

ОБЩАЯ ПЕДАГОГИКА



УДК 371.263

Научная статья

<https://doi.org/10.23947/2658-7165-2023-6-1-81-93>

Стратегии обеспечения справедливости в дизайне оценивания компетенций студентов

Н. Ф. Ефремова  

Донской государственный технический университет, Российская Федерация, г. Ростов-на-Дону, пл. Гагарина, 1

 nefremova61@donstu.ru

Аннотация

Введение. В статье представлено текущее положение рассматриваемой проблемы в работах отечественных и зарубежных авторов. Рассмотрены трудности оценивания компетенций как сложных, глубоко латентных и междисциплинарных по своей природе результатов обучения. Актуальность исследования обусловлена потребностью обеспечения надежности оценок. Новизна исследования заключается в представлении алгоритма разработки оценки компетенций на основе моделирования конструкта и условий для проявления компетенций в процессе деятельности студентов при выполнении комплексного задания.

Цель. Системное представление современных подходов к аутентичному оцениванию компетенций с опорой на теорию и практику педагогических измерений, опыт конструирования и применение средств по моделям доказательной аргументации, обеспечивающих справедливость оценочного процесса и аутентичность результатов оценки.

Материалы и методы. В качестве теоретических методов были использованы работа с источниками, сбор, обобщение, структурирование и анализ информации о существующих подходах к оцениванию компетенций обучающихся в России и за рубежом. В практической части был использован опыт организации и проведения различных форм оценки результатов обучения в условиях компетентностного обучения студентов.

Результаты исследования. Предложены современные подходы и механизмы проектирования оценочных средств по моделям аргументированного доказательства результатов оценки с опорой на теорию и практику педагогических измерений. Созданы расширенные паттерн-дизайны для преподавателей разработчиков средств и процедур оценки, обеспечивающее аутентичность результатов за счет анализа и интерпретации данных с учетом контекстной информации о деятельности испытуемых в ситуации оценки.

Обсуждение и заключения. Проанализирована взаимосвязь проблематики оценивания компетенций с их сложностью и междисциплинарностью, предложено в процессе разработки оценочного инструментария обращение к моделям С. Тулмина и Р. Мислеви. Показаны тенденции развития контрольно-оценочной деятельности для диагностики качества обучения.

Ключевые слова: компетенции, оценка, доказательства, паттерн-дизайн, модели аргументации, разработчики оценочных средств.

Для цитирования. Ефремова, Е. Ф. (2023). Стратегии обеспечения справедливости в дизайне оценивания компетенций студентов. *Инновационная наука: психология, педагогика, дефектология*, 6(1), 81–93. <https://doi.org/10.23947/2658-7165-2023-6-1-81-93>

Equity Strategies in the Design of Student Competency Assessment

N. F. Efremova  

Don State Technical University, 1, Gagarin Sq., Rostov-on-Don, Russian Federation

 nefremova61@donstu.ru

Abstract

Introduction. The article presents the current situation of the problem under consideration in the works of Russian and foreign authors. The difficulties of assessing competencies as complex, deeply latent and interdisciplinary in nature learning outcomes are considered. The study is relevant due to the need to ensure the reliability of estimates. Its novelty lies in the presentation of an algorithm for the development of competence assessment based on the modelling of the construct and conditions for the manifestation of competencies in the process of students' activities when performing a complex task.

Purpose. A systematic presentation of modern approaches to the authentic assessment of competencies based on the theory and practice of pedagogical measurements and experience in designing and using tools based on models of evidence-based argumentation that ensure the fairness of the evaluation process and the authenticity of the evaluation results.

Materials and Methods. As theoretical methods, work with sources, collection, generalization, structuring and analysis of information on existing approaches to assessing the students' competencies in Russia and abroad were used. In the practical part, the experience of organizing and conducting various forms of evaluation of learning outcomes in the context of competence-based learning of students was used.

Results. Modern approaches and mechanisms for designing evaluation tools based on models of reasoned proof of evaluation results based on the theory and practice of pedagogical measurements are proposed. Extended pattern designs have been created for teachers, and developers of assessment tools and procedures, ensuring the authenticity of the results through the analysis and interpretation of data, considering contextual information about the activities of the subjects in the assessment situation.

Discussion and Conclusion. The interrelation of the problems of competence assessment with their complexity and interdisciplinarity is analyzed, and an appeal to the models of S. Tulmin and R. Mislevi is proposed in the process of developing evaluation tools. The trends in control development and evaluation activities for the measurement of training quality are shown.

Keywords: competency, measurement, evidence, pattern-design, argument model, assessment tools developers.

For citation. Efremova, N. F. (2023). Equity Strategies in the Design of Student Competency Assessment. *Innovative Science: psychology, pedagogy, defectology*, 6(1), 81–93. <https://doi.org/10.23947/2658-7165-2023-6-1-81-93>

Введение. Одной из важнейших задач профессионального обучения является формирование компетенций как важных характеристик для профессиональной деятельности, и в современной инновационной образовательной среде оценка уровня таких компетенций, по своей сути выступающих в качестве результата получения образования, представляет особую актуальность. У работодателей растут потребности в сотрудниках, способных думать критически, работать в сотрудничестве, выдвигать идеи и использовать новые технологии. Иными словами, помимо профессиональных знаний и навыков, важными при конкурсном отборе являются мягкие навыки или компетенции. Сегодня компьютеры способны выполнять самые разнообразные виды работ, таким образом заменяя человека, поэтому на рынке труда от современных специалистов требуется овладение более сложными и востребованными навыками XXI века: критическое мышление, креативность, коммуникация и кооперация, инициатива, стрессоустойчивости, ответственность, оправданный риск, эффективное управление, организация, и владение метакогнитивными навыками — умением обучаться. П. Гриффин, профессор Мельбурнского университета, лидер в исследованиях нового содержания компетенций и навыков XXI века и новых методов преподавания и оценки достижений, с коллегами отмечают, что результаты образования все более смещаются в сторону развития универсальных компетенций (Care, Griffin, Wilson 2018; Griffin, McGaw, Care, 2012).

При гигантском технологическом и социальном сдвиге в общественном секторе такого же сдвига в методологии компетентностного обучения пока не наблюдается. Во-первых, нетривиальная задача оценки результатов обучения пока не имеет эффективного решения. Несмотря на то, что по этому вопросу уже существует значительный объем литературы, все еще остается невыясненным, как правильно формировать и оценивать компетенции. Традиционное оценивание преподавателями освоения студентами содержания предметных областей, как правило, ассоциируется с оценкой компетенций, что связано с отсутствием понимания специфических особенностей компетенций и условий, при которых возможны их формирование, проявление и оценивание. Во-вторых, имеет место отсутствие нормативно обусловленных требований к оцениванию компетенций. В силу этих и многих других причин возникает необходимость в разработке новой модели образования, как отвечающей на вызовы времени, так и выполняющей, как было всегда, свою основную задачу — наследование культуры, выработанной человечеством. Постановка акцента на формировании и надежном оценивании компетенций может послужить фундаментом для установления связи между освоением основной образовательной программы и запросами работодателей к подготовленности специалистов.

В мировой практике новые подходы к оценке компетенций находят широкое применение. Они насыщаются инновационными технологиями проектирования и разработки оценочного инструментария. В отечественном образовании такие подходы пока не имеют научного обоснования и не получают практического применения. Таким образом, цель данного исследования заключается в выявлении, обобщении и представлении принципов проектирования, разработки и применения средств и процедур, создающих условия для справедливого и надежного оценивания компетенций студентов.

Материалы и методы. При решении поставленной цели были использованы теоретические методы, такие как работа с источниками, сбор, обобщение и структурирование информации. Проведен анализ существующих в России и за рубежом методов оценивания навыков и компетенций. Использован собственный опыт организации и проведения различных форм оценки результатов обучения и качества образовательного процесса, в том числе в условиях перехода системы образования на компетентностное обучение.

Результаты исследования. Практика показывает, что с компетентностным обучением связано много проблем, и обеспечение надежности оценок компетенций является одной из них (Ефремова, 2019; Углова, Брун, Васин, 2018; Чандра, Байкина, 2018). Доминирующая традиционная система оценки образовательных достижений не обеспечивает студентам равные возможности демонстрации этих самых достижений и способностей, также отсутствует сбор подтверждающих доказательств, дающих надежную интерпретацию результатов обучения. Отличные от традиционных способов современные дизайнерские стратегии аутентичного оценивания компетенций предоставляют возможности планировать и обеспечивать условия деятельности студентов в ситуации оценки, наблюдать и оценивать фокальные знания и действия по их применению для повышения достоверности результатов.

На первом этапе планирования такой оценки необходимо создание конструкта (модели студента), позволяющего включить в оценочное средство несколько целей оценки и разработать комплексное задание на базе ситуаций, максимально приближенных к реальным жизненным условиям (Weideman, 2012). Именно природа конструкта, как отмечал С. Мессик (Messick, 1994), обеспечивает разработку релевантных заданий и оценочных шкал для подсчета баллов и оценочных суждений об уровнях сформированности компетенций.

Создание конструкта для проектирования оценки компетенций начинается с ряда вопросов:

- Какие компетенции (глубоко латентные по своей природе) подлежат оценке и какими оценочными средствами они должны измеряться?
- Какие образовательные результаты имеют ценность для управления образовательным процессом и для работодателей?
- Каким образом необходимо планировать деятельность студента, чтобы он смог проявить свои компетенции?
- Какие ситуации, близкие к реальным процессам, могут вызвать нужное поведение студента?
- Какие критерии, доказательства, аргументации и опровержения должны обеспечить аутентичность и справедливость результатов оценки?

Комплексное оценивание ставит перед студентом ряд задач, решение которых обеспечивает демонстрацию его владения знаниями и компетенциями в соответствии с целями оценки. Оценка должна проводиться строго в рамках конструкта и не требовать от студентов демонстрации знаний и действий, не относящихся к планируемым результатам. В первую очередь это обеспечивается четким обозначением в конструкте цели и задач оценочного процесса путем выделения оцениваемых атрибутов (рис. 1).

Оценочные средства проектируются с учетом оптимального сочетания между традиционной практикой и современным дизайном оценки, предусматривающим планируемую в ситуации оценки деятельность студента с применением имеющихся знаний и допустимого числа стимульных обращений к источникам за необходимой информацией. На практике же часто получается, что разработанный инструмент не показывает нужного результата, так как разработчики не проверяют соответствие задуманного конструкта и реальных результатов измерения, в то время как для корректного оценивания необходим статистический анализ эмпирических данных после предварительного применения (апробации) оценочного средства.

Модельное обоснование оценки результатов обучения студентов с использованием ориентированного на результат проектирования оценочных средств и процессов предусматривает использование наборов шаблонов (паттерн-дизайнов многократного использования), уже прошедших апробацию и показавших свою эффективность (Ефремова, 2018б; Ефремова, 2019; Mislevy et al., 2017). Их можно использовать в разных предметных областях и средах образовательной деятельности.

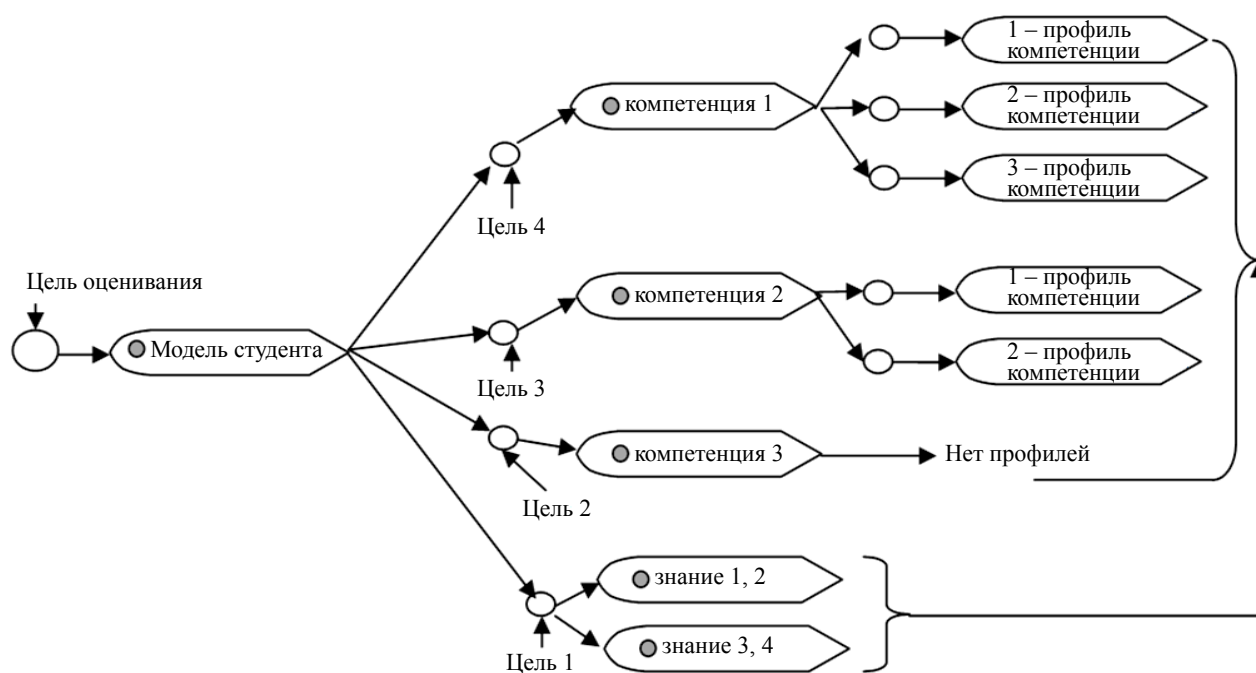


Рис. 1. Пример многомерного конструкта (Ефремова, 2018а)

С. Тулмин (Нефедов, 2017; Toulmin, 2003) предложил схему проектирования и логической структуры аргументов для оценок, не имеющих четких и однозначных ответов. Она состоит из нескольких взаимосвязанных элементов (рис. 2). Основным является утверждение *C* (claim) на базисе доказательств или предположений с учетом критериев, которые необходимо обосновать и доказать. При этом выводы и аргументации опираются на данные *D* (data), а квалификатор *Q* (qualifier) обеспечивает в утверждении степень обоснованности выводов, отвечает на вопрос, являются ли они случайными или истинными при интерпретации данных. Помимо этого, влияние на выводы могут оказывать обоснование *W* (warrant) как цепь рассуждений, устанавливающих логическую связь между данными и утверждением; поддержка *B* (backing) — аргументы и соображения, подкрепляющие обоснование; опровержение *R* (rebuttal), как контраргумент, это могут быть возможные оговорки или возражения, в том числе противоречивые, позволяющие формировать выводы с различной степенью доверия.



Рис. 2. Общая схема аргументации С. Тулмина

В поддержку могут быть включены контрольные вопросы, результаты опросов и анкетирования, дополнительные данные об испытуемых, которые можно использовать при интерпретации результатов оценки. Квалификатор подчеркивает или ограничивает достоверность утверждения об оценке при помощи набора слов (вероятно, возможно, так, поскольку, если не, невозможно, безусловно, предположительно). Допущения и предположения с учетом контекстных данных являются важным основанием для вывода, обеспечивая за счет дополнительной информации и уточнений, заключение о подготовленности испытуемого. Переход от свидетельств к утверждению считается наиболее важным в доказательствах того, что анализ данных и аргументации были проведены успешно. Такой подход из набора взаимосвязанных процедур направлен на решение двух основных проблем оценки: наблюдение за поведением испытуемого при взаимодействии с заданием (свидетельства проявления компетенции или знаний) и структурирование и документирование свидетельств, заложенных в средстве оценки (ситуации).

На основе приведенных выше рассуждений С. Тулмина, автором статьи, с учетом структуры аргументов и доказательств его теории, предложены расширенные паттерн-дизайны проектирования надежных и валидных средств оценки для преподавателей, стремящихся получить достоверные сведения об уровнях компетенций студентов (рис. 3, 4). Это позволяет планировать ожидаемое поведение студентов в ситуации оценки, фокальные действия по заявленным критериям, предоставляет достаточное количество аргументированных данных для документирования полученной информации и позволяет делать достоверные выводы о компетенциях студентов.

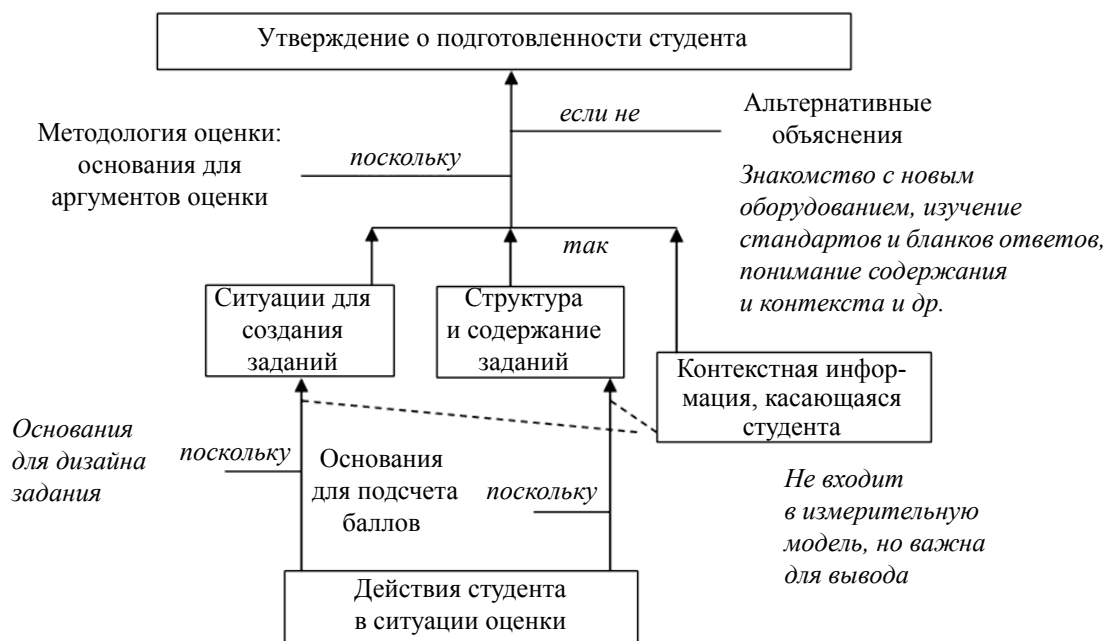


Рис. 3. Паттерн-дизайн аргументов для обоснования оценки

В средствах оценки планируемая деятельность студентов требует умений критически мыслить, взаимодействовать в малых группах с другими обучающимися и специалистами, творчески подходить к своей деятельности при решении комплексных нестандартных задач. Такая оценка, обеспечивая диагностическую обратную связь, интегрирует ее в учебный процесс и ориентирует на использование цифровых технологий в виртуальных и связанных с ними симуляционных средах обучения. При этом паттерн-дизайн должен, с одной стороны, соответствовать поставленной задаче и требованиям к освоению программы, а с другой — быть достаточно общим, чтобы учесть возможности решения подобных проблем и иметь возможность многократного использования в схожих задачах оценивания результатов обучения. Все это показывает, что на сегодняшний день существует необходимость осваивать и использовать новые технологии оценки, которые могут помочь в достижении такой сложной цели, как оценка глубоко латентных характеристик студентов, представленных в форме компетенций. Определение оценки согласно принципам доказательной аргументации стало нужным в образовательном контексте в тот момент, когда появилась необходимость в ходе оценки интегрировать большое количество данных и определять сложные результаты образования, такие как навыки и компетенции, необходимые человеку XXI века. Авторами систематического подхода к разработке оценочных заданий на основе моделей Evidence Centered Design (ECD) или методологии доказательного дизайна являются Р. Мислеви с соавт. (Almond et al., 2019; Mislevy, Almond, Lukas, 2003; Mislevy, Levy, 2007; Oliveri, Mislevy, 2019). Здесь важно отметить построение концептуальной рамки оценки (Conceptual Assessment Framework — CAF) (рис. 5).

Модель концептуальной рамки доказательного дизайна оценки отражает последовательное выстраивание модели студента (набор проверяемых компетенций и знаний), модели свидетельств и доказательств (оценочный компонент, измерительный компонент, правила определяющие существенные особенности продукта деятельности, индивидуальные баллы, средний балл по выборке), модели задания и модели презентации результатов (сборка оценочного средства). Для корректного функционирования модели необходимо описание сценария оценочного процесса, подготовка бланков наблюдений за деятельностью испытуемых для сбора опровержений и поддержки, используемых при интерпретации результатов. Наибольшую трудность для разработчиков представляет подбор ситуаций, при решении которых возможно проявление и подтверждение компетенций. Такие ситуации должны представлять структурированные проблемы, с одной стороны, имеющие готовые/известные решения, а с другой, предполагающие множество дополнительных решений.

При планировании и разработке доказательно аргументированной оценки должны быть подготовлены ответы на следующий ряд вопросов. На вопрос «Для чего измеряем?» ответ дает модель студента (в структуре конструкта — фокальные знания и компетенции, подлежащие оцениванию). На вопрос «Как измеряем?» ответом является модель доказательств (формальная основа измеряемых атрибутов оценки с учетом условий, в которых проводятся наблюдения за деятельностью студента, а также сбор доказательств для правильного вывода). «Что измеряем?» — модель задания раскрывает, какие ситуации используются и в какой форме ставятся задачи. На следующий вопрос «Чем измеряем?» отвечает концептуальная рамка оценивания или модель сборки (в целом оценочное средство), в нем должно быть учтено, как все модели взаимодействуют между собой. Наконец, на вопрос «Как обеспечиваем анализ результатов?» ответ дают модели протоколов о результатах оценки. Все эти модели являются мостом между аргументом надежной оценки и деятельностью преподавателя по проектированию и разработке комплекса доказательств. Использование принципов ECD обеспечивает заранее заданные квалитетические свойства педагогического измерителя и снижает число ошибок при измерении результатов обучения за счет обоснования проектных решений на этапе конструирования оценочного средства еще до того, как будет разработан и использован его первый вариант. Для создания комплексного задания, проверяющего набор (кластер) однотипных компетенций, могут быть использованы различные задачи: компетентностно-ориентированные, наборы элементов заданий с вариантами решений в условиях неопределенности, также возможен поиск недостающей информации, презентация, подготовка статьи для публикации в открытой печати, участие в деловой игре, решение кейсов, выполнение проектов индивидуальных или в малых группах.

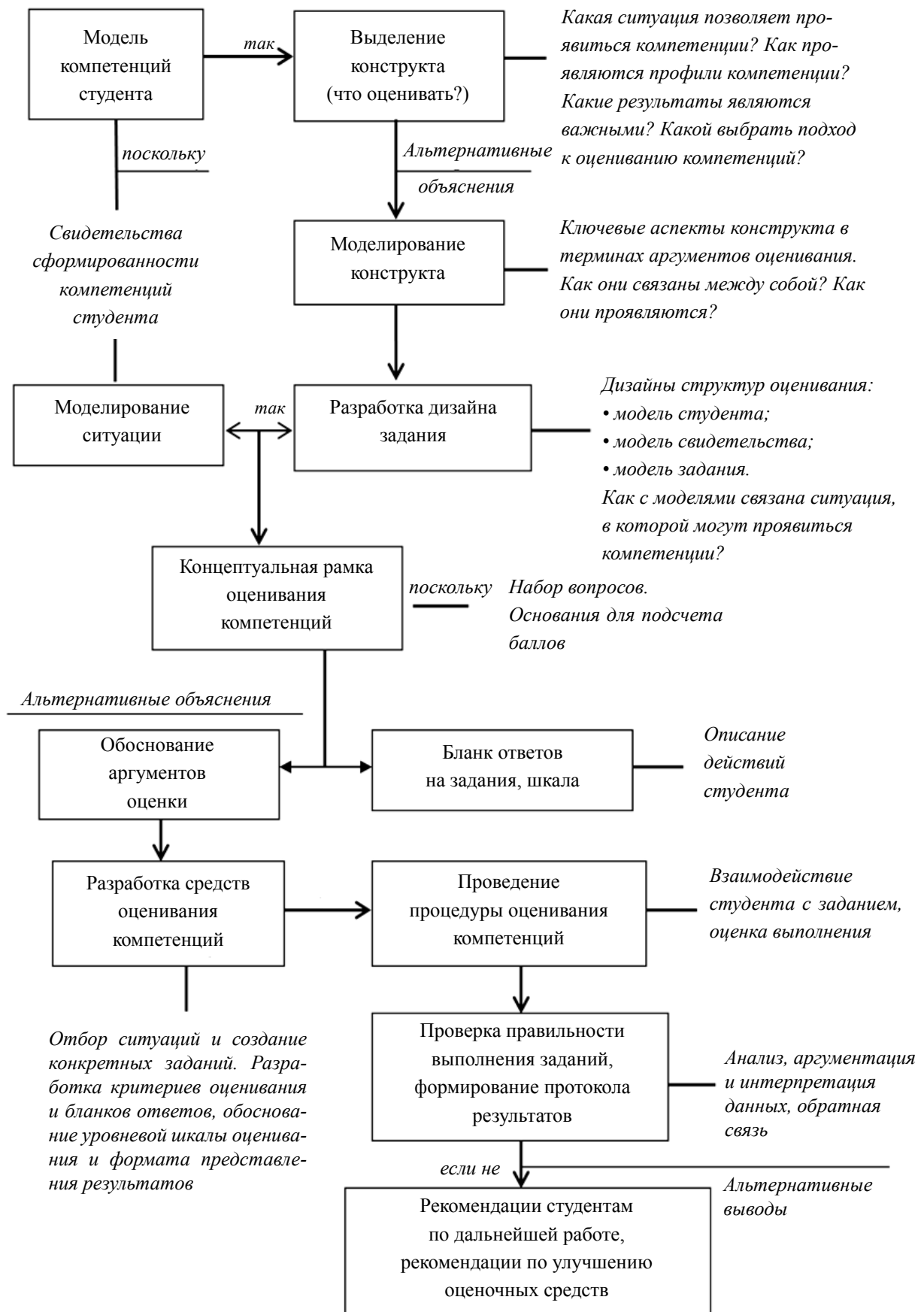


Рис. 4. Паттерн проектирования оценки компетенций

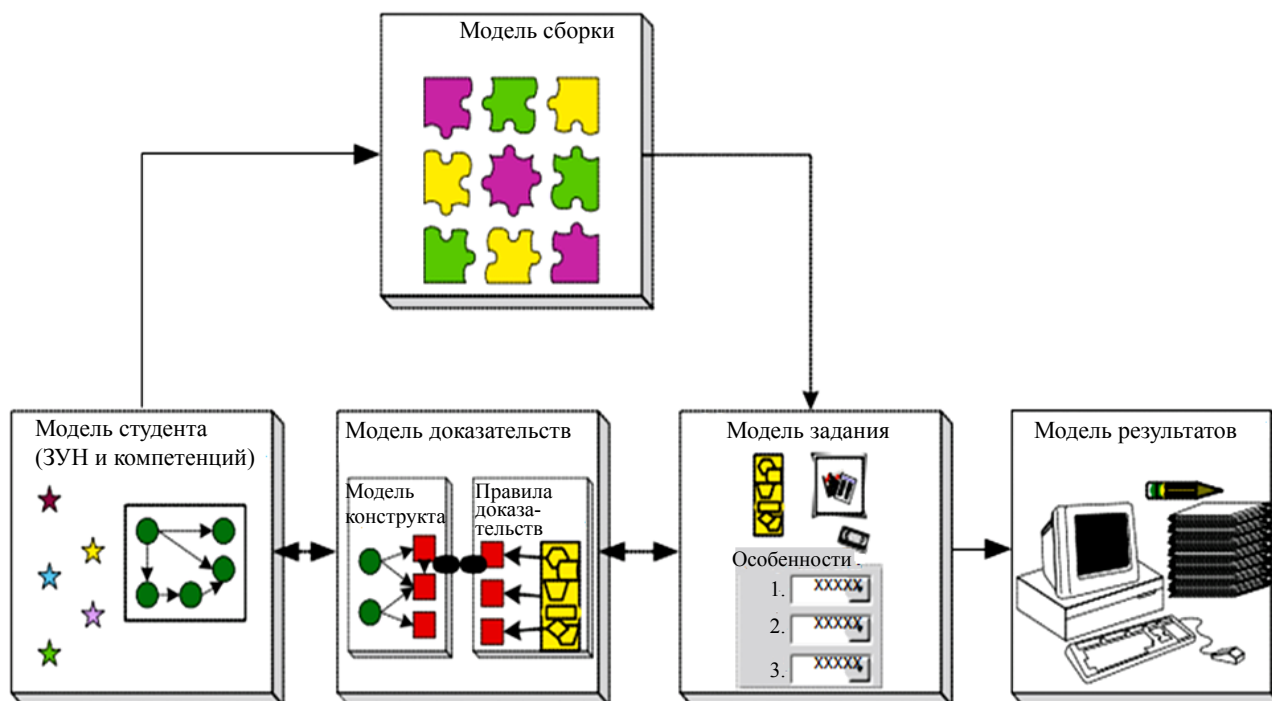


Рис. 5. Модель концептуальной рамки доказательного дизайна оценки (Р. Мислеви)

Ориентированный на доказательства дизайн предоставляет несколько преимуществ в разработке педагогических измерителей. Согласно спецификации оценочного средства, в нем обеспечивается логичность каждого последующего шага проектирования с опорой на предыдущие этапы. Основа таких рассуждений охватывает весь процесс разработки, поддерживает общий аргумент оценки, обеспечивает валидность и надежность результатов оценочных процедур. Соответствующими методами доказательств можно выявить, насколько хорошо студенты планируют способы решения проблем, выполняют задания, анализируют и интерпретируют качественные и количественные данные. К доказательствам можно отнести умозаключения, интерпретации, выводы и предложения, все, на что ссылается утверждение. Это могут быть статистические данные исследований, факты, артефакты, свидетельства, мнения экспертов, примеры — все то, с помощью чего аргументируется доказательство достижений.

Таким образом, Р. Мислеви предложил принципиально новый дизайн ECD для проектирования оценочных средств как подход, ориентированный на доказательство и аргументацию результатов оценивания. Стратегия такого дизайна оценки направлена на выявление взаимосвязи знаний студентов и умений их использования при решении проблем. Взятые вместе это обеспечивает формальную основу для интеллектуального перехода от промежуточных баллов к выводам об уровне владения компетенциями и поддержки правильного вывода об оценке. ECD — это язык для описания процесса разработки оценочного средства (ОС), который предусматривает сбор и документирование информации о структуре и силе доказательных отношений. Он задает согласование работ разработчиков и психометриков при калибровке заданий измерительной модели ОС, обеспечивающей основу для надежности результатов оценки. Оценивание компетенций в рамках методологии Evidence-Centered Design — это процесс доказательной аргументации для вывода о способности студента применять знания при решении нестандартной задачи на основе свидетельств о результатах его деятельности в ситуации оценки. Центральным понятием здесь является суждение о студенте и о том, что он знает и умеет делать после освоения дисциплины или образовательной программы.

Отличительные особенности традиционного подхода к оценке и ECD отражены в таблице 1.

Таблица 1

Традиционный подход и ECD к оценке образовательных достижений студентов	
Традиционный подход	Evidence-Centered Design
Оценка знаний/умений	Результаты обучения студентов (знания/навыки и компетенции), планируемые к оцениванию в форме конструкта
Нет аналога	Наблюдаемые переменные в ситуации взаимодействия испытуемого с заданием (деятельность)
Ответы на задания когнитивного свойства	Результаты деятельности студента в структурированной ситуации оценочной задачи
Содержание задания на оценку освоения контролируемого содержания	Контент, контекст, ситуация, средства презентации, средства администрирования, доказательства и опровержения, допущения о знаниях/компетенциях из других предметных областей
Оценка в пятибалльной шкале	Аргументы и свидетельства (доказательства) для повышения надежности оценки компетенций на уровневой шкале.

Опираясь на принципы ECD, последовательность действий разработчиков при проектировании, организации и проведении оценки компетенций студентов можно предложить в следующем виде:

- Определение цели оценки как основы для применения метода ECD;
- Построение многомерного конструкта из знаний, умений, компетенций или других оцениваемых атрибутов с опорой на определенный объем необходимых знаний (анализ содержания контролируемой области, выделение измеряемых переменных, выбор фокальных компетенций);
- Операционализация задания для создания условий обработки (желательно автоматизированной) и представления результатов оценки в удобном для пользователей виде;
- Разработка концептуальной основы инструмента оценки, состоящей из моделей студента, доказательств, задания;
- Подбор интересных студентам ситуаций и задач, разработка спецификации оценочного средства (каких и сколько ситуаций и задач необходимо, предполагаемые решения) и инструкций (в каком виде испытуемый дает ответ);
- Обоснование критериев достижений и уровневой шкалы оценивания;
- Обеспечение комфортных условий деятельности обучающихся при взаимодействии с инструментами оценки;
- Проведение оценочного процесса, сбор данных и свидетельств достижения планируемых результатов;
- Статистическая обработка результатов оценки с учетом аргументации свидетельств и опровержений, отбраживание суммы полученных баллов на уровневой шкале;
- Анализ и интерпретация результатов, подготовка протоколов и обеспечение доступности информации.

Важно отметить, что декомпозиция компетенций на уровни и индикаторы освоения имеет смысл как развернутая программа для формирования тех или иных составляющих компетенций, но она становится непригодной при оценивании. Сложная структура и междисциплинарный характер компетенций требуют интегрированного подхода к их оцениванию, перехода к модульному построению образовательных программ, когда после освоения модуля можно считать сформированными однотипные компетенции, сгруппированные в отдельный кластер. Надежная оценка достижений и их обоснованная интерпретация создают основу для обратной связи между обучением, оценкой, ресурсами образовательной программы и разработкой политики дальнейшего обучения (Рис. 6).

Необходимость освоения новых моделей оценки и аргументированного доказательства результатов оценивания достижений студентов становится все более значимой, особенно в ключе того, чтобы достоверно выявлять действительно ли они обладают необходимым набором компетенций для решения стандартных и нестандартных проблем, с которыми столкнутся в жизни и профессиональной деятельности. Сочетание индивидуальной и кооперативной деятельности при выполнении комплексных заданий является социальным аспектом обучения, в котором проверяется готовность общения и обмена идеями, идентификация проблемы и ее отдельных элементов,

определение путей и способов ее решения, достижение соглашения в группе, анализ и рефлексия результатов, оценка их последствий, подготовка и презентация итогового отчета.

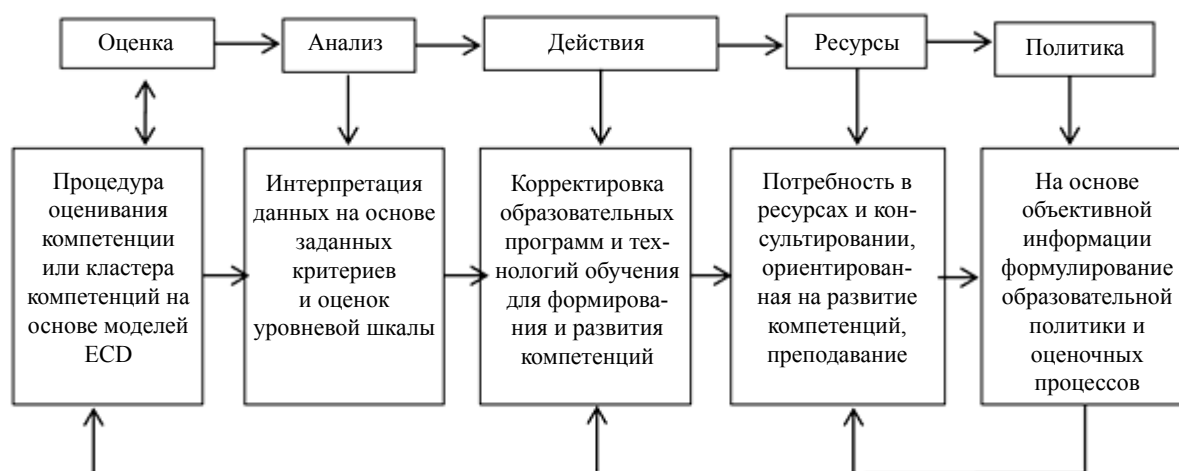


Рис. 6. Пятиступенчатый подход к оценке компетенций

Обсуждение и заключения. Переход от знаниевой парадигме к компетентностной сегодня характеризуется тем, что при широком использовании второй терминология технологии обучения и оценивание результатов обучения остаются когнитивными, а оценка знаний выдается за оценку компетенций. Поскольку компетенции имеют глубоко латентный характер, обеспечение деятельности студентов требует от разработчиков оценочных средств и обращения к теории педагогических измерений для обеспечения надежности и валидности результатов оценки, что представляет определенные сложности на теоретическом и практическом уровнях. Еще одной проблемой является то, что образовательные программы носят дисциплинарный характер, а компетенции не формируются только одной дисциплиной, они междисциплинарны по своей природе. Это затрудняет возможность проведения оценки сформированности компетенции в рамках освоения только лишь одной дисциплины, что отмечается в ряде работ отечественных авторов (Авдеева и др., 2017; Чандра, Байкина, 2018).

Современный дизайн средств оценивания результатов компетентностного обучения направлен на выявление взаимосвязи знаний студентов со стратегиями решения проблем в планируемой деятельности при выполнении комплексных оценочных заданий. Используя модели Тулмина и Мислеви в процессе разработки и применения оценочного инструментария, можно спроектировать действия студента в ситуации оценки. Разработанные таким образом оценочные средства позволят выявить, как компетентностные результаты извлекаются из хода работы студента над заданием, как эти извлечения накапливаются и по каким критериям оцениваются его действия для обеспечения надежных суждений об уровнях освоения компетенций (Jumariati, Febriyanti, Rizki, 2021).

За последние два десятилетия в мире и системе образования произошли существенные изменения, требующие от специалистов особых навыков для решения сложных проблем и «выживания» в среде неопределенности. Их оценка выходит за рамки традиционных методов контроля и ставит задачи разработки новых подходов, обеспечивающих аутентичность результатов оценивания достижений студентов. Такие компетенции, как, например, метакогнитивные (умение учиться, думать о собственном мышлении, уметь планировать, контролировать и оценивать собственное мышление и понимание), не могут быть оценены традиционными методами или средствами тестирования учебных достижений. Новые подходы к проектированию оценочных средств и процедур, направленных на определение компетенций и измерение навыков XXI века, только начинают разрабатываться и применяться. Методология ECD, получившая распространение в зарубежной практике, может использоваться в качестве основы для обеспечения надежности и валидности оценок при компетентностном обучении и в отечественном образовании. Предложенные шаблоны проектирования оценочных процессов (паттерн-дизайны) могут быть полезны преподавателям при создании инновационных заданий относительно целей оценки компетенций студентов.

Список литературы

- Авдеева, С. М., Руднев, М. Г., Васин, Г. М. и др. (2017). Оценка информационно-коммуникационной компетентности учащихся: подходы, инструмент, валидность и надежность результатов. *Вопросы образования*, 4, 104–132. <https://doi.org/10.17323/1814-9545-2017-4-104-132>
- Ефремова, Н. Ф. (2018а). Проектирование оценочных средств по модели доказательной аргументации. *Педагогические измерения*, 3, 8–15.
- Ефремова, Н. Ф. (2018б). Аргументации и доказательства надежности оценок компетенций студентов. *Известия Дагестанского государственного педагогического университета. Психолого-педагогические науки*, 12(2), 43–50.
- Ефремова, Н. Ф. (2019). Концептуальная модель оценки компетенций студентов. *Современные наукоемкие технологии*, 7, 169–174. <https://doi.org/10.17513/snt.37607>
- Нефедов, С. Н. (2017). Модель доказывания Тулмина и количественные показатели достоверности вывода. *Вопросы криминологии, криминалистики и судебной экспертизы*, 1(41), 87–93.
- Углонова, И. Л., Брун, И. В., Васин, Г. М. (2018). Методология Evidence-Centered Design для измерения комплексных психологических конструкторов. *Современная зарубежная психология*, 7(3), 18–27. <https://www.doi.org/10.17759/jmfp.2018070302>
- Чандра, М. Ю., Байкина, Е. А. (2018). Стратегии оценивания компетенций студентов в процессе освоения образовательной программы вуза. *Известия Волгоградского государственного педагогического университета*, 7(130), 10–15.
- Almond, R. G., Mislevy, R. J., Steinberg, L. et al. (2019). *Bayesian networks in educational assessment*. New York: Springer.
- Care, E., Griffin, P., Wilson, M. (2018). *Assessment and Teaching of 21st Century Skills. Research and Applications*. New York: Springer. <http://dx.doi.org/10.1007/978-3-319-65368-6>
- Griffin, P., Care, E., McGaw, B. (2012). The Changing Role of Education and Schools. In P. Griffin, B. McGaw, E. Care (eds): *Assessment and Teaching of 21st Century Skills. Research in Science Teaching* (pp. 1–15). New York: Springer. https://www.doi.org/10.1007/978-94-007-2324-5_1
- Jumariati, Febriyanti, E. R., Rizki, M. (2021). Argumentation Skills: An Analysis on EFL Students' Essays Based on Toulmin's Model of Argument. In: *Advances in Social Science, Education and Humanities Research, Proceedings of the 2nd International Conference on Education, Language, Literature, and Arts (ICELLA 2021)*, 587 (pp 86–95). <https://www.doi.org/10.2991/assehr.k.211021.013>
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Education Researcher*, 23(2), 13–23. <https://www.doi.org/10.3102/0013189x023002013>
- Mislevy, R. J., Almond, R. G., Lukas, J. F. (2003). A Brief Introduction to Evidence-centered Design. *Educational Testing Service*, 2003(1), i–29. <https://www.doi.org/10.1002/J.2333-8504.2003.TB01908.X>
- Mislevy, R. J., Haertel, G., Riconscente, M. et al. (2017). *Assessing Model-Based Reasoning using Evidence-Centered Design: A Suite of Research-Based Design Patterns*. New York: Springer. <https://www.doi.org/10.1007/978-3-319-52246-3>
- Mislevy, R. J., Levy, R. (2007). Bayesian psychometric modeling from an evidence-centered design perspective. *Handbook of statistics*, 26, 839–865. [https://doi.org/10.1016/S0169-7161\(06\)26026-7](https://doi.org/10.1016/S0169-7161(06)26026-7)
- Oliveri, M. E., Mislevy, R. J. (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), 97–102. <https://www.doi.org/10.1080/15305058.2019.1608551>
- Toulmin, S. E. (2003). *The Uses of Argument*. Cambridge: Cambridge University Press.
- Weideman, A. (2012). Validation and Validity Beyond Messick. *Per Linguam*, 28(2), 1–14. <https://doi.org/10.5785/28-2-526>

References

- Almond, R. G., Mislevy, R. J., Steinberg, L. et al. (2019). *Bayesian networks in educational assessment*. New York: Springer.
- Avdeeva, S. M., Rudnev, M. G., Vasin, G. M. et al. (2017). Assessment of information and communication competence of students: approaches, tool, validity and reliability of results. *Educational Studies*, 4, 104–132. <https://doi.org/10.17323/1814-9545-2017-4-104-132> (In Russ.).
- Care, E., Griffin, P., Wilson, M. (2018). *Assessment and Teaching of 21st Century Skills. Research and Applications*. New York: Springer. <http://dx.doi.org/10.1007/978-3-319-65368-6>
- Chandra, M. Yu., Baikina, E. A. (2018). Strategies for assessing students' competencies in the process of mastering the educational program of the university. *Bulletin of Volgograd State Pedagogical University*, 7(130), 10–15. (In Russ.).
- Efremova, N. F. (2018b). Arguments and Evidence of Reliability Assessment of Students' Competencies. *DSPU Journal, Psychological and pedagogical sciences*, 12(2), 43–50. (In Russ.).
- Efremova, N. F. (2019). Conceptual model for assessing students' competencies. *Modern science-intensive technologies*, 7, 169–174. <https://doi.org/10.17513/snt.37607> (In Russ.).
- Griffin, P., Care, E., McGaw, B. (2012). The Changing Role of Education and Schools. In P. Griffin, B. McGaw, E. Care (eds): *Assessment and Teaching of 21st Century Skills. Research in Science Teaching* (pp. 1–15). New York: Springer. https://www.doi.org/10.1007/978-94-007-2324-5_1
- Jumariati, Febriyanti, E. R., Rizki, M. (2021). Argumentation Skills: An Analysis on EFL Students' Essays Based on Toulmin's Model of Argument. In: *Advances in Social Science, Education and Humanities Research, Proceedings of the 2nd International Conference on Education, Language, Literature, and Arts (ICELLA 2021)*, 587 (pp 86–95). <https://www.doi.org/10.2991/assehr.k.211021.013>
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Education Researcher*, 23(2), 13–23. <https://www.doi.org/10.3102/0013189x023002013>
- Mislevy, R. J., Almond, R. G., Lukas, J. F. (2003). A Brief Introduction to Evidence-centered Design. *Educational Testing Service*, 2003(1), i–29. <https://www.doi.org/10.1002/J.2333-8504.2003.TB01908.X>
- Mislevy, R. J., Haertel, G., Riconscente, M. et al. (2017). *Assessing Model-Based Reasoning using Evidence-Centered Design: A Suite of Research-Based Design Patterns*. New York: Springer. <https://www.doi.org/10.1007/978-3-319-52246-3>
- Mislevy, R. J., Levy, R. (2007). Bayesian psychometric modeling from an evidence-centered design perspective. *Handbook of statistics*, 26, 839–865. [https://doi.org/10.1016/S0169-7161\(06\)26026-7](https://doi.org/10.1016/S0169-7161(06)26026-7)
- Nefedov, S. N. (2017). Toulmin's proof model and quantitative indicators of the reliability of the conclusion. *Issues of criminalistics and forensics*, 1(41), 87–93. (In Russ.).
- Oliveri, M. E., Mislevy, R. J. (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), 97–102. <https://www.doi.org/10.1080/15305058.2019.1608551>
- Toulmin, S. E. (2003). *The Uses of Argument*. Cambridge: Cambridge University Press.
- Uglanova, I. L., Brun, I. V., Vasin, G. M. (2018). Evidence-Centered Design methodology for measuring complex psychological constructs. *Modern foreign psychology*, 7(3), 18–27. <https://www.doi.org/10.17759/jmfp.2018070302> (In Russ.).
- Weideman, A. (2012). Validation and Validity Beyond Messick. *Per Linguam*, 28(2), 1–14. <https://doi.org/10.5785/28-2-526>
- Yefremova, N. F. (2018a). Evidence centred design model in the design of measuring instruments. *Pedagogical dimensions*, 3, 8–15. (In Russ.).

Об авторе:

Ефремова Надежда Федоровна, доктор педагогических наук, профессор, заведующий кафедрой «Педагогические измерения», Донской государственный технический университет (РФ, 344003, г. Ростов-на-Дону, пл. Гагарина, 1), [ORCID](https://orcid.org/0000-0001-9151-1111), nefremova61@donstu.ru

About the Author:

Efremova Nadezhda F., Dr. Sci. (Pedagogy), Professor, Head of the Department of Pedagogical Measurements, Don State Technical University (1, Gagarin Sq., Rostov-on-Don, 344003, RF), [ORCID](#), nefremova61@donstu.ru

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Автор прочитал и одобрил окончательный вариант рукописи.