

2023, № 08 (август)

Раздел 5.8. Педагогика

ART 231068

DOI: 10.24412/2304-120X-2023-11068

УДК 378.146

**Доказательно аргументированный подход  
к оцениванию компетенций  
в профессиональном образовании**

**Evidence-centered approach to competences assessment  
in vocational education**

**Автор статьи**

**Ефремова Надежда Федоровна,**  
доктор педагогических наук, заведующая кафедрой  
педагогических измерений ФГБОУ ВО «Донской  
государственный технический университет», г. Ро-  
стов-на-Дону, Российская Федерация  
nefremova61@donstu.ru  
ORCID: 0000-0003-4556-008X

**Author of the article**

**Nadezhda F. Efremova,**  
Doctor of Education, Professor, Head of the Pedagogical  
Measurements Department, Don State Engineering Uni-  
versity, Rostov-on-Don, Russian Federation  
nefremova61@donstu.ru  
ORCID: 0000-0003-4556-008X

**Конфликт интересов**

Конфликт интересов не указан

**Conflict of interest statement**

Conflict of interest is not declared

**Для цитирования**

Ефремова Н. Ф. Доказательно аргументированный  
подход к оцениванию компетенций в профессиональ-  
ном образовании // Научно-методический электрон-  
ный журнал «Концепт». – 2023. – № 08. – С. 1–14. –  
URL: <https://e-koncept.ru/2023/231068.htm>. DOI:  
10.24412/2304-120X-2023-11068

**For citation**

N. F. Efremova, Evidence-centered approach to compe-  
tences assessment in vocational education // Scientific-  
methodological electronic journal "Koncept". – 2023. –  
No. 08. – P. 1–14. – URL: <https://e-koncept.ru/2023/231068.htm>. DOI: 10.24412/2304-120X-2023-11068

Поступила в редакцию <i>Received</i>	11.06.23	Получена положительная рецензия <i>Received a positive review</i>	19.07.23
Принята к публикации <i>Accepted for publication</i>	19.07.23	Опубликована <i>Published</i>	31.08.23



**Аннотация**

В статье на основе анализа научных работ отечественных и зарубежных авторов представлено текущее состояние теории и практики оценки компетенций студентов. Отмечены трудности оценивания сложных, глубоко латентных и междисциплинарных по своей природе образовательных результатов обучающихся в терминах компетенций при освоении основных образовательных программ профессионального образования. Актуальность исследования обусловлена потребностью повышения надежности оценок. Цель исследования – показать новые подходы к проектированию, конструированию и применению оценочных средств по моделям доказательной аргументации с опорой на теорию и практику педагогических измерений, обеспечивающих справедливость оценочного процесса и аутентичность результатов оценки. Новизна исследования заключается в системном представлении алгоритма проектирования и моделирования доказательно аргументированной оценочной процедуры с включением планируемой деятельности студентов при выполнении задания. Расширены представления о методологии доказательно аргументированной оценки компетенций в целях повышения надежности и аутентичности результатов. Показано, что тенденции развития контрольно-оценочной деятельности направлены на укрепление связи между обучением и дизайном оценки, ориентированной на фактические данные. Важным этапом нового подхода к оцениванию компетенций является проектирование инструмента оценки на основе моделирования элементов оценки и использования ситуативных задач. При этом моделирование оценочного конструкта предметной области и планируемой деятельности испытуемых в процессе их взаимодействия с оценочным средством является системообразующим элементом инструмента оценки, деятельность обучающихся в ситуации оценочной процедуры обеспечивается использованием симуляций в цифровой среде. Анализ и самооценка результатов становятся обязательными составляющими в оценочном процессе.

**Abstract**

Based on the analysis of domestic and foreign scientific works, the article presents the current state of the theory and practice of assessing students' competences. It is noted that there are difficulties in evaluating complex, deeply latent and interdisciplinary by their nature educational results of students in terms of competences in the development of basic educational programs of vocational education. The relevance of the study is due to the need to improve the reliability of estimates. The aim of the study is to show new approaches to the design, construction and application of assessment tools according to evidence-centered argumentation models based on the theory and practice of pedagogical measurements that ensure the fairness of the assessment process and the authenticity of the assessment results. The novelty of the study lies in the systemic representation of the design algorithm and simulation of an evidence-centered assessment procedure with the inclusion of the planned work of students when completing the task. The ideas about the methodology of evidence-centered assessment of competences have been expanded in order to increase the reliability and authenticity of the results. It is shown that the trends in the development of control and assessment activities are aimed at strengthening the connection between training and the design of evidence-centered assessment. An important step in the new approach to assessing competences is the design of an assessment tool based on the modeling of assessment elements and the use of situational tasks. At the same time, modeling the assessment construct of the subject area and the planned activity of the subjects in the process of their interaction with the evaluation means is a system-forming element of the assessment tool, the activity of students in the situation of the assessment procedure is ensured by the use of simulations in the digital environment. Analysis and self-assessment of the results become an obligatory component in the assessment process.

**Ключевые слова**

оценка компетенции, дизайн оценочных средств, доказательство, аргументации, моделирование, цифровая среда

**Key words**

assessment of competence, design of assessment tools, evidence, reasoning, modeling, digital environment

**Благодарности**

Автор выражает благодарность О. Д. Федотовой, заведующей кафедрой образования и педагогических наук ФГБОУ ВО «Донской государственный технический университет», за помощь в организации и проведении апробации многокомпонентного оценочного средства.

**Acknowledgements**

The author is grateful to O. D. Fedotova, Head of the Department of Education and Pedagogical Sciences, Don State Engineering University, for her help in organizing and testing the multicomponent assessment tool.

**Введение / Introduction**

Чем сложнее становится мир, тем больше потребность в специалистах, владеющих навыками решения проблем, умением общаться, сотрудничать и творить, и они будут становиться все более востребованными в высокотехнологичной среде. Основная особенность этих навыков заключается в том, что они позволяют индивиду, используя информацию, решать проблемные ситуации, выполнять различные виды деятельности и эффективно функционировать в производственной сфере. Эти навыки, так называемые мягкие, включают творчество, критическое мышление и решение проблем, навыки сотрудничества, способности работать с информационными и цифровыми

технологиями, а также социальную, культурную и метакогнитивную осведомленность. Их сложная структура и междисциплинарный характер определяют проблемы аутентичного оценивания образовательных достижений. Включение «мягких навыков» и компетенций как важных результатов освоения основных образовательных программ ставит задачу их надежного оценивания. Проблема заключается в том, следует ли оценивать такие результаты и как делать это правильно. Запрос на аутентичное оценивание результатов обучения в высших учебных заведениях возрастает во всем мире, в том числе и в системе отечественного образования. Ключевым достижением в этом направлении стало признание оценки важной составляющей образовательной деятельности и способом обеспечения гарантии качества обучения.

В выборе методов и технологий оценки возникает ряд вопросов. Как зависит освоение компетенций от способов организации и проведения учебного процесса? Можно ли проявление компетенций в ситуации оценки обеспечить планируемыми действиями испытуемых при решении сложных задач? Какие конкретные, ситуативные и многокомпонентные задания позволяют испытуемым проявить сложные и глубоко латентные результаты обучения? Можно ли считать достоверными оценки сформированности компетенций, полученные методами тестирования? Насколько сложными должны быть планируемые действия испытуемых при оценивании, чтобы одним педагогическим инструментарием выявлять несколько аспектов знаний, компетенций и навыков? Каким должен быть дизайн оценочных задач, чтобы они вызывали интерес обучающихся и не только мотивировали на решение конкретной задачи, но и повышали стремление углублять познания и развивать способы деятельности? Как разрабатывать комплексные задания с широким спектром новых идей в изучаемой области, используя достижения науки и технологий в области педагогических измерений для создания и применения различных форм оценивания? На эти вопросы частично есть ответы, но на многие другие еще предстоит найти в рамках теории и практики педагогических измерений.

Целью этого исследования является систематизация научной информации о методах доказательной аргументации и обеспечения достоверности результатов оценивания компетенций в образовании, а также представление опыта конструирования и использования многокомпонентного задания, предусматривающего этап самооценки студентами своих компетенций в цифровой среде.

### Обзор литературы / Literature review

Внедрение компетентностного подхода в образовательный процесс привело к новым результатам освоения основных образовательных программ в терминах компетенций как характеристик обучающихся, отражающих способности деятельности на основе применения усвоенных знаний. А соответственно, актуализируется проблема разработки и применения оценочных средств, способных такую деятельность задавать и обеспечивать надежность ее оценивания. Традиционные формы оценивания с помощью тестовых заданий с выбором правильного ответа или множественным выбором долгое время служили определенным социальным целям, контексту и группам населения. Их основная цель в основном сводится к оценке когнитивных результатов освоения содержания отдельных дисциплин. Достижения в области теории и практики педагогических измерений привели к появлению стандартизированных процедур тестирования и лучшему получению оценок когнитивных результатов обучения. Таким образом, классическая теория конструирования и применения тестов с

заданиями ответов на вопросы соответствует модели когнитивной диагностики, но она не может обеспечить оценку компетенций. Следует отметить, что именно такие технологии позволили лучше понимать задачи и проблемы оценки результатов освоения образовательных программ, улучшить дизайн оценок для удовлетворения различных потребностей образовательной деятельности и запросов работодателей, на основе опыта разработать новые методы оценки трудно поддающихся измерению компетенций и навыков XXI века. Эти проблемы сегодня решаются за счет использования достижений педагогики, психологии, теории и практики педагогических измерений и цифровых технологий.

Заявленные результаты освоения образовательных программ в виде компетенций как никогда важны в цифровую эпоху, а это диктует необходимость изменений в оценивании, переосмысления роли и значения аутентичной оценки в обеспечении и гарантии качества обучения. Поскольку компетенции по определению это способности действовать с использованием полученных при обучении знаний и умений и принимать решения в проблемных ситуациях, то и в оценочном процессе необходимо планировать и оценивать деятельность обучающихся. Именно совокупность данных о поведенческих характеристиках испытуемых позволяет делать выводы о сформированности компетенций. Концептуализацию методов оценивания компетенций и навыков в масштабируемых пространствах задает проектирование дизайна оценки по моделям доказательной логики и аргументации.

По мнению Р. Мислеви, образовательная оценка в настоящее время переживает своего рода ренессанс [1]. За последние десятилетия внимание к оценке результатов обучения усилилось во всем мире. В сфере получения образовательных результатов стало возможным широкое использование формативного оценивания на основе электронных учебников и онлайн-ресурсов, веб-портфолио и анализа учебной аналитики с целью получения системной качественной и количественной оценки и отслеживания прогресса обучения на основе анализа объективных данных об уровнях подготовленности обучающихся. Такой подход к оценке на основе объективной аналитики образовательных достижений на каждом значимом этапе обучения способствует своевременной корректировке и повышению эффективности образовательной деятельности. А развитие теории педагогических измерений и практики конструирования оценочного инструментария, а также технологические возможности обеспечивают выполнение задач аутентичного оценивания образовательных достижений и определения психометрических характеристик оценочных средств, тем самым способствуя повышению качества оценочного инструментария.

Согласно С. Тулмину, при оценивании образовательных достижений для доказательства результата необходимо выделять оценочные аргументы на основе эпистемологической и эмпирической логики [2]. Сущность достоверной оценки, по мнению С. Мессика, следует искать в полноте представления конструкта оценочного средства в рамках обоснованного дизайна, когда оценивание понимается как построение аргумента, включающего в себя некое свойство испытуемого, доказательства надежности, достоверности и справедливости оценки и рациональное использование критериев и оценочных рубрик [3].

Сегодня цифровые технологии бросают вызов области традиционных образовательных измерений, позволяя в значительной степени автоматизировать оценочный процесс, сделать его независимым, более достоверным и увлекательным. В последние десятилетия информационное образовательное пространство становится одним из

важных инструментов не только обучения, но и самоконтроля, независимая оценка компетенций позволяет обучающимся определиться с образовательной траекторией и выбором программ обучения. В работах П. Гриффина и Э. Кэр инновационные подходы к оценке рассматриваются с точки зрения возможностей использования технологических платформ, обеспечивающих тесные связи между оценкой и обучением [4]. На основе широкого использования масштабируемости среды онлайн-обучения повышается достоверность образовательных результатов, создаются условия для самооценки и взаимной оценки, организации обратной связи.

Метод Evidence-Centered Design (ECD) проектирования дизайна оценки, ориентированный на получение фактических данных, предложен группой ученых (Р. Мислеви, Р. Алмонд и Дж. Лукас) и уже более двух десятилетий широко используется в зарубежной образовательной практике [5]. Методология Evidence-Centered Design концептуально основана на проектировании и конструировании оценочных средств, требующих от испытуемых выполнения определенной деятельности с использованием знаний контролируемого содержания изучаемой области. Идея создания инструмента оценки компетенций, основанная на методе доказательной аргументации ECD, предоставляет возможности из наблюдений за поведением испытуемого получать свидетельства о сформированности компетенции. Дизайн оценки строится на принципах доказательств и опровержений, а также системе аргументаций наблюдаемого поведения испытуемых при решении многокомпонентного задания. Для его реализации Р. Мислеви, Л. Стейнберг и Р. Алмонд предложили использование шаблонов проектирования оценочных средств на основе профессиональных ситуаций, наблюдений и оценки решения заданий, чтобы охватить более широкий набор целевых навыков испытуемых в ситуации их взаимодействия с инструментом оценки [6]. Принципиальная структура проектирования дизайна оценочного средства, согласно Ю. Цой и Р. Мислеви, опирается на согласованность глубокого анализа содержания предметной области, его интеграцию с образовательной программой и учебным планом и обеспечивает наблюдение прогресса обучения в системе онлайн-оценки [7]. Организация обучения и формирования компетенций согласуется с репрезентативным отображением контролируемого содержания при решении задания и выполнении планируемой деятельности обучающегося, заложенных в отдельных измерительных моделях оценочного средства. Оценивание выстраивается на основе теоретических предположений о природе конструкта и эмпирических свидетельствах, наблюдаемых в процессе действий испытуемых. По мнению М. Оливери и Р. Мислеви, проблемы и возможности в разработке оценки следующего поколения навыков XXI века становятся все более актуальными, требующими новых подходов [8].

В качестве одного из таких подходов Р. Шавельсон, О. Златкин-Троицкая, К. Бек, С. Шмидт и Дж. Марино предлагают оценку действий испытуемых при выполнении «критических» заданий с использованием реальных ситуаций, при решении которых студенты учатся проявлять свои компетенции и критические навыки как в академических, так и в профессиональных областях [9]. При этом М-Т. Нагель, О. Златкин-Троицкая и Дж. Фишер обращают внимание на важность валидации задач для оценки критического онлайн-рассуждения студентов и выпускников университетов [10]. Основное внимание в этих работах направлено на укрепление связи между оценкой и обучением, независимую оценку «мягких навыков» (soft skills) и компетенций, таких как критическое мышление, навыки коммуникации, креативность.

В этой области только в последние годы появились работы отечественных ученых и практиков. Сделаны попытки адаптировать методологию Evidence-Centered



Design к оцениванию критического мышления школьников и информационно-коммуникационной компетентности учащихся. Расширяются возможности использования комплексных компетентностных заданий. В исследовании Н. В. Родиной, А. А. Постоваловой и А. В. Долматова акцент сделан на независимом характере оценки компетенций в цифровом мире с использованием метода Evidence-Centered Design [11], показано, что индикаторы достижений связаны общим контекстом и контентом. Контекст определяет проблемную ситуацию задания и ее развитие, последовательность действий испытуемых по этапам выполнения задания. Контент задает содержательную основу предметной области в заданиях ситуационного типа в рамках метода доказательной аргументации. В заданиях такого типа могут быть использованы симуляторы поисковой системы.

Проблемы разработки содержательного конструкта и качественного инструмента измерения критического мышления взрослых отражены в трудах А. Р. Садовой, Ю. С. Хиль, Т. В. Пащенко и К. В. Тарасовой [12]. Качество заданий подтверждается психометрическим анализом и оценкой размерности инструмента измерения на основании данных статистики согласия.

Одним из преимуществ методологии ECD К. В. Тарасова и Е. А. Орел считают высокую конструктивную валидность, обеспеченную особенностями дизайна проектирования оценочного инструмента, а фиксируемое поведение студентов в процессе выполнения оценочного задания создает условия для сбора свидетельств сформированности критического мышления [13]. Это позволяет проводить оценивание трудно наблюдаемых латентных навыков испытуемых с обеспечением психометрического качества оценки. Таким образом, ситуативные задания позволяют испытуемым проявить, а оценщикам собрать необходимые доказательства для обоснования вывода о компетенции.

Важным преимуществом методологии ECD И. Л. Угланова, И. В. Брун и Г. М. Васин считают целостность и внутреннюю согласованность всей разработки оценочного инструмента, структуру и содержание заданий, правила начисления баллов и условия взаимодействия испытуемых с инструментом оценки [14]. Для оценки критического мышления и организации обратной связи многокомпонентное задание ими разрабатывалось с построением психологического конструкта на базе диаграмм С. Тулмина.

Подходы к разработке педагогического инструментария для оценивания информационно-коммуникационной компетенции с высокой валидностью и надежностью представлены в работе С. М. Авдеевой, М. Г. Руднева, Г. М. Васиной, К. В. Тарасовой и Д. М. Пановой [15].

На необходимость переосмысления оценки не только обучающихся, но и университетов в цифровом пространстве обращают внимание М. Бирман и Ф. Доусон с соавторами [16]. Они выделяют три новых направления оценки в цифровом мире: новая ориентация на будущее; меняющийся характер оценочной работы преподавателей в цифровом мире; новые способы мышления о масштабируемости оценок результатов обучения в цифровом пространстве.

В концептуальной модели оценки компетенций студентов Н. Ф. Ефремовой показаны основные этапы проектировочной деятельности с использованием паттерн-дизайна при разработке аргументированной оценки компетенций студентов [17]. Приемы доказательной аргументации оценки компетенций, по мнению Н. Ф. Ефремовой, обеспечивая надежность и достоверность оценок, повышают доверие испытуемых к результатам и, как следствие, мотивацию учебной деятельности [18] за счет обеспечения справедливости, заложенной в дизайне оценивания компетенций студентов [19].

Особое отношение ECD может иметь, по мнению Д. Алмонд и Дж. Бланкен-Уэбб, к преобразующему потенциалу вузов [20]. Технологические инновации меняют способы обучения и оценивания, а дизайн, ориентированный на фактические данные, может оказаться полезным для переосмысления университетской оценки в цифровом мире. Поэтому возможности современной независимой оценки компетентностных и когнитивных результатов обучения студентов в цифровом образовательном пространстве вузов могут рассматриваться как ключевой способ преподавания и обучения будущих специалистов, адаптированных к цифровому миру. Определение и оценка необходимых навыков в производственных условиях также начинают опираться на доказательную базу, обеспечивая условия для трансформации и реализации стратегии непрерывного совершенствования. Ф. Ачерби, М. Росси и С. Терци считают, что пока лишь немногие исследования включали эффективные инструменты оценки навыков и компетенций рабочей силы для повышения и закрепления конкурентного преимущества компаний [21].

Таким образом, краткий обзор работ в этом направлении показывает, что при планировании и разработке оценки по методологии ECD решаются две основные задачи. Первая – подбор свидетельств в оценочном процессе, при помощи которых можно наблюдать за действиями испытуемых для аргументированного заключения о достижении компетентностных результатов обучения. Вторая связана с подбором профессионально важных ситуаций, которые могут быть заложены в оценочное средство, чтобы испытуемый при выполнении задания мог проявить свои компетенции. Связь между этапами разработки оценки с обязательным планированием деятельности испытуемых при выполнении заданий может способствовать созданию нового типа оценочных средств и дальнейшему развитию аутентичного оценивания образовательных достижений в условиях компетентностного обучения студентов. Как правило, это достигается построением сложных конструкторов, в которых для целей проверки используются обособленные атрибуты как когнитивных, так и компетентностных результатов обучения, в том числе с использованием симуляций и внешних цифровых платформ.

### **Методологическая база исследования / Methodological base of the research**

В теоретическом плане проведен анализ работ отечественных и зарубежных авторов по проблемам оценивания образовательных достижений, существующих инновационных подходов к выявлению компетенций студентов профессионального образования, проведены сбор, обобщение и структурирование научной информации о существующих методах и технологиях проектирования и конструирования оценочного инструментария с использованием моделирования элементов оценочного средства. Основная проблема оценивания компетенций связана с необходимостью проектирования и разработки оценочных средств, способных выявлять и обеспечивать наблюдение за проявлением глубоко латентных характеристик испытуемых. Методология доказательно аргументированного дизайна оценки, первоначально представленная в трудах Р. Мислеви [22], а также Р. Мислеви и Г. Хартель [23], хорошо зарекомендовала себя в мировой образовательной практике при оценке когнитивных результатов обучения. А затем основные положения этой методологии стали использоваться и для оценивания компетентностных достижений обучающихся. Дизайн ECD, ориентированный на доказательства, оказался мощной основой для разработки достоверных и надежных задач и оценок навыков XXI века, отмечают В. Шуте, Ю. Ким и Р. Раззук [24]. Как аналитический и творческий процесс такая оценка обеспечивает

студентов и преподавателей возможностями экспериментировать, создавать и прототипировать, собирать отзывы, анализировать результаты.

В практической части в качестве примера на основе опыта организации и проведения различных форм оценки результатов компетентностного обучения разработано и апробировано многокомпонентное задание для независимой оценки цифровой грамотности студентов с использованием внешней цифровой платформы. Предусмотрены условия для самооценки испытуемыми их способности работать в цифровом образовательном пространстве, а также уровня владения информационной и коммуникативной компетенциями в цифровой среде.

### Результаты исследования / Research results

Особенностью настоящего времени, как отмечают Н. Е. Сердитова и А. В. Белоцерковский, являются масштабы и скорость изменений в основе технологической жизни и деятельности, что диктует особые требования к результатам образования [25]. Все большее распространение получают технологии обучения в цифровой среде, требующие отслеживания цифрового следа обучающихся для получения данных о взаимодействии обучающихся с цифровыми технологиями обучения и оценки для решения образовательных проблем. По мнению М. Чукуровой, М. Маврикиса и Р. Лукина, методология оценки по моделям ECD может стать хорошей основой для формирования больших данных, обеспечивая субъектов образования информацией о взаимодействии обучающихся с системой оценки, отслеживания и выявления процессов и результатов обучения в цифровой среде [26]. Использование ECD, как показано в статье Э. Сноу, Д. Рутштейна, С. Басу, М. Биенковски Х. Эверсона, представляет собой принципиально новый подход к измерению критического мышления и обеспечению связи между учебными планами и дизайном оценивания [27]. Это способствует планированию образовательной деятельности и оценки, основанной на фактических данных, разработке и администрированию оценки, обеспечивает прочную основу для аргументации ее достоверности. Но, как справедливо отмечают С. Бечард, А. Кларк и Р. Ромин, такие измерения могут потребовать много времени, ресурсов и сложности при реализации [28]. Однако эти затраты окупятся за счет создания и применения многократно используемых шаблонов для однотипных методов разработки оценочных средств.

Обоснованный дизайн оценочного средства по компетенциям в своей структуре содержит несколько моделей: конструкт с набором аргументов и характеристик испытуемых, которые предполагается оценить; перечень данных для доказательно аргументированного заключения; структуру ситуативных заданий, позволяющих выявить интересующие характеристики в соответствии с целью оценочного процесса. По методологии Evidence-Centered Design в структуре оценочного средства имеются три основные модели: модель студента, модель доказательств и модель задания. Вначале создается модель студента, которая содержит конструкт с оцениваемыми атрибутами. Затем в соответствии с конструктом выстраивается модель доказательств, отражающая планируемое поведение или действия испытуемых в ситуации оценки, виды аргументации и критерии сформированности компетенций. И только после этого разрабатывается модель задания (шаблон проектирования), описывающая структуру и содержание ситуационных задач, способных вызвать планируемые действия и адекватное поведение испытуемых, доказывающие наличие и уровни сформированности компетенции. Оценка компетенций предусматривает наблюдение за действиями испытуемых в конкретных ситуациях при



решении задания, сбор данных, использование доказательств, аргументации и опровержения наблюдаемых результатов, интерпретацию действий испытуемых, а также обоснование вывода о способности человека действовать и быть готовым находить решения в этих и других подобных ситуациях. Для этого задание конструируется так, чтобы испытуемый мог проявить компетенции, а проверяющий – их наблюдать и оценивать, обеспечивая равенство испытаний для всех обучающихся и справедливость оценки. По мнению Д. Руссо, методология оценки поведенческих компетенций опирается на несколько возможностей, таких как моделирование, психометрический тест, упражнения для принятия критических решений и другие, должна проводиться как непрерывный процесс постоянного наращивания знаний и навыков [29].

Средства оценивания компетенций становятся все более широкомасштабными, инновационными и высокотехнологичными. Обращение к перспективным концепциям разработки дизайна оценки в системе образования способствует овладению методами построения многовариантных форм контроля и установления связей объектов в системах обучения и контроля достижений. Сбор, анализ и интерпретация данных обеспечивают доказательную основу оценки (см. рисунок).



Основные этапы проектирования оценочного средства

Проектирование многокомпонентного задания по такой методологии состоит из последовательности нескольких этапов:

- анализ контролируемого содержания предметной области (понятия, законы и принципы, основные положения, методы и технологии, концепции и др.) и обоснование конструкта;
- выбор формы и метода проведения оценочного процесса (устный, бланочный, компьютерный и др.), определение средств и технологий предъявления заданий, сбора ответов и обработки результатов;
- определение структуры конструкта из одного или нескольких аргументов и предполагаемого в ситуации оценки поведения испытуемых, свидетельствующего о достижении планируемых результатов обучения;
- разработка показателей и критериев образовательных достижений;
- подбор проблемных ситуаций, в которых необходимые для их разрешения действия обеспечат проявление испытуемыми запланированных к оцениванию компетенций и важных для данной предметной области знаний и умений;
- конструирование многокомпонентного задания, релевантного цели и задаче оценочного процесса;
- создание уровневой шкалы для отображения результатов и обоснование пороговых значений положительных оценок по уровням освоения компетенций.

В качестве примера такого оценочного средства в данном исследовании предложено многокомпонентное задание по оценке дефицитов цифровой грамотности будущих педагогов (магистрантов). Разработка концептуальной основы оценки начинается с формулирования цели. Задание состоит из ряда элементов: контекста для введения испытуемого в проблему и обеспечения мотивации выполнения задания; аргумента оценки, задающего наблюдаемое поведение испытуемого и обеспечивающего свидетельство сформированности оцениваемого конструкта; задач, которые должен выполнить испытуемый; рефлексивной части, обеспечивающей самоанализ выполненных действий и самооценки полученных результатов. В данном случае это самоанализ и самооценка дефицитов цифровой грамотности как квалификационной характеристики педагога. В задании использованы элементы аргументации и независимой оценки на внешней платформе в цифровой среде.

Вводная часть (контекст): главной предпосылкой успешного развития российской цифровой экономики является подготовка подрастающего поколения к жизни и деятельности в цифровом обществе. Одним из путей решения этой задачи выступают согласованные действия и постоянные коммуникации педагогов и учащихся в цифровой среде. Важными компетенциями учителей для решения этой задачи являются информационные и коммуникационные, которые обеспечивают прогрессивные уровни образовательной деятельности, в том числе с использованием цифровой среды. Именно они подлежат самооценке при выполнении данного задания.

Конструкт задания состоит из оцениваемых атрибутов:

- проявление цифровой грамотности при выполнении задания на внешней цифровой платформе;
- владение способами получения, хранения и передачи информации, в том числе с использованием интерактивной онлайн-доски Miro;
- использование педагогического общения и коммуникации в электронной среде;
- рефлексивный анализ и самооценка, краткое резюме об отношении к работе.

Выполнение задания предполагает четыре этапа, задающих деятельность испытуемых (краткая инструкция).

1. Пройдите тест «ИКТ компетентность педагогических работников» на платформе <https://fgosonline.ru/test-po-ikt-kompetentnosti-pedagogicheskikh-rabotnikov/>. Тест состоит из 15 заданий, в каждом из которых необходимо выбрать правильный вариант ответа. Выполнение заданий оценивается дихотомически: 0 – задание не выполнено; 1 – выполнено.

2. Получите результат выполнения теста на платформе и впишите его в таблицу.

3. Оцените свои дефициты цифровой грамотности, значимость и уровень владения информационной и коммуникативной компетенциями по пятибалльной шкале (значимость компетенций выделить/подчеркнуть, а уровень владения выразить в числовом эквиваленте в табл. 1).

4. Опишите результаты рефлексивного анализа работы с заданием и разместите их на онлайн-доске Miro. Приглашение на доску предоставляется по Ф. И. О. и имейлу участника.

В этом задании предусмотрена работа испытуемых на внешней многопользовательской цифровой платформе, обеспечившей внешнюю независимую оценку цифровой грамотности, способности испытуемых получать и передавать информацию в онлайн формате. После работы с ситуативной частью задания предоставлена возмож-

ность самооценки и выражения отношения к заданию и значению цифровых технологий. В апробации задания на добровольной основе приняли участие магистранты педагогического направления подготовки (табл. 2).

Таблица 1

### Самооценка и оценка важности развития цифровых компетенций

Компетенции	Значимость	Уровень владения компетенцией (самооценка)		
		Низкий (0-1)	Средний (2-3)	Высокий (4-5)
Информационной	Высокая (4-5)			
	Средняя (2-3)			
	Низкая (0-1)			
Коммуникационной	Высокая (4-5)			
	Средняя (2-3)			
	Низкая (0-1)			
Средний результат самооценки				
Балл за выполнение теста				

Таблица 2

### Результаты тестирования и самооценки

№ участника	Балл за тест	Самооценка компетенций		Выдержки из краткого резюме участников
		информационной	коммуникативной	
1	12	4	5	Работа с платформой вызвала незначительные трудности
2	15	5	5	Задания для меня были очень интересными
3	14	5	5	Рада, что познакомилась с полезной интерактивной платформой и Miro
4	12	4	5	Задания были познавательными, но вызвали затруднения
5	12	4	5	Задания были интересные и познавательные, работа с платформой вызвала затруднения
6	14	5	5	Работать с платформой было весьма интересно, но имелись некоторые трудности
7	14	5	5	Задания не вызвали трудностей, были интересными
8	15	5	5	Задания не вызвали трудностей, были интересными и достаточно простыми. Работа с Miro вызвала небольшие затруднения, но была интересна

Участники апробации отметили достижения и трудности при взаимодействии с заданием, продемонстрировали заинтересованное отношение к заданию и работе с онлайн-доской Miro, способности передавать информацию и выразили готовность общаться между собой в цифровой среде, показали достаточно высокий уровень владения цифровой грамотностью, информационной и коммуникативной компетенциями. Как отмечают Дж. Кансал и Н. Джейн, анализ результатов является лучшим инструментом для выявления индивидуальных потребностей в обучении и планирования дальнейшего развития [30].

### Заключение / Conclusion

Показано, что пересмотр технологии оценки компетенций студентов может значительно усилить надежность результатов, повысить интерес и доверие к оценке со стороны испытуемых, чтобы она была аутентичной и пригодной для повышения учебной мотивации и прогнозирования успешности молодых специалистов. Особое

внимание в таком оценочном процессе уделяется роли вероятностных рассуждений о выполнении заданий при анализе накопленных данных, характеризующих знания, навыки и/или способности обучающихся, а также возможности обеспечения масштабируемости как адаптации условий проведения оценочного процесса к измененным размерам выборки испытуемых с сохранением ее репрезентативности и использованием стандартных уровневых шкал. При проектировании компетентностной оценки по моделям ECD с точки зрения сбора доказательственных аргументов о деятельности испытуемых в процессе взаимодействия с оценочным средством обеспечивается связь между требованиями к освоению содержания предметной области и дизайном оценки, ориентированным на фактические данные.

Создание заданий по моделям ECD требует от разработчиков довольно много усилий, чтобы получать необходимые аргументации и доказательства владения компетенцией. Следует учитывать, что сложный конструкт снижает доказательную базу по каждому атрибуту оценки, а упрощенный может не обеспечить выполнение задач оценки, хорошо сбалансированные оценочные рубрики обеспечивают интерес к заданию со стороны испытуемых и дают необходимую информацию преподавателю по результатам выполнения задания. Структура и уровень детализации задания становятся эффективными не только для оценки, но и для поддержки процесса обучения, что оправдывает затраты на разработку. Таким образом, важно найти идеальный баланс между объемом контролируемой области и структурированным сбором данных о когнитивных и компетентностных результатах обучения и агрегировать все виды данных (качественных и количественных) для анализа состояния объектов оценки.

Все это представляет собой довольно сложную задачу, ее решение требует от разработчиков соответствующей подготовки и дополнительного времени для того, чтобы для конструирования заданий подбирать актуальные ситуации, в которых выпускники будут выполнять задачи профессиональной деятельности.

### Ссылки на источники / References

1. Mislevy R., Haertel G., Riconscente M., et. al. Assessing Model-Based Reasoning using Evidence-Centered Design // A Suite of Research-Based Design Patterns. – Springer Briefs in Statistics, 2017. – 137 p.
2. Toulmin S. The Uses of Argument. Updated Edition. – Cambridge University Press, 1958, 2003. – 259 p.
3. Messick S. The interplay of evidence and consequences in the validation of performance assessments // Educational Researcher. – 1994. – No 23(2). – P. 13–23.
4. Griffin P., Care E. Assessment and Teaching of 21st Century Skills. Educational Assessment in an Information Age // Methods and Approach. – Springer Science+Business Media Dordrecht, 2015. – 314 p.
5. Mislevy R. J., Almond R. G., Lukas J. F. A brief introduction to evidence centered design // ETS Research Report Series. – 2003. – 29. DOI: 10.1002/j.2333-8504.2003.tb01908.x
6. Mislevy R. J., Steinberg L. S., Almond R. A. On the structure of educational assessments // Measurement Interdisciplinary Research and Perspectives. – 2003. – No 1. – P. 3–67.
7. Choi Y., Mislevy R. Evidence centered design framework and dynamic bayesian network for modeling learning progression in online assessment system // Frontiers Psychology. – 2022. – 13:742956. – P. 1–18. DOI: 10.3389/fpsyg.2022.742956
8. Oliveri M., Mislevy R. Introduction to “Challenges and Opportunities in the Design of ‘Next-Generation Assessments of 21st Century Skills’”. Special Issue // International Journal of Testing. – 2019. – 19(2). – P. 97–102. DOI: 10.1080/15305058.2019.1608551
9. Shavelson R., Zlatkin-Troitschanskaia O., Beck K. et. al. Assessment of critical thinking: Next generation performance assessment // International Journal of Testing. – 2019. – No 19(9). – P. 337–362. DOI: 10.1080/15305058.2018.1543309
10. Nagel M.-T., Zlatkin-Troitschanskaia O., Fischer J. Validation of newly developed tasks for the assessment of generic Critical Online Reasoning (COR) of university students and graduates. Psychology // Frontiers in Education. – 2022. – P. 1–15. DOI:10.3389/educ.2022.914857
11. Родина Н. В., Постовалова А. А., Долматов А. В. Метод Evidence-Centered Design. Независимая оценка компетенций в цифровом мире // Образовательная политика. – 2020. – DOI: 10.22394/2078-838X-2020-3-52-59

12. Садова А. Р., Хиль Ю. С., Пашенко Т. В., Тарасова К. В. Измерение критического мышления взрослых: методология и опыт разработки // Современная зарубежная психология. – 2022. – № 11(4) – С. 105–116. DOI: 10.17759/jmfp.2022110409
  13. Тарасова К., Орел Е. Измерение критического мышления студентов в открытой онлайн-среде: методология, концептуальная рамка и типология заданий // Вопросы образования. Educational Studies Moscow. – 2022. – № 3. – URL: <https://doi.org/10.17323/1814-9545-2022-3-187-212>
  14. Углова И. Л., Брун И. В., Васин Г. М. Методология «Evidence-Centered Design» для измерения комплексных психологических конструктов // Современная зарубежная психология. – 2018. – Т. 7. – № 3. – С. 18–27. DOI: 10.17759/jmfp.2018070302
  15. Авдеева С. М., Руднев М. Г., Васин Г. М. и др. Оценка информационно-коммуникационной компетентности учащихся: подходы, инструмент, валидность и надежность результатов // Вопросы образования. Educational Studies Moscow. – 2017. – № 4. – С. 104–132.
  16. Dawson P., Bearman M. et al. Re-imagining University Assessment in a Digital World // The Enabling Power of Assessment. – 2020. – No 7. – P. 291–296. – URL: [https://doi.org/10.1007/978-3-030-41956-1\\_20](https://doi.org/10.1007/978-3-030-41956-1_20)
  17. Ефремова Н. Ф. Концептуальная модель оценки компетенций студентов // Современные наукоемкие технологии. – 2019. – № 7. – С. 169–174. DOI: 10.17513/snt.37607
  18. Ефремова Н. Ф. Приемы доказательной аргументации оценки компетенций // Инновационная наука: психология, педагогика, дефектология. – 2020. – № 3(2). – С. 112–124.
  19. Ефремова Е. Ф. Стратегии обеспечения справедливости в дизайне оценивания компетенций студентов // Инновационная наука: психология, педагогика, дефектология. – 2023. – № 6(1). – С. 81–93. – URL: <https://doi.org/10.23947/2658-7165-2023-6-1-81-93>
  20. Almond D., Blanken-Webb J. The Transformative Potency of Vulnerabilities in Strengths-Based Approaches to Higher Education // Academia Letters. – 2021. – Article 2540. – URL: <https://doi.org/10.20935/AL2540>
  21. Acerbi F., Rossi M., Terzi S. Identifying and Assessing the Required I4.0 Skills for Manufacturing Companies' Workforce // Front. Manuf. Technol. – 2022. – 2:921445. DOI: 10.3389/fmtec.2022.921445
  22. Mislevy R. J. Evidence and inference in educational assessment, Psychometrika. – 1994. – No 12. – P. 341–369.
  23. Mislevy R. J., Haertel G. D. Implications of evidence-centered design for educational testing // Educational Measurement: Issues and Practice. – 2006. – No 25(4). – P. 6–20.
  24. Shute V. J., Kim Y. J., Razzouk R. ECD for Dummies // J. Gee & D. Davidson (Eds.) Working Examples. – 2010. – 37 p. – URL: <https://studylib.net/doc/5805072/what-is-ecd%3F>
  25. Сердитова Н. Е., Белоцерковский А. В. Образование, качество и цифровая трансформация // Высшее образование в России. – 2020. – № 29(4). – С. 9–15. DOI: <https://doi.org/10.31992/0869-3617-2020-29-4-9-15>
  26. Cukurova M., Mavrikis M., Luckin R. Evidence-Centered Design and Its Application to Collaborative Problem Solving in Practice-based Learning Environments // UCL Knowledge Lab. – University College London, 2017. – P. 1-16.
  27. Snow E., Rutstein D., Basu S. et al. Leveraging Evidence-Centered Design to Develop Assessments of Computational Thinking // International Journal of Testing. – P. 103–127. – URL: <https://doi.org/10.1080/15305058.2018.1543311>
  28. Bechard S., Clark A., Romine R. Use of Evidence-Centered Design to Develop Learning Maps-Based Assessments // International Journal of Testing. – 2019. – 19(2). – P. 188-205. DOI: 10.1080/15305058.2018.1543310
  29. Russo D. Competency Measurement Model. European Conference on Quality in Official Statistics (Q2016). – Madrid, 2016. – P. 2–29.
  30. Kansal J., Jain N. Development of Competency Model and Mapping of employees Competencies for Organizational Development: A New Approach // Journal of Scientific & Industrial Research. – 2019. – Vol. 78. – P. 22–25.
- 
1. Mislevy, R., Haertel, G., Riconscente, M. et. al. (2017). "Assessing Model-Based Reasoning using Evidence-Centered Design", *A Suite of Research-Based Design Patterns*, SpringerBriefs in Statistics, 137 p. (in English).
  2. Toulmin, S. (1958, 2003). "The Uses of Argument, *Updated Edition*, Cambridge University Press, 259 p. (in English).
  3. Messick, S. (1994). "The interplay of evidence and consequences in the validation of performance assessments", *Educational Researcher*, 23(2), pp. 13–23 (in English).
  4. Griffin, P., Care, E. (2015). "Assessment and Teaching of 21st Century Skills", *Educational Assessment in an Information Age. Methods and Approach*, Springer Science+Business Media Dordrecht, 314 p. (in English).
  5. Mislevy, R. J., Almond, R. G., Lukas, J. F. (2003). "A brief introduction to evidence centered design", *ETS Research Report Series*, 29. DOI: 10.1002/j.2333-8504.2003.tb01908.x (in English).
  6. Mislevy, R. J., Steinberg, L. S., & Almond, R. A. (2003). "On the structure of educational assessments", *Measurement Interdisciplinary Research and Perspectives*, 1, pp. 3–67 (in English).
  7. Choi, Y., Mislevy, R. (2022). "Evidence centered design framework and dynamic bayesian network for modeling learning progression in online assessment system", *Frontiers Psychology*, 13:742956, pp. 1–18. DOI: 10.3389/fpsyg.2022.742956 (in English).
  8. Oliveri, M. & Mislevy, R. (2019). "Introduction to "Challenges and Opportunities in the Design of 'Next-Generation Assessments of 21st Century Skills'", *Special Issue*, *International Journal of Testing*, 19(2), pp. 97–102. DOI: 10.1080/15305058.2019.1608551 (in English).



9. Shavelson, R., Zlatkin-Troitschanskaia, O., Beck, K. et al. (2019). "Assessment of critical thinking: Next generation performance assessment", *International Journal of Testing*, 19(9), pp. 337–362. DOI: 10.1080/15305058.2018.1543309 (in English).
10. Nagel, M.-T., Zlatkin-Troitschanskaia, O., Fischer, J. (2022). "Validation of newly developed tasks for the assessment of generic Critical Online Reasoning (COR) of university students and graduates", *Psychology. Frontiers in Education*, pp. 1–15. DOI: 10.3389/feduc.2022.914857 (in English).
11. Rodina, N. V., Postovalova, A. A., & Dolmatov, A. V. (2020). "Metod Evidence-Centered Design. Nezavisimaya ocenka kompetencij v cifrovom mire" [The Evidence-Centered Design method. Independent assessment of competences in the digital world], *Obrazovatel'naya politika*. DOI: 10.22394/2078-838H-2020-3-52-59 (in Russian).
12. Sadova, A. R., Hil', Yu. S., Pashchenko, T. V., & Tarasova, K. V. (2022). "Izmerenie kriticheskogo myshleniya vzroslyh: metodologiya i opyt razrabotki" [Measuring critical thinking in adults: methodology and development experience], *Sovremennaya zarubezhnaya psihologiya*, № 11(4), pp. 105–116. DOI: 10.17759/jmfp.2022110409 (in Russian).
13. Tarasova, K., & Orel, E. (2022). "Izmerenie kriticheskogo myshleniya studentov v otkrytoj onlajn-srede: metodologiya, konceptual'naya ramka i tipologiya zadaniy" [Measuring students' critical thinking in an open online environment: methodology, conceptual framework and typology of tasks], *Voprosy obrazovaniya. Educational Studies Moscow*, № 3. Available at: <https://doi.org/10.17323/1814-9545-2022-3-187-212> (in Russian).
14. Uglanova, I. L., Brun, I. V., & Vasin, G. M. (2018). "Metodologiya "Evidence-Centered Design" dlya izmereniya kompleksnyh psihologicheskikh konstruktov" [Evidence-Centered Design methodology for measuring complex psychological constructs], *Sovremennaya zarubezhnaya psihologiya*, t. 7, № 3, pp. 18–27. DOI: 10.17759/jmfp.2018070302 (in Russian).
15. Avdeeva, S. M., Rudnev, M. G., Vasin, G. M. et al. (2017). "Ocenka informacionno-kommunikacionnoj kompetentnosti uchashchihsya: podhody, instrument, validnost' i nadezhnost' rezul'tatov" [Assessment of students' information and communication competence: approaches, tools, validity and reliability of results], *Voprosy obrazovaniya. Educational Studies Moscow*, № 4, pp. 104–132 (in Russian).
16. Dawson, P., Bearman, M., et al. (2020). "Re-imagining University Assessment in a Digital World", *The Enabling Power of Assessment*, 7, pp. 291–296. Available at: [https://doi.org/10.1007/978-3-030-41956-1\\_20](https://doi.org/10.1007/978-3-030-41956-1_20) (in English).
17. Efremova, N. F. (2019). "Konceptual'naya model' ocenki kompetencij studentov" [A conceptual model for assessing students' competences], *Sovremennye naukoemkie tekhnologii*, № 7, pp. 169–174. DOI: 10.17513/snt.37607 (in English).
18. Efremova, N. F. (2020). "Priemy dokazatel'noj argumentacii ocenki kompetencij" [Methods of evidence-centered argumentation of competence assessment], *Innovacionnaya nauka: psi-hologiya, pedagogika, defektologiya*, № 3(2), pp. 112–124 (in Russian).
19. Efremova, E. F. (2023). "Strategii obespecheniya spravedlivosti v dizajne ocenivaniya kompetencij studentov" [Strategies for ensuring equity in the design of student competency assessment], *Innovacionnaya nauka: psi-hologiya, pedagogika, defektologiya*, № 6(1), pp. 81–93. Available at: <https://doi.org/10.23947/2658-7165-2023-6-1-81-93> (in Russian).
20. Almond, D., & Blanken-Webb, J. (2021). "The Transformative Potency of Vulnerabilities in Strengths-Based Approaches to Higher Education", *Academia Letters*, Article 2540. Available at: <https://doi.org/10.20935/AL2540> (in English).
21. Acerbi, F., Rossi, M., & Terzi, S. (2022). "Identifying and Assessing the Required I4.0 Skills for Manufacturing Companies' Workforce", *Front. Manuf. Technol.*, 2:921445. DOI: 10.3389/fmtec.2022.921445 (in English).
22. Mislevy, R. J. (1994). "Evidence and inference in educational assessment", *Psychometrika*, 12, pp. 341–369 (in English).
23. Mislevy, R. J., & Haertel, G. D. (2006). "Implications of evidence-centered design for educational testing", *Educational Measurement: Issues and Practice*, 25(4), pp. 6–20 (in English).
24. Shute, V. J., Kim, Y. J., & Razzouk, R. (2010). "ECD for Dummies", in Gee, J., & Davidson D. (Eds.). *Working Examples*, 37 p. Available at: <https://studylib.net/doc/5805072/what-is-ecd%3F> (in English).
25. Serditova, N. E., & Belocerkovskij, A. V. (2020). "Obrazovanie, kachestvo i cifrovaya transformaciya" [Education, quality and digital transformation], *Vyshee obrazovanie v Rossii*, № 29(4), pp. 9–15. DOI: <https://doi.org/10.31992/0869-3617-2020-29-4-9-15> (in Russian).
26. Cukurova, M., Mavrikis, M., & Luckin, R. (2017). "Evidence-Centered Design and Its Application to Collaborative Problem Solving in Practice-based Learning Environments", *UCL Knowledge Lab*, University College London, pp. 1–16 (in English).
27. Snow, E., Rutstein, D., Basu, S., et al. "Leveraging Evidence-Centered Design to Develop Assessments of Computational Thinking", *International Journal of Testing*, pp. 103–127. Available at: <https://doi.org/10.1080/15305058.2018.1543311> (in English).
28. Bechard, S., Clark, A., & Romine, R. (2019). "Use of Evidence-Centered Design to Develop Learning Maps-Based Assessments", *International Journal of Testing*, 19(2), pp. 188–205. DOI: 10.1080/15305058.2018.1543310
29. Russo, D. (2016). *Competency Measurement Model. European Conference on Quality in Official Statistics (Q2016)*. Madrid, pp. 2–29 (in English).
30. Kansal, J., & Jain, N. (2019). "Development of Competency Model and Mapping of employees Competencies for Organizational Development: A New Approach", *Journal of Scientific & Industrial Research*, vol. 78, pp. 22–25 (in English).