

МЕТОД EVIDENCE-CENTERED DESIGN

Независимая оценка компетенций в цифровом мире

Наталья Родина, Алина Постовалова, Александр Долматов

АННОТАЦИЯ В статье рассмотрены основные понятия, этапы и примеры использования метода доказательной аргументации Evidence-Centered Design (далее — ECD) для разработки инструментов оценки компетенций. Актуальность применения метода ECD обусловлена потребностью в развитии у граждан компетенций цифровой экономики и необходимостью наличия у них минимального уровня цифровой грамотности. Компетенции цифровой экономики — сложносоставные и внепредметные. Требуются специальные методы для их оценки. По опыту зарубежных стран, признанным методом оценки таких компетенций является Evidence-Centered Design, или ECD. В статье рассмотрен опыт зарубежных и российских специалистов в применении метода ECD для оценки различных компетенций, отдельно выделены примеры использования метода в рамках игровой оценки (оценки компетенций и усвоения знаний посредством компьютерных игр).

КЛЮЧЕВЫЕ СЛОВА Независимая оценка компетенций, компетенции цифровой экономики, цифровые компетенции, метод Evidence-Centered Design (ECD), метод доказательной аргументации.

DOI 10.22394/2078–838X-2021–3-52-59

Актуальность и значимость оценки

Происходящие в последние десятилетия изменения — массовое использование цифровых инструментов, растущая ценность образования, изменение содержания разных профессий — обуславливают актуальность введения независимой оценки компетенций, которая бы позволяла гражданам планировать свои образовательные и трудовые траектории, а организациям — снижать затраты на поиск необходимых работников.

Фактически оценка компетенций человека осуществляется на разных этапах его жизни. Различные организации (в зависимости от целей своей деятельности) создают собственные подходы к оценке компетенций. Так, в высшем образовании имеет место переход на компетентностный подход, при котором результаты освоения студентами образовательных программ оцениваются по компетенциям, на формирование которых нацелена программа. Крупные государственные и коммерческие организации (например, «Сбер», РЖД) создают целевые модели компетенций сотрудников, служащие ориентиром при найме

соискателей, а также при развитии имеющих-ся кадров. В ряде отраслей существуют профессиональные сертификации, признаваемые при найме и продвижении сотрудников по службе (например, в сфере финансов — CIMA, ACCA, CFA и CIA, сертификаты для дизайнеров, подтверждающие владение программами пакета Adobe (Adobe Certified Associate (ASA)), сертификация Microsoft Technology Associate (MTA) для специалистов в области информационных технологий). Зарубежные компании (имеющие представительства и в России) проводят сертификацию навыков и компетенций, особо востребованных в современном мире (так называемых «навыков XXI века») — критического мышления, коммуникаций и пр.¹ Все эти примеры говорят о существовании различных внутрикорпоративных, внутриотраслевых подходов, но не о единой общенациональной системе независимой оценки компетенций.

Примером создания единых рамок оценки компетенций являются общенациональные

¹ Например, компания Certiport оценивает коммуникационные навыки — сертификат Communication Skills for Business (CSB).

Наталья Викторовна
РОДИНА

аналитик, к. э. н., «Университет Национальной технологической инициативы 2035»
(121205, РФ, Москва, территория инновационного центра «Сколково», ул. Нобеля, д. 1).
E-mail: n.rodina@2035.university

Алина Анатольевна
ПОСТОВАЛОВА

«Университет Национальной технологической инициативы 2035», руководитель направления независимой оценки компетенций
(121205, РФ, Москва, территория инновационного центра «Сколково», ул. Нобеля, 1).
E-mail: a.postovalova@2035.university

Александр Владимирович
ДОЛМАТОВ

«Университет Национальной технологической инициативы 2035», руководитель отдела внедрения независимой оценки компетенций
(121205, РФ, Москва, территория инновационного центра «Сколково», ул. Нобеля, 1).
E-mail: a.dolmatov@2035.university

проекты оценки цифровой грамотности. Цифровизация экономики требует, чтобы граждане обладали минимально необходимыми цифровыми компетенциями, были готовы пользоваться общественными услугами в цифровом виде и отвечать на вызовы рынка труда. В европейском сообществе разработаны рамки цифровых компетенций, признаваемые разными странами.² В России эти подходы также используются для оценки уровня цифровой грамотности населения в целом и отдельных профессиональных групп.³

Среди последних мировых тенденций в оценке компетенций — использование для этих целей компьютерных игр и симуляторов. Хотя в нашей стране применение таких инструментов не распространено⁴, в зарубежной литературе набирает обороты дискуссия о возможностях их использования для оценки компетенций взрослых и детей, а также в целях обучения (Dichev & Dicheva, 2017).

Перечисленные примеры говорят о широком применении оценки компетенций в

современном мире и о будущем спросе на нее. Для того чтобы результаты оценки признавались различными участниками рынка труда и оцениваемыми людьми, должны использоваться валидные инструменты, а процесс оценки должен быть открытым и понятным участникам.

Современная научная и экспертная литература представлена многими публикациями о методах оценки навыков и компетенций, таких как тесты, кейс-тесты, ассессмент-центры и пр. Однако русскоязычных работ о методах оценки сложносоставных компетенций (к которым относятся цифровые компетенции и «навыки XXI века» — цифровая грамотность, критическое мышление, креативное мышление и др.) немного.

В данной статье на основании синтеза работ, опубликованных на русском и английском языках, рассматриваются возможности метода доказательной аргументации Evidence-Centered Design (далее — ECD) для разработки инструментов оценки сложносоставных компетенций, в том числе цифровых компетенций. Метод ECD охарактеризован кратко — об этом есть хорошие публикации на русском языке (Авдеева и др., 2017; Углова и др., 2018), основное внимание уделяется примерам его использования.

Метод ECD и примеры его использования

Метод ECD был разработан в компании Educational Testing Service (ETS; www.ets.org)

² Европейской Комиссией разработана рамка цифровых компетенций для граждан DigComp, а также ее обновленные версии DigComp 2.0 и DigComp 2.1. Для отдельных целевых групп Европейской Комиссией разработаны рамки для педагогов (DigCompEdu), для потребителей (DigCompConsumers), для высших учебных заведений (Opened), для предпринимательства (EntreComp). Рамки цифровых компетенций имеются в Канаде (British Columbia’s Digital Literacy Framework), Чили (ICT SIMCE competences), Коста-Рике (Student performance standards in learning with digital technologies), Венгрии (Digital Education Strategy), Индии (National Digital Literacy Mission).

³ Примеры исследований цифровой грамотности работников отраслей: учителей, преподавателей высшего и среднего профессионального образования (НАФИ, портал «Цифровой гражданин Югры», интернет-издание «Профобразование», платформа «Фоксфорд», «Яндекс»), работников музеев (Российский комитет Международного совета музеев (ИКОМ России) совместно с Ассоциацией электронных коммуникаций (РАЭК), Национальным исследовательским университетом «Высшая школа экономики» (НИУ ВШЭ) и Microsoft в России при информационной поддержке проекта PRO.КУЛЬТУРА.РФ).

⁴ Среди российских компаний, использующих игровую оценку компетенций, Школа предпринимательства «ТехноСпарка» (тренажер «Построй компанию. Продай компанию»), Oblako Group.

в начале 2000-х гг.; за рубежом используется в таких тестологических компаниях, как Cisco, GlassLab, CRESST, Labs, CITO (Mislevy et al., 2003). Метод наиболее актуален для измерения комплексных, многокомпонентных конструкторов (Авдеева и др., 2017; Угланова и др., 2018)⁵. Для оценки таких конструкторов в фокус оценщика попадают не только результаты решения тестовых заданий, но и сам способ решения.

Основная цель ECD — обосновать дизайн теста, процесс его разработки, шкалирования и использования с помощью принципов формальной логики. Измерение рассматривается как процесс аргументации какого-либо вывода о респондентах на основе свидетельств о том, что респонденты делают/говорят в ситуации тестирования. Центральное понятие здесь — суждение (вывод; claim) о респонденте, о том, что он знает или умеет делать в реальной жизни. Весь процесс разработки теста направлен на то, чтобы собрать как можно больше доказательств того, что вывод, сделанный о респонденте на основе его тестового балла, отражает действительность (Угланова и др., 2018).

Разработка оценочного инструмента в рамках этого подхода происходит «послойно»: она начинается с набора очень общих идей об оцениваемой компетенции (с определений и составляющих, с построения модели компетенции) и заканчивается спецификацией теста (конкретными правилами его разработки). Авторы метода (Mislevy et al., 2003) выделяют 5 этапов (слоев, layers)⁶ его применения. Мы рассмотрим три основных понятия (которые позволяют сформировать представление о методе и описать примеры его использования), это: модель компетенции (Competency Model, в работах разных авторов используются смежные понятия — Student model, Proficiency Model), модель доказательств (Evidence Model), модель заданий (Task Model). В совокупности

5 Специалисты-психометрики используют термин «конструкт» для обозначения явно не наблюдаемых, латентных психологических черт или характеристик, компетенций или навыков, которые недоступны непосредственному наблюдению, но могут быть оценены через релевантное наблюдаемое поведение. Многомерный конструкт — такой конструкт, составляющие которого являются взаимосвязанными (Угланова и др., 2018).

6 1) анализ области компетенций (Domain Analysis); 2) моделирование области компетенций (Domain Modeling); 3) разработка концептуальной рамки инструмента (Conceptual Assessment Framework — CAF); 4) сборка инструмента (Assembly Model); 5) доставка инструмента (Delivery Model).

перечисленные модели включаются в концептуальную рамку инструмента оценки (Conceptual Assessment Framework — CAF).

Модель компетенции описывает, какая совокупность знаний, умений и/или компетенций оценивается в рамках используемого инструмента, включает формализацию уровней развития компетенций. *Модель доказательств* отвечает на вопросы: как мы измеряем компетенции, какое поведение или действия тестируемого выявляют различные уровни владения оцениваемой компетенцией. Она обосновывает, как ответы тестируемого и то, что он делает при решении заданий, доказывают наличие компетенции. *Модель заданий* описывает, какие задачи должны вызывать те действия и поведение оцениваемых, которые доказывают наличие компетенции. Основная цель модели заданий — получить доказательства (которые наблюдаемы) о компетенциях (которые непосредственно не наблюдаемы).

В совокупности эти модели позволяют определить: (1) знания, навыки и/или компетенции, которые должны быть измерены; (2) задачи, которые позволяют обнаружить у тестируемых эти знания, навыки, компетенции; (3) данные и доказательства, получаемые из решения этих задач, и то, как интерпретировать их, чтобы сделать выводы о компетенциях оцениваемого.

Как показывает анализ литературы, в России не много авторов, имеющих опыт применения ECD при разработке оценочных инструментов. Данный метод используется авторами из НИУ «Высшая школа экономики» для разработки инструментов измерения креативности и критического мышления (Угланова и др., 2018), из «Национального фонда подготовки кадров» для оценки информационно-компьютерной компетентности (Авдеева и др., 2017; Авдеева и др., 2009; Худенко, Авдеева, 2014). В своих статьях Ефремова Н. Ф. (Донской государственный технический университет) обосновывает перспективы использования метода ECD для оценки компетенций студентов высших учебных заведений (Ефремова, 2018; Ефремова, 2019а; Ефремова, 2019b).

За рубежом опыт использования ECD гораздо богаче. Компания Educational Testing Service использовала ECD на протяжении последних 10-15 лет при проектировании инструментов оценки. Об опыте применения данного метода свидетельствует внушительный перечень отчетов, размещенных в открытом доступе на сайте ETS⁷.

БОЛЬШИНСТВО АВТОРОВ, ПРИМЕНЯЮЩИХ ECD, ИСПОЛЬЗУЮТ ДЛЯ ПОДСЧЕТА БАЛЛОВ ТЕСТИРУЕМЫХ СЕТИ БАЙЕСА

Метод ECD — основополагающий при разработке тестов, проводимых Cisco Networking Academy. Департаментом образования штата Нью-Йорк метод ECD применялся для разработки государственных тестов, которые в обязательном порядке сдают учащиеся общеобразовательных школ (например, для разработки государственного теста по всеобщей истории и географии (University of the State of New York, 2019)). Метод также использовался при разработке компьютерных игр, служащих для обучения и оценки компетенций.

Отчеты и статьи зарубежных проектировщиков тестов о применении метода ECD включают описания того, как разработчики прошли основные логические этапы разработки теста: что было включено в компетентностную модель оцениваемого (какие компетенции и какие между ними существуют взаимосвязи), как выбирались задания теста и какова их связь с оцениваемыми компетенциями, какие данные собираются в ходе применения оценочного инструмента для получения вывода об уровне развития компетенций оцениваемого.

Обеспечение прозрачности оценки

К ключевым компетенциям цифровой экономики, в соответствии с Приказом Минэкономразвития от 24 января 2020 г. № 41 «Об утверждении методик расчета показателей федерального проекта “Кадры для цифровой экономики” национальной программы “Цифровая экономика Российской Федерации”», относятся: коммуникация и кооперация

в цифровой среде, саморазвитие в условиях неопределенности, креативное мышление, управление информацией и данными, критическое мышление в цифровой среде. Все эти компетенции представляют собой совокупность паттернов мышления, поведения и реакций, которые формируются и проявляются на самых разных уровнях, они внепредметные (их формирование и проявление не ограничивается какой-то одной дисциплиной или сферой деятельности).

Как пишут зарубежные авторы (Behrens, Mislevy, et al., 2010; Mislevy et al., 2003), использование ECD при создании инструментов оценки цифровых компетенций позволяет обеспечить валидность и прозрачность оценки и, следовательно, сформировать доверие к ней пользователей. Валидность инструмента обеспечивается за счет обоснованности связей между моделью компетенции, моделью доказательств и моделью заданий. Прозрачность — за счет понятных схем скоринга (подсчета итоговых баллов).

Michael J. Zieky (2014), специалист компании ETS, указывает на то, что методология ECD не требует применения какого-либо одного подхода к скорингу результатов тестируемых. Выбор такого подхода определяется уровнем сложности оцениваемой компетенции и компетентностной моделью. Так, по мнению M. J. Zieky, инструмент, нацеленный на оценку одной компетенции и имеющий модель сбора доказательств, в которой каждая задача напрямую связана с этой единственной компетенцией, может включать оценку в виде подсчета количества правильных ответов и суммарного итогового балла. В то же время есть более сложные диагностические задачи, нацеленные на определение развития нескольких компетенций и предполагающие сложные модели сбора свидетельств, в которых существует несколько связей между задачами и различными знаниями, умениями, навыками. Такие инструменты предполагают использование более сложных подходов скоринга, например, сетей Байеса.

Отметим, что проведенный нами анализ литературы показывает: большинство авторов, применяющих ECD, используют для подсчета баллов тестируемых сети Байеса.

Байесовские сети представляют собой инструмент для описания достаточно сложных процессов и событий с неопределенностями, удобство которого заключается в графической

7 Результаты поиска отчетов на сайте компании ETS по запросу «Evidence-Centered Design» можно посмотреть по следующей ссылке: <http://search.ets.org/researcher/query.html?col=rsrchr&q1=a&fl0=AU%3A&ty0=w&op0=&tx0=+&fl1=TI%3A&ty1=w&op1=&tx1=+Evidence-Centered&fl2=RN%3A&ty2=w&op2=&tx2=+&fl3=PY%3A&ty3=w&op3=&tx3=+&fl4=SN%3A&ty4=w&op4=&tx4=+&q5=KW%3A&ty5=w&op5=&tx5=+&fl6=AB%3A&ty6=w&op6=&tx6=+&q7=+&nh=25&charset=utf-8>

наглядности. Основной идеей построения сети является разложение сложной системы на простые элементы. Для объединения отдельных элементов в систему используется математический аппарат теории вероятностей, позволяющий количественно описывать силу причинно-следственных отношений между элементами системы (Хлопотов, 2014).

Байесовская сеть — это ациклический ориентированный граф, в котором каждая вершина (узел сети) представляет переменную, дуги обозначают существование причинно-следственных зависимостей между соединенными переменными, а сила этих зависимостей количественно выражается в виде условных вероятностей (это вероятность события при условии, что произошло другое событие).

Использование байесовской сети в инструментах оценки цифровых компетенций предполагает, что в сети определены наблюдаемые переменные (отражающие владение компетенцией), направления взаимоотношений между наблюдаемыми переменными, заданиями и оцениваемыми компетенциями, а также сила этих отношений. Сила отношений определяется эмпирическим путем (при наличии соответствующих исследований) или методом экспертного опроса.

Изменение полученного балла по одной из наблюдаемых переменных в сети Байеса приводит к обновлению суммарного балла по оцениваемой компетенции в целом.

Таким образом, при использовании сетей Байеса в оценочных средствах для измерения компетенций отчеты о полученных тестируемыми баллах включают информацию о связи результатов выполнения заданий с выводами об уровне владения компетенцией. Такой подход позволяет достигать прозрачности оценки.

Игровая оценка компетенций

В зарубежных публикациях можно найти примеры использования метода ECD для разработки игровых инструментов оценки (в формате компьютерной игры). Например, Shute V. J., Masduki I., Donmez O. в своей статье (2010) рассматривают пример существующей игры — Quest Atlantis: Taiga Park — и демонстрируют, как собираются и интерпретируются доказательства в отношении такой компетенции, как навык системного мышления.

Во-первых, авторы на основании работ других ученых и исследователей определяют, что подразумевается под системным мышлением. Они понимают эту компетенцию как: а) способность эффективно действовать в сложных ситуациях, б) способность понимать отношения между элементами в сложно устроенной среде. Во-вторых, авторы оценочного инструмента определяют структуру оцениваемой компетенции. У Shute V. J., Masduki I., Donmez O. модель компетенции «системное мышление» представлена тремя переменными первого уровня:

- 1) определение составляющих системы и проблем в системе,
- 2) моделирование системы,
- 3) тестирование модели системы (симуляция).

Каждая из перечисленных переменных детализируется на переменные второго и третьего уровней.

«Тайга» в игре — это красивый парк, через который протекает река. Парк населен несколькими группами людей, которые используют речные ресурсы и зависят от реки, это: а) мулу (коренные жители), живущие за счет занятий сельским хозяйством; б) лесозаготовительная компания; в) туристическая компания. Сюжетная линия игры фокусируется на снижении популяции рыбы в Таежной реке. Тестируемые помогают выяснить, как решить проблему умирающей рыбы и таким образом спасти парк.

В процессе игры тестируемые выполняют специальное задание: они должны нарисовать схему причинно-следственных связей, которая показывает компоненты системы, уменьшающие популяцию рыбы в реке. Модель доказательств (в терминологии ECD) в игре Quest Atlantis: Taiga Park — это: а) созданные участниками схемы причинно-следственных связей, б) правила оценки представленных участниками работ и в) весовые коэффициенты для определения вклада результатов выполненной работы в оцениваемую компетенцию.

Оценка составленных участниками игры схем причинно-следственных связей происходит автоматически, через сопоставление с эталонной (экспертной) схемой. При обработке данных используются байесовские сети. Вывод об уровне владения компетенцией формируется на основании процента совпадений между двумя схемами — участника и эксперта. Задаются

МЕТОД ECD УСПЕШНО ИСПОЛЬЗУЕТСЯ ЗА РУБЕЖОМ ДЛЯ СОЗДАНИЯ ИНСТРУМЕНТОВ ОЦЕНКИ КОМПЕТЕНЦИЙ

также границы отсечения по уровням владения компетенцией: например, 0–33 % — низкий, 34–66 % — средний, 67–100 % — высокий.

В других работах следующих авторов: 1) Groff J., Clarke-Midura J., Owen V. E., Rosenheck L., Beall M. (2015), 2) Conrad Sh., Clarke-Midura J., Klopfer E. (2014), — описывается применение метода ECD для оценки в игре The Radix Endeavor, созданной в целях повышения интереса к STEM-наукам у старшеклассников.

В игре The Radix Endeavor оценка компетенций является не самоцелью, а скорее промежуточным звеном в процессе обучения старшеклассников. В игре существует семь квестов по биологии и шесть по математике. Каждый квест содержит серию более мелких квестов из определенной области контента (например, системы человеческого тела, геометрия и т. д.). Участникам игры нужно пройти все квесты.

Команда разработчиков игры The Radix Endeavor опиралась на циклическое понимание развития компетенций. Цикл начинается с набора задач, выбранных администратором из большой библиотеки. Участникам (оцениваемым) предлагаются задания для работы. После выполнения задания участник отправляет свои ответы на обработку. После проверки ответов происходит обновление совокупного балла учащегося, и для него выбирается новый набор задач. Цикл может продолжаться до тех пор, пока не будет сформирована значимая целевая компетентностная модель участника игры.

Исходя из таких целей, модель компетенции (в терминологии метода ECD) трансформируется разработчиками игры в целевую модель учащегося (что старшеклассник должен знать, уметь, какими компетенциями владеть после обучения, в том числе после прохождения обучающей игры).

Методологическая рамка ECD разработчиками игры была использована в качестве руководства для проектирования квестов: они составлялись таким образом, чтобы метод сбора

и интерпретации доказательств в игровом процессе соответствовал целям обучения. Чтобы регламентировать процесс получения данных о действиях учащихся, команда Radix использует шаблоны квестов в виде таблиц. В таких таблицах использована терминология ECD: формулировки целей обучения (первый столбец таблицы; в терминологии методологии ECD это модель компетенции), выполняемые учащимся задания (Task Model — второй столбец таблицы), собираемые данные о действиях учащихся (Data Collected — третий столбец таблицы) и варианты продолжения игры в зависимости от правильности выполнения учащимся задания (Interpreting Evidence — четвертый столбец таблицы).

На практике, в игре применение шаблона квеста приводит к итерационному процессу взаимодействия игрока с игрой в зависимости от успешности выполнения участником заданий игры. Применение шаблона происходит таким образом, чтобы команда оценщиков знала: 1) какие данные нужно получить от учащегося, выполняющего задания, 2) понимают ли студенты задание, предметную область, 3) нужно ли в игре давать учащемуся задания на повторение и выдать новые задания.

Заключение

Таким образом, анализ литературы показывает, что метод ECD успешно используется за рубежом для создания инструментов оценки компетенций, в том числе сложносоставных. Метод Evidence Centered Design может успешно применяться для разработки оценочных средств, заложенных в компьютерную игру или симулятор, может служить не только для оценки, но и для образования (как средство промежуточного контроля и база для формирования индивидуального маршрута обучения).

В нашей стране имеются перспективы для применения метода в рамках создаваемой системы независимой оценки компетенций, для оценки компетенций в системе высшего и среднего профессионального, дополнительного образования. Минэкономразвития России определен перечень ключевых компетенций цифровой экономики, который состоит из сложносоставных, внепредметных компетенций, для оценки которых необходимо использование соответствующих методов, например метода ECD.

Применение ECD для оценки цифровых компетенций позволит:

- сформировать понимание состава компетенций цифровой экономики (что к ним относится, какие подкомпетенции они включают и на каких уровнях проявляются);
- разработать валидные инструменты для оценки компетенций цифровой экономики;
- проводить прозрачную оценку владения компетенциями цифровой экономики, позволяющую увидеть взаимосвязи между решением конкретных задач и полученными результатами (результатами) об уровне владения компетенцией.

Литература

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. Получено из <https://vo.hse.ru/data/2017/12/20/1159979685/Avdeeva.pdf>.

Авдеева, С. М., Барышникова, М. Ю., Коваленко, С. К., и др. (2009). Об опыте реализации модели оценки ИКТ-компетентности. *Информатизация образования и науки*, 2, 62–71.

Ефремова, Н. Ф. (2018). Проектирование оценочных средств по модели доказательной аргументации. *Педагогические измерения*, 3. Получено из <https://cyberleninka.ru/article/n/proektirovanie-otse-nochnyh-sredstv-po-modeli-dokazatelnoy-argumentatsii/viewer>.

Ефремова, Н. Ф. (2019а). Концептуальная модель оценки компетенций студентов. *Современные наукоемкие технологии*, 7, 169–174. Получено из <https://www.top-technologies.ru/ru/article/view?id=37607>.

Ефремова, Н. Ф. (2019б). Моделирование объективной оценки достижений студентов. *Образовательные технологии*, 3. Получено из <https://cyberleninka.ru/article/n/modelirovanie-obektivnoy-otsenki-dostizheniy-studentov/viewer>.

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. DOI: 10.17759/jmfp.2018070302.

Углонова, И. Л., Орел, Е. А., Брун, И. В. (2020). Измерение креативности и критического мышления в начальной школе. *Психологический журнал*, 41(6), 96–107. DOI: 10.31857/S020595920011124-2.

Хлопотов, М. В. (2014). Применение байесовской сети при построении моделей обучающихся для оценки уровня сформированности компетенций. *Науковедение*, 5(24). Получено из <https://cyberleninka.ru/article/n/primenenie-bayesovskoy-seti-pri-postroenii-modely-dlya-otsenki-urovnya-sformirovannosti-kompetentsiy/viewer>.

Худенко, Л. А., Авдеева, С. М. (2014). Исследование информационно-коммуникационной компетентности учащихся 9 классов Республики Беларусь: цель и основные результаты. *ВЕСНІК адукацыі*, 1, 3–10.

Получено из https://www.ntf.ru/sites/default/files/%D0%A0%D0%91_%D1%81%D1%82%D0%B0%D1%82%D1%8C%D1%8F_%D0%B8%D0%BA%D0%BA.pdf.

Behrens, J. T., Mislevy, R. J., Di Cerbo, K. E., et al. (2010). *An evidence centered design for learning and assessment in the digital world*. (CRESST Report 778). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://files.eric.ed.gov/fulltext/ED520431.pdf>.

Conrad, Sh., Clarke-Midura, J., & Klopfer, E. (2014). A Framework for Structuring Learning Assessment in a Massively Multiplayer Online Educational Game: Experiment Centered Design. *International Journal of Game-Based Learning*. DOI: 10.4018/IJGBL.2014010103. Retrieved from https://www.researchgate.net/publication/261960598_A_Framework_for_Structuring_Learning_Assessment_in_a_Massively_Multiplayer_Online_Educational_Game_Experiment_Centered_Design.

Dichev, C., & Dicheva, D. (2017). Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *International Journal of Educational Technology in Higher Education*, (14)9. DOI: 10.1186/s41239-017-0042-5.

University of the State of New York, State Education Department. (2019). *Educator Guide to the Regents Examination Global History and Geography II Grade 10*. Retrieved from <http://www.p12.nysed.gov/assessment/ss/hs/framework/ghg2/educator-guide-ghg2-19.pdf>.

Groff, J., Clarke-Midura, J., Owen, V. E., et al. (2015). *Better Learning in Games. A Balanced Design Lens for a New Generation of Learning Games*. Learning Games Network, MIT Education Arcade. Retrieved from <https://education.mit.edu/wp-content/uploads/2018/10/BalancedDesign-Guide2015.pdf>.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). *A brief introduction to evidence-centered design: Research Report*. Princeton, N.J.: Educational Testing Service. Retrieved from <https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2333-8504.2003.tb01908.x>.

Shute, V. J., Masduki, I., & Donmez, O. (2010). Conceptual Framework for Modeling, Assessing and Supporting Competencies within Game Environments. *Tech., Inst., Cognition and Learning*, 8, 137–161.

Zieky, M. J. (2014). An introduction to the use of evidence-centered design in test development. *Psicología Educativa*, 20, 79–87. Retrieved from <https://journals.copmadrid.org/psed/art/j.pse.2014.11.003>.

Possibilities of the Evidence-Centered Design method for independent assessment of competencies in the digital world

Natalia V. RODINA.
Analyst, Candidate of Economic Sciences, University of the National Technological Initiative 2035 (1, Nobelya ul., Russian Federation, 121205, Moscow, Skolkovo Innovation Center).
E-mail: n.rodina@2035.university

Alina A. POSTOVALOVA.
Head of the Independent Competence Assessment Department, University of the National Technological Initiative 2035 (1, Nobelya ul., Russian Federation, 121205, Moscow, Skolkovo Innovation Center).
E-mail: a.postovalova@2035.university

Alexander V. DOLMATOV.
Head of the Independent Competence Assessment Department, University of the National Technological Initiative 2035, (1, Nobelya ul., Russian Federation, 121205, Moscow, Skolkovo Innovation Center).
E-mail: a.dolmatov@2035.university

Abstract

The article discusses the main concepts, stages and examples of using the Evidence-Centered Design (or ECD) method, which is used for the development of competence assessment tools. The relevance of using the ECD method is due to the increasing need for the development of digital economy competencies among citizens and the need for them to have a minimum level of digital literacy. The competencies of the digital economy belong to complex and non-subject competencies, special methods are required for their assessment. According to the experience of foreign authors and experts, the recognized method of evaluating such competencies is the Evidence-Centered Design. The article considers the experience of foreign and Russian specialists in the application of the ECD method for assessing various competencies, separately highlights examples of using the method in the framework of game assessment (assessment of competencies and knowledge acquisition through computer games).

Keywords: independent assessment of competencies, digital economy competencies, digital competencies, Evidence-Centered Design (ECD) method, evidence-based argumentation method.

References

Avdeeva, S. M., Baryshnikova, M. Yu., Kovalenko, S. K., et al. (2009). About the experience of implementing the ICT competence assessment model. *Informatization of Education and Science*, 2, 62–71. (In Russian).

Avdeeva, S. M., Rudnev, M. G., Vasin, G. M., et al. (2017). Assessing Information and Communication Technology Competence of Students: Approaches, Tools, Validity and Reliability of Results. *Educational Studies Moscow*, 4, 104–132. Retrieved from <https://vo.hse.ru/data/2017/12/20/1159979685/Avdeeva.pdf>. (In Russian).

Behrens, J. T., Mislevy, R. J., Di Cerbo, K. E., et al. (2010). *An evidence centered design for learning and assessment in the digital world*. (CRESST Report 778). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://files.eric.ed.gov/fulltext/ED520431.pdf>.

Conrad, Sh., Clarke-Midura, J., & Klopfer, E. (2014). A Framework for Structuring Learning Assessment in a Massively Multiplayer Online Educational Game: Experiment Centered Design. *International Journal of Game-Based Learning*. DOI: 10.4018/IJGBL.2014010103. Retrieved from https://www.researchgate.net/publication/261960598_A_Framework_for_Structuring_Learning_Assessment_in_a_Massively_Multiplayer_Online_Educational_Game_Experiment_Centered_Design.

Dichev, C., & Dicheva, D. (2017). Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *International Journal*

of Educational Technology in Higher Education, (14)9. DOI: 10.1186/s41239-017-0042-5.

Efremova, N. F. (2018). Design of evaluation tools based on the model of evidence-based argumentation. *Pedagogicheskie izmereniya*, 3. Retrieved from <https://cyberleninka.ru/article/n/proektirovanie-otsenochnyh-sredstv-po-modeli-dokazatelnoy-argumentatsii/viewer>. (In Russian).

Efremova, N. F. (2019). Conceptual Model of Evaluation of Competence of Students. *Modern high-tech Technologies*, 7, 169–174. Retrieved from www.top-technologies.ru/ru/article/view?id=37607. (In Russian).

Efremova, N. F. (2019). Modeling of an objective assessment of students' achievements. *Obrazovatelnye tekhnologii*, 3. Retrieved from <https://cyberleninka.ru/article/n/modelirovanie-obektivnoy-otsenki-dostizheniy-studentov/viewer>. (In Russian).

Groff, J., Clarke-Midura, J., Owen, V. E., et al. (2015). *Better Learning in Games. A Balanced Design Lens for a New Generation of Learning Games*. Learning Games Network, MIT Education Arcade. Retrieved from <https://education.mit.edu/wp-content/uploads/2018/10/BalancedDesign-Guide2015.pdf>.

Khlopotov, M. V. (2014). Bayesian network in student model engineering for competence level evaluation. *Naukovedenie*, 5(24). Retrieved from <https://cyberleninka.ru/article/n/primenenie-bayesovskoy-seti-pri-postroenii-modeley-dlya-otsenki-urovnya-sformirovannosti-kompetentsiy/viewer>. (In Russian).

Khudenko, L. A., & Avdeeva, S. M. (2014). Research of information and communication competence of 9th grade students of the Republic of Belarus: purpose and main results. *Bulletin of Education*, 1, 3–10. Retrieved from https://www.ntf.ru/sites/default/files/%D0%A0%D0%91_%D1%81%D1%82%D0%B0%D1%82%D1%8C%D1%8F_%D0%B8%D0%BA%D0%BA.pdf. (In Russian).

Mislevy, R. J., Almond R. G., & Lukas, J. F. (2003). *A brief introduction to evidence-centered design: Research Report*. Princeton, N.J.: Educational Testing Service. Retrieved from <https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2333-8504.2003.tb01908.x>.

Shute, V. J., Masduki, I., & Donmez, O. (2010). Conceptual Framework for Modeling, Assessing and Supporting Competencies within Game Environments. *Tech., Inst., Cognition and Learning*, 8, 137–161.

Uglanova, I. L., Brun, I. V., & Vasin, G. M. (2018). Evidence-Centered Design method for measuring complex psychological constructs. *Journal of Modern Foreign Psychology*, 7(3), 18–27. DOI: 10.17759/jmfp.2018070302. (In Russian).

Uglanova, I. L., Orel, E. A., & Brun, I. V. (2020). Measuring creativity and critical thinking in Primary School. *Psikhologicheskii Zhurnal*, 41(6), 96–107. DOI:10.31857/S020595920011124-2. (In Russian).

University of the State of New York, State Education Department. (2019). *Educator Guide to the Regents Examination Global History and Geography II Grade 10*. Retrieved from www.p12.nysed.gov/assessment/ss/hs/framework/ghg2/educator-guide-ghg2-19.pdf.

Zieky, M. J. (2014). An introduction to the use of evidence-centered design in test development. *Psicología Educativa*, 20, 79–87. Retrieved from <https://journals.copmadrid.org/psed/art/j.pse.2014.11.003>.