполнения каждого этапа многостадийного оценивания), также задает правила отбора и включения в оце-

ночный инструментарий заданий определенной трудности. При критериально-ориентированном подходе инструментарий должен с максимальной полнотой отразить все то, что можно принять за полный объем планируемых к освоению результатов образовательной программы [Porham, 1978]. Доля правильно выполненных заданий рассматривается как уровень освоения общего объема (то, что можно принять за 100 %). При итоговой аттестации встает вопрос, что следует принять за полный объем. Строго говоря, это должен быть весь перечень компетенций (универсальных, общепрофессиональных и профессиональных), число которых часто превосходит 30 наименований, но оценить их в рамках одного аттестационного испытания невозможно. К этому выводу приводит не только здравый смысл, но и практические попытки многих университетских коллективов на начальном этапе реализации стандартов третьего поколения и оценивания компетенций. Позднее появились подходы, связанные с кластеризацией и паспортизацией компетенций, которые предусматривали укрупнение их по смыслу и по содержанию, что положительным образом сказалось на понимании процессов оценивания компетенций [Ефремова, 2010; Звонников, 2012]. С другой стороны, в действующих нормативных документах, регламентирующих порядок проведения итоговой аттестации по программам высшего образования, предусматривается, что государственный экзамен проводится по одной или нескольким дисциплинам (модулям) образовательной программы, результаты освоения которых имеют *определяющее значение для профессиональной деятельности* выпускников [Об утверждении Порядка..., 2015]. Таким образом, на уровне регулятора закладывается норма, согласующаяся со стремлением получить в рамках этого аттестационного испытания информацию об уровне компетентности выпускника и при этом соотносить знания, умения и опыт практической деятельности с требованиями, заложенными во ФГОС и профессиональных стандартах. В качестве результатов, имеющих определяющее значение для профессиональной деятельности, в первую очередь отдается предпочтение профессиональным компетенциям и тем дисциплинам и модулям, которые их формируют. Поэтому за содержательную область, подлежащую операционализации и конструированию для нее измерителей, выбираются соответствующие дисциплины учебного плана образовательной программы, которые иногда называют «ядерными» или базовыми. Содержание таких измерителей должно быть достаточно полным, в него включают

все то, что условно можно принять за 100% свой критериальный или пороговый балл, пре- объем, планируемый к освоению выпускниками образовательной программы. Основную часть заданий измерителей для аттестации делают значительно легче средней трудности, особенно при планировании процента аттестуемых, кото- рые не пройдут за пороговый балл. Если про- цент не аттестованных не должен быть больше 10 % и пороговый балл установлен на уровне 70 % (т.е. выпускники, которые выполнили меньше 70 % заданий становятся не аттестова- ными), то в измеритель включают не менее 70 % легких заданий, с которыми справится 90 % выпускников [Звонников, 2012].

Другими словами, каждый этап многоста- дийного оценивания носит критериально- ориентированный характер и должен иметь

свой критериальный или пороговый балл, пре- одоление которого свидетельствует о соответ- ствии выпускника определенному уровню ком- петентности.

Дифференциация методов и инструментария для итоговой аттестации может быть проведена по различным основаниям. В случае нашего исследования, в котором сделан выбор в пользу бипарадигмального подхода к образовательным измерениям в условиях компетентностной трак- товки результатов освоения выпускниками об- разовательных программ, такая дифференциа- ция проводится сообразно принятым уровням измерений. Первая группа методов и средств обеспечивает количественный уровень измере- ний, вторая – качественный (рисунок 2).



**Рисунок 2.** Методы и инструментарий для многостадийного оценивания в соответствии с уровнями образовательных измерений

**Figure 2.** Methods and tools for multistage assessment in accordance with educational measurement levels

В группу, соответствующую количественно- му уровню измерений и таким же шкалам, относятся стандартизированные оценочные зада- ния, проверка правильности выполнения кото- рых может быть автоматизирована. Традицион- но это задания закрытой формы (с выбором од- ного или нескольких правильных ответов, на установление соответствия или правильной по- следовательности, ситуационные задания с вы- сокой степенью детерминированности контек- ста), носящие репродуктивный и частично по- исковый характер. В принятой нами терминологии это задания минимального и базового уров- ней компетентности, т. е. при условии их ус-

пешного выполнения соответствующие оцен- кам «удовлетворительно» и «хорошо». При проведении обследований качества образова- ния, условий реализации образовательных про- грамм и удовлетворенности обучающихся, про- фессорско-преподавательского состава и рабо- тодателей целесообразно дополнять оценки вы- пускников по итоговой аттестации результа- тами стандартизированного анкетирования. Этот метод также относится к количественному уровню, а сами анкеты как инструментарий тре- буют при их разработке проведения апробации, доказательства надежности и валидности. При стремлении получения максимально разносто-

ронней информации о результатах реализации образовательных программ в целях повышения качества образования и подготовки специалистов сочетание методов тестирования и анкетирования, конечно, будет приносить положительный эффект.

Вторая группа качественного уровня измерений включает такие методы и инструментарий, реализация которых осуществляется с участием экспертов. В первую очередь, речь идет об оценке результатов выполнения заданий качественного уровня измерений. В условиях итоговой аттестации такую оценку необходимо проводить не одному человеку, а группе экспертов. В этом и есть одно из предназначений создания экзаменационных комиссий для итоговой аттестации, когда классификационные решения и общая оценка уровня подготовленности студента и компетентности выпускника принимаются групповым методом, коллегиально. К качественному уровню следует отнести задания практических умений, кейс-измерители, ситуационные задания с низкой степенью детерминированности контекста, структурированное интервью, собеседование по профессиональной проблеме, портфолио.

Вне зависимости от уровней измерения можно сформулировать ряд требований к аттестационным процедурам:

– методы и инструментарий должны обеспечивать информацией, на каком уровне и в каком объеме освоена образовательная программа и компетенции;

– оценочные задания должны быть релевантны измеряемым образовательным результатам и быть согласованы с требованиями образовательных и профессиональных стандартов;

– по полученным результатам можно корректно зафиксировать прирост и динамику освоения образовательной программы на протяжении всего периода обучения;

– механизмы применения методов и средств оценивания не должны приводить к неоднозначным оценкам и нарушениям принципов многостадийного оценивания.

Определенный для каждого уровня измерений инструментарий требует некоторых замечаний. Так, необходимо уточнить понятие «тест» в рамках критериально-ориентированного подхода. Под тестом в современном понимании рассматривают совокупность различных форм заданий (не только с выбором одного или нескольких правильных ответов), обладающих системообразующими свойствами и устойчивыми характеристиками (прежде всего, трудность и дифференцирующая способность) [Звонников, 2006; Linden, 2016]. Заявленный критери-

ально-ориентированный подход требует уточнений термина «тест». Поэтому под *тестом* на количественном уровне образовательных измерений для многостадийного оценивания понимается система таких заданий, которые предъявляются в определенной стратегии и по алгоритмам, обеспечивающим надежность и содержательную валидность результатов по отношению к установленным стандартам выполнения конкретного этапа оценивания. Такой тест в процедурах итоговой аттестации является междисциплинарным, поскольку для измерения заявляется сразу несколько компетенций. Часть профессиональных компетенций, как правило, соотносится с трудовыми функциями, при условии, если результаты освоения образовательной программы проектировались с учетом требований профессиональных стандартов. Аттестационный тест, который в многостадийном оценивании предъявляется аттестуемому на первом этапе для измерения уровня компетентности выпускника на минимальном уровне (оценка «удовлетворительно»), в силу гетерогенности состоит из тестлетов – однородных по содержанию субтестов, служащих измерению отдельного конструкта. Количество таких тестлетов будет соответствовать числу выделенных базовых дисциплин (модулей) учебного плана образовательной программы, результаты освоения которых имеют определяющее значение для профессиональной деятельности выпускников [Малыгин, 2018]. Высокий уровень значимости аттестационных испытаний, в которых используется такой инструментарий, требует серьезного подхода к разработке оценочных заданий, их обязательной апробации желательно на репрезентативных выборках, что не всегда становится возможным в силу разных обстоятельств. Такая апробация должна завершаться коррекцией оценочных заданий, повышая содержательную валидность инструментария и приводя его к профессиональному уровню.

Вторая категория оценочных заданий, относящихся к количественному уровню измерений и предусматривающих в ходе их выполнения не просто воспроизводство конкретных знаний и демонстрацию умений (репродуктивный уровень), а решение задач проблемно-поискового характера – это *ситуационные или профессионально-ориентированные задания*. Этот инструментарий используется на втором этапе многостадийного оценивания, и уровень его трудности соответствует базовому уровню компетентности выпускника. Следует отметить, что ситуационные задания являются частным случаем кейсов, более широкой категории оценочных средств, но их отличие состоит в степени детер-

минированности контекста [Семенова, 2019]. По сравнению с кейсами, допускающими изменение контекста со стороны аттестуемых, ситуационные задания строго стандартизированы, если речь идет об автоматизированной проверке, но также носят междисциплинарный характер. Как и в случае разработки теста, перед созданием ситуационного задания необходимо определить число и характер конструкторов (компетенций), подлежащих измерению, сформировать проблемный контекст, который имеет аутентичный характер и актуальный для того вида профессиональной деятельности, к которой готовился выпускник при освоении образовательной программы. Основная часть задания содержит текст с ситуацией и контекстом из реальной профессиональной деятельности. В общем виде ситуация представляет собой совокупность событий, связанных в целое одной проблемой. Эти события объединены причинно-следственными связями, способствующими или препятствующими разрешению противоречия. Ситуация содержит явные и латентные противоречия и характеризуется контекстом, под которым понимаются граничные условия, описывающие количественно или качественно особенности действий, компонентов и участников ситуации. Профессионально-ориентированное задание предполагает ответ на несколько вопросов, которые направлены на оценку измеряемых профессиональных компетенций. Оценочные рубрики, т. е. правила по которым следует проверять правильность выполнения задания, должны содержать одинаковое число градаций в одном измерителе. Как правило, таких градаций вводят не более четырех. Экспертизу ситуационных заданий желательно проводить специалисту, который не участвовал в его разработке. Этот оценочный инструмент также должен проходить апробацию на представительной выборке, чтобы повысить качество структуры и содержания ситуационных заданий.

После того как аттестуемый подтверждает компетентность на базовом уровне, что соответствует оценке «хорошо», он может выходить на следующий этап оценивания – высокого уровня компетентности (оценка «отлично»). Этот уровень предусматривает структурированное интервью или свободное собеседование на заданную тему (поставленную проблему). Кроме того, целесообразно добавлять к этим качественным методам оценивания также портфолио, выступающее одновременно и методом, и средством оценивания.

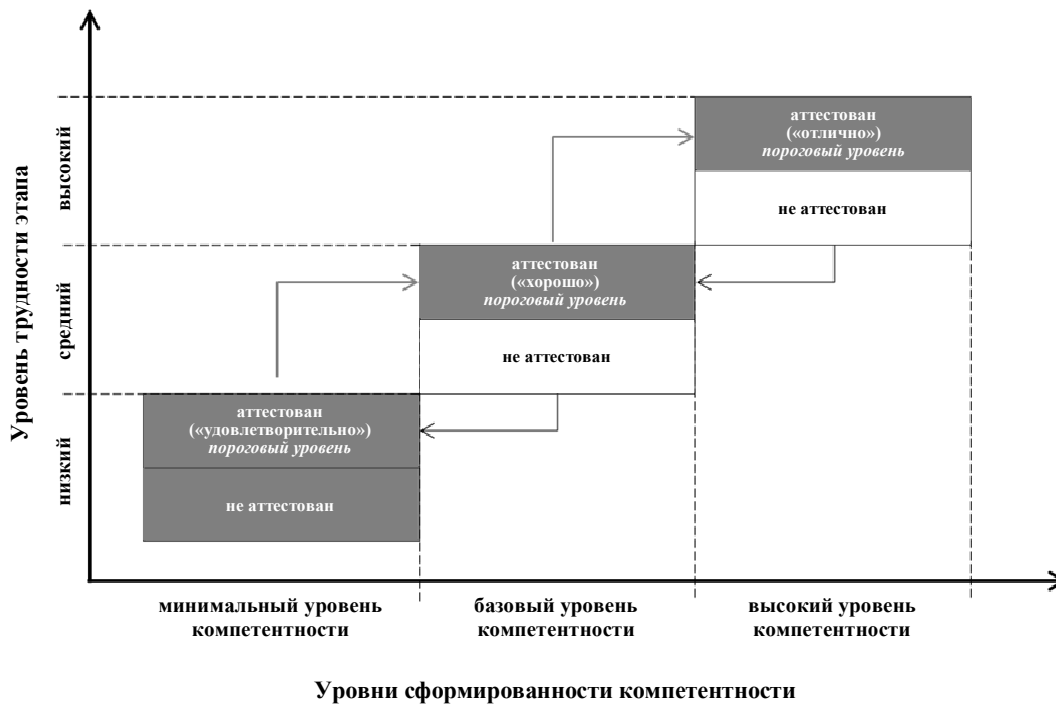
Интервью следует отнести к научно обоснованному методу сбора данных, которое прово-

дится на качественном уровне. Приписывание каких-либо чисел ответам аттестуемых не делает этот метод количественным и не предусматривает использование метрической шкалы. Интервью классифицируют на неструктурированные (открытые) и структурированные (закрытые). В силу специфики аттестационных испытаний, ограниченности временных ресурсов, возможности открытых диалогов с выпускниками на защите выпускной квалификационной работы, на аттестационной процедуре в форме государственного экзамена, строящегося на многостадийном оценивании, предпочтение следует отдать структурированному интервью. Такой метод предполагает последовательность вопросов – от общего к частному, сводящих к предмету обсуждения. Для представления данных интервьюирования применяют неметрические шкалы (например, Лайкерта). Конечно, проведение интервью является трудозатратным и может оказаться малоэффективным, повышая субъективность при принятии классификационных решений и выставления итоговой оценки выпускнику. Кроме того, сами эксперты, выполняющие функции экзаменаторов, должны быть готовы к проведению такого метода оценивания. В этой связи целесообразнее рассматривать структурированное собеседование с выпускником, который прошел два этапа оценивания и подтвердил соответствие уровню базовой компетентности. Такое собеседование проводится по определенной профессиональной проблеме, которую может предложить как сам аттестуемый выпускник, так и экзаменационная комиссия. Дополнять ответы аттестуемого призвано портфолио, которое формируется самим обучающимся на протяжении всего периода обучения. Метод портфолио и оценка его достаточно проработаны как в научном, так и методическом плане [Воронова, 2019; Звонников, 2006; Казакова, 2018]. Его следует рассматривать как дополнительную аргументацию при подготовке выводов об уровне компетентности выпускника на высоком уровне и выставлении оценки «отлично». Принято считать, что портфолио (или можно встретить название «паспорт профессиональных достижений») включает такие работы, которые раскрывают достижения его владельца, свидетельства самостоятельной работы. В связи с большими трудозатратами на анализ и оценку материалов портфолио, выпускники могут представить его заблаговременно в аттестационную комиссию. Такой метод аутентичного оценивания благоприятным образом сказывается на аттестационной процедуре высокой значимости, снижая психологическое напряжение аттестуемых.



На рисунке 3 представлен последовательный дизайн многостадийного оценивания. Первый этап содержит задания репродуктивного характера для оценки знаний и умений, необходимых выпускнику для выполнения трудовых действий, который реализуется с помощью компьютерного тестирования на основе алгоритмов современной теории тестов с критериально-ориентированным подходом к интерпретации результатов [Dorozhkin, 2016; Lord, 1980]. В случае успешного преодоления критериального балла аттестуемый будет отнесен к минимальному уровню компетентности (оценка «удовлетворительно»), после чего он может приступить к выполнению заданий второго этапа оценивания. На этой стадии выпускнику предлагается выполнить квазипрофессиональные ситуацион-

ные задачи проблемно-поискового характера. При успешном выполнении необходимого количества заданий (также устанавливается критериальное значение (балл) для второго этапа) выпускник будет отнесен к категории аттестованных и соответствовать базовому уровню сформированности компетентности (оценка «хорошо»). Третья стадия оценивания соотносится с высоким уровнем сформированности компетентности выпускника и предусматривает свободное собеседование на заданную тему с членами государственной экзаменационной комиссии. При успешном прохождении такой формы оценивания, которая может быть дополнена, как описано выше, портфолио из работ выпускника, выставляется оценка «отлично».



**Рисунок 3.** Последовательный дизайн многостадийного оценивания  
**Figure 3.** Sequential design of multistage assessment

**Заключение.** Представленная в данной статье технология многостадийного оценивания служит обеспечению качества в аттестационных процедурах высокой значимости, которые имеют существенное значение для принятия различного рода управленческих решений. Именно за счет нескольких стадий оценивания, предусматривающих применение разнообразных оценочных средств на каждом этапе сообразно выделяемому уровню сформированности компетентности (минимальному – «удовлетворительно», базовому – «хорошо», высокому – «отлично»), обеспечивается повышение

надежности, валидности и аутентичности результатов аттестуемых выпускников. Реализация такой технологии основывается на бипарадигмальном подходе к образовательным измерениям и теории тестов. Разработка инструментария и его использование в многостадийном оценивании требует учета специальных положений, обеспечивающих научное обоснование качества всех компонентов процесса, средств и результатов образовательных измерений. В обоснование качества сложность вносит латентный характер подлежащих измерению компетенций, поэтому интерпретация результатов

измерений при итоговой аттестации без общения с выпускником, без предыстории его обучения может привести к неверным выводам относительно уровня его компетентности. В этой связи необходимо сочетать количественный и качественный уровни измерений. Обработка, анализ и интерпретация результатов измерений различного уровня должна проводиться в строгом соответствии с правилами и группами допустимых операций, выделенных для каждого уровня измерения, и с учетом специфики оцениваемых переменных. Невыполнение правил обработки данных может привести к несправедливой оценке выпускников и некорректным аттестационным выводам.

#### Список источников

- Болотов В. А. Прошлое, настоящее и возможное будущее российской системы оценки качества образования // Вопросы образования. 2018. № 3. С. 287–297.
- Воронова Т. А., Малыгин А. А. Контрольно-оценочная система в вузе: традиции и требования современности // Научный поиск. 2015. № 3.1. С. 32–34.
- Воронова Т. А., Малыгин А. А. Актуальные направления в оценивании готовности выпускников к профессиональной педагогической деятельности // Высшее образование сегодня. 2019. № 12. С. 10–16. <https://doi.org/10.25586/RNU.HET.19.12.P.10>
- Елина Е. Г., Ковтун Е. Н., Родионова С. Е. Компетенции и результаты обучения: логика представления в образовательных программах // Высшее образование в России. 2015. № 1. С. 10–20.
- Ефремова Н. Ф., Звонников В. И., Чельшкова М. Б. Педагогические измерения в системе образования // Педагогика. 2006. № 2. С. 14–22.
- Ефремова Н. Ф. Подходы к оцениванию компетенций в высшем образовании: учеб. пособие. Москва: Исслед. центр пробл. качества подгот. специалистов, 2010. 216 с.
- Звонников В. И. Измерения и качество образования. Москва: Логос, 2006. 312 с.
- Звонников В. И., Малыгин А. А., Чельшкова М. Б. О доказательном подходе и его видах в образовании // Вестник Ивановского государственного университета. Серия «Естественные, общественные науки». 2021. Вып. 2. С. 46–52.
- Звонников В. И., Чельшкова М. Б. Оценка качества результатов обучения при аттестации: (компетентностный подход): учеб. пособие. 2-е изд., перераб., доп. Москва: Логос, 2012. 280 с.
- Зимняя И. А. Компетенция и компетентность в контексте компетентностного подхода в образовании // Ученые записки национального общества прикладной лингвистики. 2013. № 4 (4). С. 16–31.
- Казакова Е. И., Тарханова И. Ю. Оценка универсальных компетенций студентов при освоении образовательных программ // Ярославский педагогический вестник. 2018. № 5. С. 127–135.
- Малыгин А. А. Оценивание при аттестации студентов в современном высшем образовании: цели и подходы // Мир университетской науки: культура, образование. 2023. № 9. С. 80–94. <https://doi.org/10.18522/2658-6983-2023-9-80-94>.
- Малыгин А. А. Современные форматы образовательного тестирования // Высшее образование сегодня. 2018. № 6. С. 15–18. <https://doi.org/10.25586/RNU.HET.18.06.P.15>.
- Малыгин А. А., Соловьева Е. А., Травина Ю. Ю. Современные практики оценивания результатов обучения в высшем образовании // Научный поиск: личность, образование, культура. 2021. № 4 (42). С. 3–13.
- Новиков А. М., Новиков Д. А. Методология. Москва: КРАСАНД, 2014. 632 с.
- Об образовании в Российской Федерации. Федер. закон № 273-ФЗ от 29 декабря 2012 г. : с изменениями 2023 г. (включает все изменения до 1 января 2024 г.). URL: <http://consultant.ru/> (Дата обращения: 15.02.2024).
- Об утверждении Порядка проведения государственной итоговой аттестации по образовательным программам высшего образования – программам бакалавриата, программам специалитета и программам магистратуры : приказ Минобрнауки России № 636 от 29 июня 2015 г. URL: <http://consultant.ru/> (Дата обращения: 15.02.2024).
- Семенова Т. В., Сизова Ж. М., Малахова Т. Н., Малыгин А. А., Чельшкова М. Б., Князева С. А. Интерактивные множественные кейсы в аккредитации специалистов здравоохранения // Медицинский вестник Северного Кавказа. 2019. № 1. С. 118–122. <https://doi.org/10.14300/mnnc.2019.14066>
- Чельшкова М. Б. Аттестация выпускников вузов в рамках компетентностного подхода // Вестник Костромского государственного университета им. Н. А. Некрасова. 2012. Т. 18, № 6. С. 270–273.
- Crocker L., Algina J. Introduction to Classical and Modern Test Theory. Pacific Grove (CA): Wadsworth, 2006. 527 p.
- Dorzhkin E.M., Chelyshkova M.B., Malygin A.A., Toymentseva I.A., Anopchenko T.Y. Innovative Approaches to Increasing the Student Assessment Procedures Effectiveness // International Journal of Environmental and Science Education. 2016. Vol. 11, № 14. P. 7129-7144.
- Hambleton R. K., Swaminathan H., Rogers H. J. Fundamentals of Item Response Theory. New York: Sage Publications, 1991. 174 p.
- Linden W. J. van der. Handbook of Item Response Theory : Models / W. J. van der Linden. New York : CRC Press, 2016. 624 p.
- Lord F. M. Application of Item Response Theory to Practical Testing Problems. Hillsdale (NJ): Lawrence Erlbaum Ass., Publ., 1980. 274 p.

Popham W. J. Criterion-referenced measurement. Englewood Cliffs. New York: Prentice Hall, 1978.

## References

- Bolotov V. A. Past, present and possible future of the Russian system for assessing the quality of education. *Voprosy obrazovaniya = Educational Studies*. 2018. No. 3. Pp. 287–297. (In Russ).
- Voronova T. A., Malygin A. A. Control and assessment system at a university: traditions and modern requirements. *Nauchnyy poisk = Scientific search*. 2015. No. 3.1. Pp. 32–34. (In Russ).
- Voronova T. A., Malygin A. A. Current directions in assessing the readiness of graduates for professional teaching activities. *Vysshye obrazovaniye segodnya = Higher education today*. 2019. No. 12. Pp. 10–16. (In Russ) <https://doi.org/10.25586/RNU.HET.19.12.P.10>
- Elina E. G., Kovtun E. N., Rodionova S. E. Competencies and learning outcomes: logic of presentation in educational programs. *Vysshye obrazovaniye v Rossii = Higher education in Russia*. 2015. No. 1. Pp. 10–20. (In Russ).
- Efremova N.F., Zvonnikov V.I., Chelyshkova M.B. Pedagogical measurements in the education system. *Pedagogika = Pedagogy*. 2006. No. 2. Pp. 14–22. (In Russ).
- Efremova N. F. Approaches to assessing competencies in higher education: textbook. allowance. Moscow: Research. problem center quality of preparation specialists, 2010. 216 p. (In Russ).
- Zvonnikov V.I. Measurements and quality of education. Moscow: Logos, 2006. 312 p. (In Russ).
- Zvonnikov V.I., Malygin A.A., Chelyshkova M.B. To the Evidence-Based Approach and its Types in Education. *Vestnik Ivanovskogo gosudarstvennogo universiteta. Ser.: Estestvennyye, obshchestvennyye nauki = Bulletin of Ivanovo State University. Series: Natural and Social Sciences*. 2021. Issue 2. Pp. 46–52. (In Russ).
- Zvonnikov V. I., Chelyshkova M. B. Assessing the quality of learning outcomes during attestation (competency-based approach). Moscow: Logos Publ., 2012. 280 p. (In Russ).
- Zimnyaya I. A. Competence and competency in the context of the competency-based approach in education. *Uchenyye zapiski natsional'nogo obshchestva prikladnoy lingvistiki = Scientific notes of the National Society of Applied Linguistics*. 2013. No. 4 (4). Pp. 16–31. (In Russ).
- Kazakova E.I., Tarkhanova I.Yu. Assessment of universal competencies of students when mastering educational programs. *Yaroslavskiy pedagogicheskiy vestnik = Yaroslavl Pedagogical Bulletin*. 2018. No. 5. Pp. 127–135. (In Russ).
- Malygin A. A. Assessment during certification of students in modern higher education: goals and approaches. *Mir universitetskoy nauki: kul'tura, obrazovaniye = World of university science: culture, education*. 2023. No. 9. Pp. 80–94. (In Russ). <https://doi.org/10.18522/2658-6983-2023-9-80-94>.
- Malygin A. A. Modern formats of educational testing. *Vysshye obrazovaniye segodnya = Higher education today*. 2018. No. 6. Pp. 15–18. (In Russ). <https://doi.org/10.25586/RNU.HET.18.06.P.15>.
- Malygin A. A., Solovyova E. A., Travina Yu. Yu. Modern practices of assessing learning outcomes in higher education. *Nauchnyy poisk: lichnost', obrazovanie, kul'tura = Scientific search: personality, education, culture*. 2021. No. 4 (42). Pp. 3–13. (In Russ).
- Novikov A. M., Novikov D. A. Methodology. Moscow: KRASAND, 2014. 632 p. (In Russ).
- On education in the Russian Federation. Federation Law No. 273-FZ of December 29, 2012: as amended in 2023 (includes all changes until January 1, 2024). URL: <http://consultant.ru/> (Date of application: 15.02.2024) (In Russ).
- On approval of the Procedure for conducting state final certification for educational programs of higher education – bachelor's programs, specialty programs and master's programs: order of the Ministry of Education and Science of Russia No. 636 dated June 29, 2015. URL: <http://consultant.ru/> (Date of application: 15.02.2024) (In Russ).
- Semenova T.V., Sizova Zh.M., Malakhova T.N., Malygin A.A., Chelyshkova M.B., Knyazeva S.A. Interactive multiple cases in the accreditation of healthcare professionals. *Meditinskiiy vestnik Severnogo Kavkaza = Medical Bulletin of the North Caucasus*. 2019. No. 1. Pp. 118–122. (In Russ). <https://doi.org/10.14300/mnnc.2019.14066>. (In Russ).
- Chelyshkova M. B. Attestation of University Graduates within the Framework of a Competency-based Approach. *Vestnik Kostromskogo gosudarstvennogo universiteta im. N.A. Nekrasova = Bulletin of Kostroma State University named after N.A. Nekrasov*. 2012. Vol. 18, No. 6. Pp. 270–273. (In Russ).
- Crocker L., Algina J. Introduction to Classical and Modern Test Theory. Pacific Grove (CA): Wadsworth, 2006. 527 p. (In Eng).
- Dorozhkin E.M., Chelyshkova M.B., Malygin A.A., Toymentseva I.A., Anopchenko T.Y. Innovative Approaches to Increasing the Student Assessment Procedures Effectiveness. *International Journal of Environmental and Science Education*. 2016. Vol. 11, № 14. Pp. 7129–7144. (In Eng).
- Hambleton R. K., Swaminathan H., Rogers H. J. Fundamentals of Item Response Theory. New York: Sage Publications, 1991. 174 p. (In Eng).
- Linden W. J. van der. Handbook of Item Response Theory : Models / W. J. van der Linden. New York : CRC Press, 2016. 624 p. (In Eng).
- Lord F. M. Application of Item Response Theory to Practical Testing Problems. Hillsdale (NJ): Lawrence Erlbaum Ass., Publ., 1980. 274 p. (In Eng).
- Popham W. J. Criterion-referenced measurement. Englewood Cliffs. New York: Prentice Hall, 1978. (In Eng).
- Статья поступила в редакцию 07.05.2024; одобрена после рецензирования 14.05.2024; принята к публикации 20.05.2024.
- The article was submitted 07.05.2024; approved after reviewing 14.05.2024; accepted for publication 20.05.2024.