

МОДЕЛИРОВАНИЕ ОБЪЕКТИВНОЙ ОЦЕНКИ ДОСТИЖЕНИЙ СТУДЕНТОВ¹

Ефремова Надежда Фёдоровна,

доктор педагогических наук, профессор, заведующая кафедрой педагогических измерений, Донской государственной технической университет, г. Ростов-на-Дону, e-mail: nefremova61@donstu.ru

В статье рассмотрены особенности проектирования оценочных средств по модели Evidence Centered Design (ECD). Применение принципов доказательной аргументации с использованием паттерн-дизайнов позволяет шире использовать методы классификации, упорядочения и планирования оценки когнитивных и компетентностных результатов освоения студентами основных профессиональных образовательных программ. Использование моделей ECD способствует разработке оценочных средств, обеспечивающих надёжность результатов оценки в условиях компетентностного подхода, что, в свою очередь, способствует повышению учебной мотивации студентов.

Ключевые слова: компетенции, качество обучения, паттерн проектирования, моделирование оценки, результаты обучения, доказательный дизайн, надёжность оценки, учебная мотивация.

Во все времена в образовании существовала проблема надёжности оценивания результатов обучения. Эта проблема в настоящее время обострилась в связи с необходимостью оценивать не только степень освоения содержания учебных дисциплин, но и уровни сформированности компетенций как результатов, показывающих способности студентов и выпускников действовать и решать нетрадиционные задачи с использованием как известных и усвоенных знаний, так и новых, полученных из информационных источников. Однако при гигантском технологическом сдвиге в общественном секторе пока не наблюдается такого же сдвига ни в методологии компетентностного обучения, ни в технологиях оценивания его результатов [1]. Несмотря на уже достаточно длительный период перехода к компетентностному обучению, в системе высшего образования доминирует дисциплинарно-когнитивная

¹ Исследование выполнено при поддержке гранта программы ЭРАЗМУС+575456-EPP-1-2016-1-RU-EPPJMO-CHAIR



структура образовательной деятельности, но в то же время активно используется компетентностная риторика, а традиционное оценивание усвоенного содержания учебных дисциплин выдаётся за оценку компетенций. В первую очередь это связано с непониманием особенностей компетенций, имеющих сложную междисциплинарную, глубоко латентную природу и деятельностный характер, что затрудняет наблюдения за их проявлением, потому что формирование и оценивание компетенций возможны только в условиях включения студентов в активную самостоятельную деятельность, а оценочные средства должны иметь свои специфические особенности.

Целью данной статьи является изложение метода Evidence Centered Design [2] (ECD), обеспечивающего возможности доказательной аргументации оценок, заложенной при проектировании педагогических измерителей на основе паттерн-дизайна, шаблона многократного использования, используемого для конструирования оценки. Главное достоинство методики состоит в обеспечении относительной полезности, валидности и надёжности оценочного средства. Понимание принципов доказательной аргументации и использование шаблонов позволяет преподавателям эффективнее применять методы классификации, упорядочения и планирования оценки когнитивных и компетентностных

результатов освоения студентами основных профессиональных образовательных программ. Одна из ключевых идей дизайна оценки, ориентированной на фактические данные (ECD), заключается в том, что функциями задачи можно манипулировать для изменения психометрических свойств оценочного средства, определять переменные функции-задачи, включая определение фокуса доказательств, формы заданий, сложности входящих в задание задач и их дискриминативные возможности. Разработчики оценок могут использовать эти функции задач для управления психометрическими свойствами как традиционных форматов оценок, так и сложных компетентностно-ориентированных задач, встроенных в комплексное задание.

В данной статье достаточно схематично обрисованы основные моменты и преимущества методики ECD по разработке паттернов. Отметим, что паттерн обеспечивает единообразие структурного признака в пределах одного иерархического уровня оценочного средства. А достижение надёжности возможно, если вся конструкция задания без потерь охвачена полностью совокупностью её элементов.

Новая методология оценки должна предложить широкий набор инструментов для оценивания работы студентов в группах и индивидуально; способностей самооценки и оценки достижений других, рефлексии достижений; умений

принятия рисков и поиска нестандартных решений задач; других важных характеристик согласно заданной модели будущего специалиста.

Компетентностный подход к построению и реализации основных профессиональных образовательных программ поставил перед преподавателями целый ряд вопросов: какой комплекс знаний, навыков или других атрибутов следует формировать и оценивать? какие оценочные средства использовать для выявления латентных (когнитивных, так называемых жёстких) и глубоко латентных (компетентностных, «мягких») результатов обучения? какие задачи или ситуации, действия или выступления могут вызывать такое поведение испытуемых, чтобы результат обучения стал явным, однозначным и измеримым? как аргументированно доказать, что результаты оценивания достижений обучающихся действительно отображают достигнутый уровень планируемых к освоению компетенций? С каждым годом спрос на доказательства достоверности и надёжности оценок компетенций выпускников вузов становится всё более актуальным, а получение ответов на эти вопросы остаётся неопределённым. Получение ответов на данные вопросы — очень сложная задача. Понимание путей решения указанных проблем может обеспечить разработчикам оценочных средств выбор контролируемых

рубрик и необходимых ситуаций для организации деятельности студентов, обоснование критериев и шкал, отображающих действительный уровень подготовленности испытуемых путём использования доказательств и аргументаций при проявлении трудноизмеримых латентных характеристик в ситуации оценки. Всё это показывает, что сегодня требуется разработка новых технологий оценки, соответствующих компетентностному обучению.

Несмотря на важность получения объективных результатов обучения, в системе высшего образования всё ещё доминирует традиционная практика оценивания учебных достижений, которая, как правило, не опирается на научно обоснованный подход и должное методическое обеспечение оценочной деятельности; практически повсеместно используются неапробированные и нестандартизированные оценочные материалы, так называемые авторские. Результаты, полученные таким способом, не являются предметом содержательного анализа индивидуальных достижений, не вызывают доверия и не обеспечивают положительного воздействия на учебную мотивацию студентов. Отметим, что оценка «мягких» результатов обучения из-за их глубоко латентной природы более сложна, чем оценка «жёстких», для которых процессы оценки могут быть значительно сложнее. Тем не менее основная контрольно-оценочная функция должна иметь дальнейшее



совершенствование и уточнение в рамках исследований и разработок, направленных на демонстрацию того, что процессы, лежащие в основе оценки как когнитивных, так и компетентностных результатов обучения, требуют более глубокого понимания.

Что касается когнитивно-ориентированного получения данных, то это в основном сводится к оценке степени усвоения содержания учебных дисциплин и традиционному дизайну оценочных средств.

В становлении когнитивной диагностики в середине прошлого века большую роль сыграли две теории тестов: классическая и современная IRT (англ. Item Response Theory) [3, 4]. Обе они оказались особо востребованными при проведении крупномасштабных оценочных процедур с так называемыми высокими ставками. Это, прежде всего, итоговая государственная аттестация выпускников основной школы и квалификационный отбор специалистов. Однако как традиционные, так и указанные выше методы конструирования и параметризации тестов IRT становятся недостаточными для выявления уровней сформированности компетенций. Оценка компетенций требует организации деятельности студентов и их привлечения к решению комплексных задач, взаимодействия с другими обучающимися и специалистами. Поэтому дизайн компетентностно-ориентированных оценок должен быть нацелен как

на выявление степени освоения учебных дисциплин, так и на взаимосвязь междисциплинарных знаний и стратегии решения проблем, обеспечивая наблюдения и доказательства правильности ответов и эффективности действий студентов при выполнении заданий. Решение проблемы оценивания достижений обучающихся заключается в поисках методов разработки надёжных оценочных средств, обеспечивающих активную деятельность испытуемых в процессе оценки и повышение интереса при выполнении заданий.

В зарубежной образовательной практике одним из таких методов стал предложенный Р. Мислеви (Mislevy R., 2003) [5, 6] метод доказательной аргументации Evidence-Centered Design (ECD). Он опирается на конструкцию задания, описанную С. Мессиком (Messick S., 1994) [7], и идеи структуризации аргументов оценки в виде универсального паттерна (множественно проверенного шаблона) С. Тулмина (Toulmin S., 1958) [8]. Согласно С. Мессику, именно природа конструкта обуславливает создание релевантных заданий, а также разработку основанной на нём системы подсчёта баллов и оценочных рубрик. Данные становятся свидетельством подготовленности испытуемых и условием повышения мотивации, только когда их значение для выводов однозначно установлено. Поэтому чем лучше задания отображают цели проверки, тем эффективнее процесс

и результат в контексте оценки. При этом следует учитывать, что одни и те же данные могут быть хорошим доказательством для одного вывода, но плохими для доказательства другого (Schum D., 1994) [9].

Таким образом, ECD — это, по сути, проектирование разработки комплексных оценочных средств на основе шаблонов, допускающих их многократное использование в подобных ситуациях, чтобы при отсутствии у преподавателей знаний и опыта в области педагогических измерений облегчить им задачи конструирования педагогического инструментария и достижения надёжности оценивания когнитивных и компетентностных результатов. Такой метод создания шаблонов для конструирования средств оценки представляет собой доказательный дизайн и аргументацию фактических результатов испытуемых. Его алгоритм обеспечивается на основе сопряжения модели студента (его

знаний и компетенций) и модели значимого задания с набором доказательств и аргументаций наблюдаемых результатов деятельности испытуемого при выполнении задания. Использование доказательного дизайна ECD создаёт надёжную основу для установления связи между оценкой и корректировкой образовательной деятельности.

В этом процессе важную роль играет концептуальная рамка оценивания (КРО) (рис. 1), которая обеспечивает сопряжение модели обучающегося, модели свидетельств и модели задания, а также задаёт принципы перехода от баллов на уровне шкалах к выводам об освоении знаний и сформированности компетенций [10, 11].

К основным элементам КРО (модели сборки оценочного средства) в ECD относят: модель студента, модель свидетельств, модель задания, статистическую модель и модель презентации данных.



▲ Рис. 1. Концептуальная рамка оценивания



Модель студента описывает конструкт исследуемых характеристик: набор латентных переменных, подлежащих оценке. Модель задания отражает структуру задания в целом (спецификацию и план) и аргументов оценки. Подбираются интересные ситуации и выстраиваются комплексные задания из набора задач; проектируются виды деятельности при выполнении задания; планируются причинно-следственные связи между ситуацией и деятельностью студента в условиях оценки; определяется, что и как должно быть выявлено при решении задач, как представить ответы на задание в целом. Связь оцениваемых аргументов с функциями задания обеспечивает проявление знаний и компетенций при выполнении запланированной деятельности. В рамках структуры оценки, ориентированной на доказательство, используются такие задачи, как решение проблем, поиск информации, анализ, общение и др. Поэтому всё задание должно быть дополнительно расчленено на целевые объекты для оценки, каждой задаче соответствуют свои критерии выполнения и шкала оценивания.

Модель свидетельств является важным оценочным компонентом задания: она задаёт общий балл по всему заданию, баллы за выполнение отдельных задач, опровержения и доказательства оценки продуктов и видов деятельности, сбор свидетельств и аргументов выполнения задания, правила оценки про-

цессов или продуктов деятельности, обоснование шкалы измерения. Статистическая модель задаёт правила обработки результатов и их отображения на шкалу оценивания. Модель презентации определяет форматы представления окончательных оценок пользователям (студентам, преподавателям и стейкхолдерам) [12, 13].

Алгоритм Evidence-Centered Design воплощает все основные процессы и проектные решения для выделения оцениваемых характеристик испытуемого: спецификацию задачи, элементы заданий и аргументацию качества их выполнения, свидетельства для обоснования оценки, обоснование шкалы и форматы представления выходных данных. В новой модели оценки применимы шаблоны проектирования оцениваемых достижений студента, психометрические модели заданий с наборами различных задач и процессов их выполнения для надёжной аргументации. Такой шаблон может применяться при разработке любой оценки, где априорное определение конструкций оценки и связанные с ними переменные имеют смысл. Все этапы разработки оценочного средства от операционализации конструкта оцениваемых рубрик и построения концептуальной рамки оценивания до подготовки задания и формулирования условий его выполнения должны давать свидетельства того, что измеряется и оценивается действительно то, что



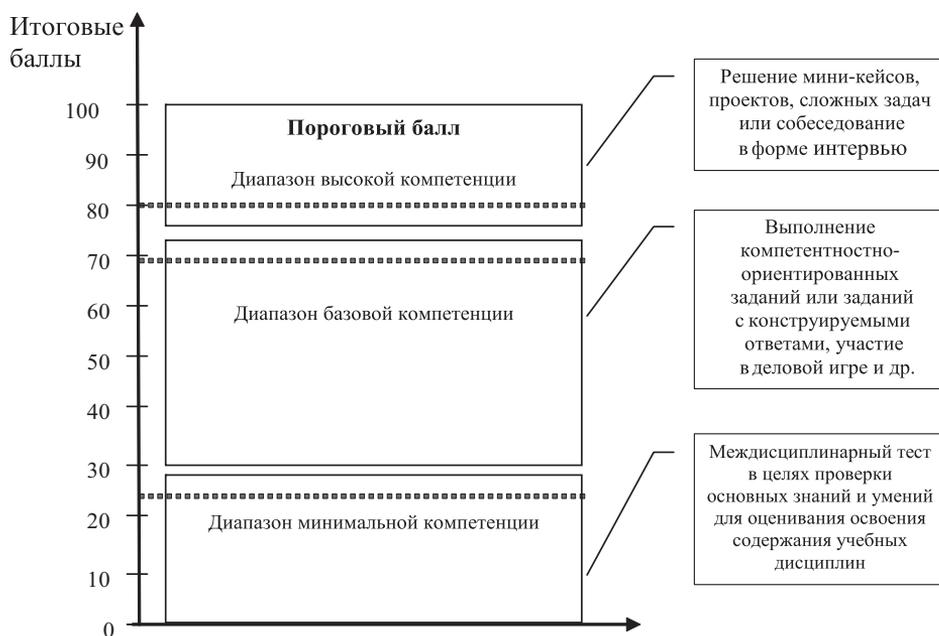
затем эти оценки накапливают, чтобы сделать окончательные оценочные суждения о знаниях и компетенциях на уровневой шкале (рис. 3).

Комплексный подход к оценке обеспечивает надёжность результатов, доверие студентов к оценке и повышение их мотивации в приобретении новых знаний. Связь между процессами разработки системы оценки и включение планируемой деятельности для выполнения задания создают условия, в которых студенты могут показать, что и как они могут делать, используя полученные знания и находя новые, необходимые для решения задания.

Дизайн, основанный на доказа-

тельствах, — это новый способ разработки высококачественных задач и оценок. Акцент делается на обеспечении того, чтобы каждый элемент и задание содержал свидетельство о цели оценки, которая может быть использована для подтверждения заявления о сформированности знаний, навыков и (или) способностей, требуемых стандартом.

Разработка средств оценки является трудоёмким процессом, именно поэтому существует потребность многократного использования удачных шаблонов дизайна оценочного средства, связанных с конкретными наборами утверждений и являющихся мостом между моделью обучающегося и моделью конструируемого задания. Такой подход может



▲ Рис. 3. Уровеньная шкала оценивания компетенций по итоговому баллу

успешно применяться при оценке как знаний студентов, так и компетенций, а процедура оценки сама является обучающим, развивающим и мотивирующим компонентом в образовательном процессе. Связь между этапами разработки оценки с обеспечением деятельности испытуемых может способствовать созданию нового типа оценочных средств и дальнейшему развитию аутентичного оценивания образовательных достижений студентов.

Литература

1. *Бояцис Р.* Компетентный менеджер. Модель эффективной работы / пер с англ. — М.: ГИППО, 2008. — 352 с.
2. *Mislevy R.J., Almond R.G., Lukas J.F.* A brief introduction to evidence centered design // ETS Research Report Series. 2003(1).
3. *Rasch G.* Probabilistic Model for Some Intelligence and Attainment Tests. With a Foreword and Afteword by B.D. Wright. — Chicago & London: The Univ. of Chicago Press, 1980. — 199 p.
4. *Lord F.M.* Application of Item Response Theory to Practical Testing Problems. Hillsdale. — N.Y.: Lawtrence Erlbaum Ass. Publ, 1980. — 266 p.
5. *Mislevy R.J., Almond R.G., Lukas J.F.* A brief introduction to Evidence-Centered Design. (ETS Research Report RR-03-16). — Princeton NJ: Educational Testing Service, 2003. — 37 p.
6. *Mislevy R.J., Behrens J.T., Dicerbo K.E., Levy R.* Design and discovery in educational assessment: Evidence-Centered Design, psychometrics, and data mining. J Educ Data Min. 2012; 4(1):11-48.
7. *Messick S.* The interplay of evidence and consequences in the validation of performance assessments. Education Researcher, 1994;32(2):13-23.
8. *Toulmin S.E.* The Uses of Argument. — UK: Cambridge University Press, 1958. — 259 p.
9. *Schum D.A.* Book review: Evidential foundations of probabilistic reasoning. — N.Y.: Wiley, 1994. — 545 p.
10. Assessing Model-Based Reasoning using Evidence-Centered Design: A Suite of Research-Based Design Patterns by R.J.Mislevy, G. Haertel, M. Riconscente, D. Rutstein, C. Ziker. — Springer, 2017. — 130 p.
11. *Mislevy R.J., Levy R.* Bayesian psychometric modeling from an evidence-centered design perspective // Rao C.R., Sinharay S. eds. Handbook of statistics. — Amsterdam: Elsevier, 2007. — P. 839-865.
12. *Ефремова Н.Ф.* Проектирование оценочных средств по модели доказательной аргументации // Педагогические измерения. — 2018. — № 3. — С. 8-15.
13. *Ефремова Н.Ф.* Patterns of Designing Evaluation of Students Competencies. American Journal of Education. 2018;4(2), (August). Vol. 124. The University of Chicago Press. — Pp. 1048-1065.
14. *Ефремова Н.Ф.* Reliability Assessments Competences of Students by means of Evidential Reasoning // Problems of the development of modern science: theory and practice. — Madrid, Spain: Cartero Publishing House, 2018. — Pp. 194-197.
15. *Звонников В.И., Малыгин А.А., Чельшкова М.Б.* Оценивание в высшем образовании: от линейности к адаптивности // Известия вузов. Серия «Гуманитарные науки». — 2014. — 5(2). — С. 166-171.