

ПРОЕКТИРОВАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ПО МОДЕЛИ ДОКАЗАТЕЛЬНОЙ АРГУМЕНТАЦИИ¹

**Ефремова
Надежда Федоровна**

доктор педагогических наук, профессор,
заведующая кафедрой «Педагогические измерения»
Донского государственного технического универси-
тета, г. Ростов-на-Дону, Россия
nefremova61@donstu.ru

Ключевые слова: знания, компетенции, качество обучения, паттерны проектирования, модели ECD, результаты обучения.

В настоящее время одной из проблем является оценка компетенций, остро ощущается необходимость инновационных подходов к разработке и конструированию контрольно-оценочных средств для системы образования надежными способами, так как положительное влияние оценки проявляется только в том случае, если она действительно отображает реальный результат обучения. Поэтому особенно актуальным является поиск приемов и моделей создания надежного педагогического инструментария не только для выявления усвоенных знаний, но и глубоко латентных характеристик обучающихся, так называемых навыков XXI века. А для получения надежной оценки необходимы организационная среда и набор функциональных требований для обеспечения эффективных действий обучающихся при выполнении заданий. Как утверждает П. Гриффин², вероятностная интерпретация результатов обучения может обеспечить основу для установления связи между оценкой, обучением, ресурсами учебного плана и разработкой образовательной политики. Это утверждение в сочетании с выводами Л.С. Выготского (1931) свидетельствует, что аутентичная оценка помогает указать обучающемуся зону проксимального (ближайшего) развития с точки зрения достигнутого уровня и определить, над чем ему необходимо работать дальше.

Функция оценки является одним из важнейших факторов активизации деятельности при усвоении знаний, приобретении умений и формировании компетенций. Знание объективных результатов обучения – обязательное условие развития самооценки возможностей и дальнейшего развития обучающихся. В том случае, если оценки надежны и вызывают доверие обучающихся, они являются мощным фактором повышения учебной мотивации³. Поэтому общая тенденция многих исследований проблем оценивания достижений обучающихся как раз и

¹ Исследование выполнено при поддержке гранта программы ЭРАЗМУС+

² Griffin P., McGaw Barry, E. Care. Editors Assessment and Teaching of 21st Century Skills. Springer Dordrecht Heidelberg. London-New York. 2012. 362 p. DOI 10.1007/978-94-007-2324-5.

³ Ефремова Н.Ф. Мотивационный аспект независимого оценивания достижений обучающихся / Российский психологический журнал. – 2017. – Т. 14, № 2. – С. 227–244.

состоит в поисках надежных методов разработки надежных средств, обеспечивающих достоверные оценки.

Цель данной статьи — дать представление о методе проектирования оценочных средств на основе доказательной аргументации Evidence-Centered Design (ECD), получившем широкое распространение в зарубежной образовательной практике и мало известном в отечественной. Метод доказательной конструкции ECD обеспечивает связь между требованиями высокой точности и большим набором данных для надежной оценки. Предлагаемые в нем задачи должны представлять структурированные ситуации как с однозначными решениями, так и проблемы с множественными решениями в незнакомых ситуациях. Такие комплексные задания позволяют получать как качественную, так и количественную информацию об уровне сформированности когнитивной и компетентностной составляющих результатов обучения.

Метод Evidence-Centered Design, или паттерн-дизайн оценки, ориентированный на доказательство, принципиально новый подход к проектированию и разработке оценочных средств на основе аргументации выводов для обоснования оценки⁴. В предложенном Р. Мислеви (Mislevy R.J., 1994 г.) подходе формирование оценочного средства идет с обоснованием его относительной важности, полезности, надежности и условий применения с наглядными демонстрациями результатов обучения при решении заданий ECD. Для разработки дизайна оценки на доказательствах методики ECD используется сборка различных учебных элементов по целям и задачам оценки, статистические методы и экспертный анализ аргументов надежности самого оценочного средства. В элементах задания указывается, сколько и какие типы задач будут использоваться, как и в каких пределах будут отбираться утверждения, какие дополнительные доказательства, аргумента-

ции и опровержения в оценочном процессе обеспечат психометрические качества и достоверность педагогического инструментария. Ключевая идея ECD заключается в том, чтобы дать обоснование аргумента оценки, основанного на создании конструкта, задающего выбор или построение соответствующих задач, а также обоснование критериев и оценочных рубрик, в которых основное внимание уделяется доказательствам в качестве основы построения оценочных задач и шкал⁵.

Подход ECD дает наиболее точные выводы, поскольку в нем меньше внимания уделяется оценке непосредственно знаний и навыков, а больше фокусируется внимание на том, что обучающиеся могут делать, применяя для этого необходимые знания. В ECD используются три основные модели: модель обучающегося (знания, навыки, компетенции и другие атрибуты, которые подлежат оцениванию); модель доказательств и правил оценивания (оценочный компонент, измерительный компонент, продукты деятельности, баллы, средний балл по выборке, правила анализа взаимодействия обучающегося с заданием и фиксации ответов на заданную проблему, что должно наблюдаться, как это наблюдается, какие критерии используются, как совместить баллы с результатами наблюдений); и модель задания, обеспечивающего интеллектуальный переход от баллов к выводам (сколько и какие типы задач будут использованы для обоснования каждого утверждения, какие задачи должны быть выполнены и какие действия необходимы для их выполнения). Модели обеспечивают связь между аргументом оценки и деятельностью обучающихся при выполнении заданий для получения надежной оценки. Все это приводит к разработке «концептуальной основы оценки (Conceptual Assessment Framework — CAF и модели сборки)» или формальному дизайну оценки.⁶

⁴ Mislevy, R. J. (1994). Evidence and inference in educational assessment. *Psychometrika*, 59, Pp. 439-483.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A Brief Introduction to Evidence-centered Design. Educational Testing Service, Princeton, N-J. 37 p.

Rupp, A. A.; Gushta, M.; Mislevy, R. J.; Shaffer, D.W. (2010). Evidence-centered Design of Epistemic Games: Measurement Principles for Complex Learning Environments, *JTLA*, Vol 8, No 4. 48 p.

⁵ Snow, E., Fulkerson, D., Feng, M., Nichols, P., Mislevy, R., & Haertel, G. (2010). Leveraging evidence-centered design in large-scale test development (Technical Report 4). Menlo Park: SRI International, Center for Technology in Learning.

A Novel Model of Cognitive Presence Assessment Using Automated Learning Analytics Methods (2017). V. Kovanović, D. Gašević, University of Edinburgh Marek Hatala, Simon Fraser University George Siemens, University of Texas at Arlington. 33 p. DOI 10.13140/RG.2.2.10611.32804.

⁶ Rupp, A. A.; Gushta, M.; Mislevy, R. J.; Shaffer, D.W. (2010). Evidence-centered Design of Epistemic Games: Measurement Principles for Complex Learning Environments,

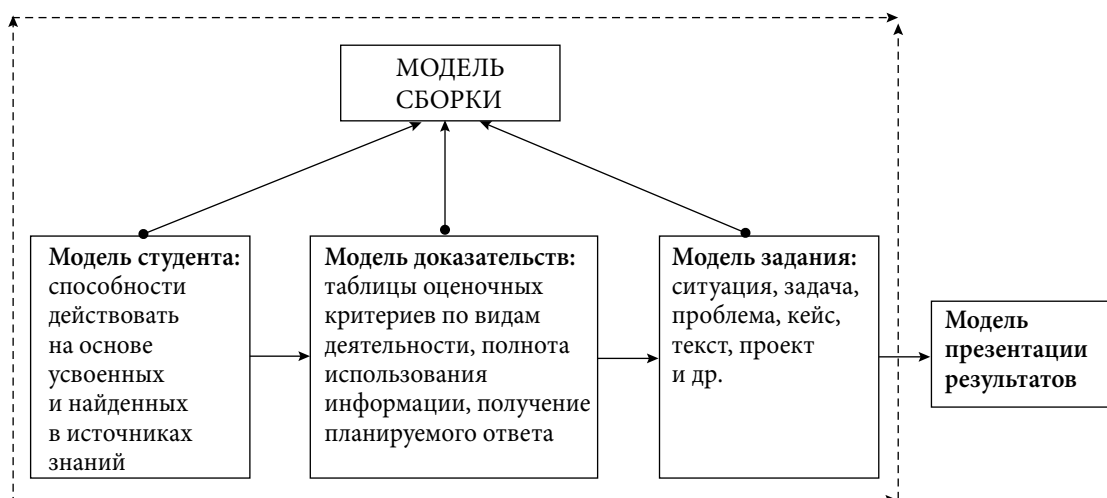


Рис. 1. Основные проектные объекты концептуальной основы оценки (CAF)

Итак, основные элементы паттерн-дизайна Evidence-Centered Design: модель студента, модель доказательств, модель задания и модель презентации результатов (рис. 1).

Модель обучающегося описывает конструкт исследуемых характеристик как набор латентных переменных (количество аргументов, структура конструкта), виды деятельности, причинно-следственные связи между ситуацией и деятельностью обучающегося.

Модель задания обеспечивает структуру аргументов оценки (валидные ситуации и сценарии, акцент на наиболее важном, ключевые аргументы, принципы отбора заданий, спецификация, планируемые действия испытуемого), связь аргументов с функциями задания, симуляционные среды, в которых обучающийся действует, чтобы обеспечить доказательства своей подготовленности. Разработчик оценки в модели задания указывает формы и ключевые особенности задач, материалы стимула и особенности презентации результатов выполнения заданий.

Модель доказательств и свидетельств задает оценочный компонент задания, баллы по задачам, общий балл по заданию, продукты деятельности, сбор свидетельств и аргументов оценки по результатам выполнения задания и/или оценке продукта деятельности, правила оценки процессов или продуктов деятельности, отображение результатов на шкале (для компетенций – на уровневой шкале).

Концептуальная основа оценивания указывает на сопряжение модели обучающегося,

модели свидетельств и модели задания (выбор при наличии проверенных шаблонов или построение оригинального паттерн-дизайна), а также на переход от баллов к выводам об оценке.

Отличительной чертой ECD является формализация процесса разработки средств оценки относительно целевых конструкций. Для снижения затрат на разработку таких заданий можно создавать многократно используемые шаблоны оценочных средств, структура которых основывается на идее С. Мессика, согласно которой «характер оцениваемой конструкции должен определять выбор или построение соответствующих задач и требований к их решениям, а также рациональную разработку критериев и рубрик, основанных на выбранном конструкте⁷, общности аргументов, структур, процессов и материалов оценки.

Шаблоны проектирования структурируют информацию о целевых навыках, эффективности действий испытуемых и использовании ими ситуаций с точки зрения аргументации структуры оценки (рис. 2).

Процесс разработки оценочного средства предусматривает ряд этапов проектирования: документирование информации о структуре и силе доказательных отношений; согласование работы разработчиков и психометриков при калибровке заданий измерительной

⁷ Messick, S. (1996). Validity of Performance Assessments. In G.W. Phillips (edn), Technical Issues in Large-Scale Performance Assessment. Washington. NCES Publication No. 96-802. Pp. 1–18. – С.17.

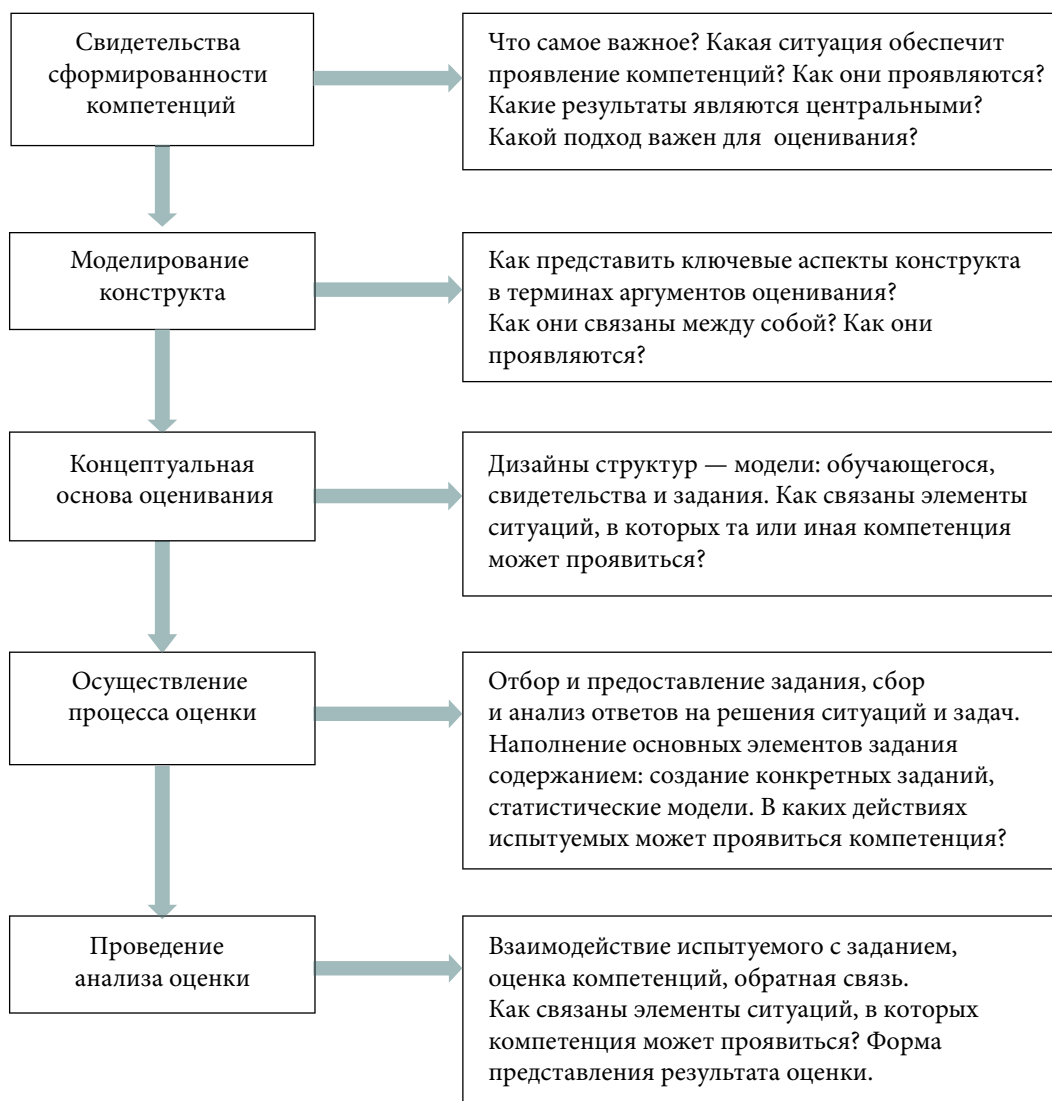


Рис. 2. Основные этапы разработки оценочных средств

модели; поиск необходимой ситуации и научной информации, обеспечивающих основу и надежность оценки. Этот метод может использоваться для разработки всех видов оценок и особенно для таких, которые включают сложные модели компетенций и динамические интерактивные среды, лежащие за пределами аналитических возможностей при использовании более простых моделей оценки. Модели ECD дают возможность обосновать дизайн оценки и способ рассуждений о подготовленности и успешности деятельности обучающегося. Аргумент оценки обеспечивает формальную основу для выявления знаний, навыков и компетенций обучающихся, а также требований к условиям, в которых будет проводиться оценочный процесс (наблю-

дения, доказательства, сбор данных для поддержки вывода об оценке).

Таким образом, в ECD оценка рассматривается как последовательность запланированной деятельности обучающихся и рассуждений, связывающих доказательства с утверждениями, чтобы сделать выводы об их знаниях, навыках и компетенциях. Для оценочных средств, отвечающих целям оценивания тех или иных образовательных достижений, удобство проектирования по моделям с дизайном ECD заключается в том, что может быть разработан ряд шаблонов (паттернов), используемых многократно для разных предметных областей. Для этого прежде всего важно сделать выбор переменных измерения и смоделировать конструкт оцениваемых

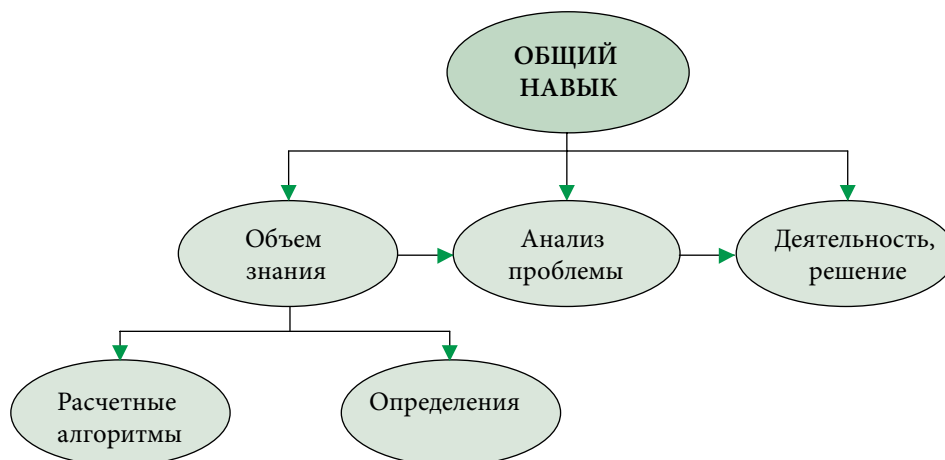


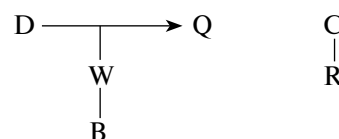
Рис. 3. Пример конструкта для разработки шаблона оценочного средства

характеристик обучающихся, один из примеров которых приведен на рис. 3.

Алгоритм ECD обеспечивает все проектные решения, связанные со спецификацией задачи, конкретным заданием, дизайном и администрированием, представлением и доставкой элементов задачи, а также системой оценки и шкалирования результатов, их анализа, интерпретации и составления отчетов. Для этого применимы: компетентностно-ориентированные задания, наборы элементов заданий с несколькими вариантами их решения в условиях неопределенности, поиск информации для написания эссе на заданную тему, выступление и презентация, подготовка статьи для публикации в открытой печати, участие в деловой игре, решение кейсов, выполнение проектов индивидуальных или в малых группах и др. При этом шаблон проектирования оценочного средства является ядром решения проблемы, которая неоднократно повторяется в образовательной практике, когда высок уровень общности целей оценки настолько, что подход может применяться во многих других ситуациях при некоторой адаптации к конкретному случаю. Каждый шаблон проектирования основывается на общей форме аргумента оценки, касающейся знания или умения, видов наблюдений, которые могут служить доказательством знаний или навыков, а также особенностей выполнения задач, которые позволяют обучающимся продемонстрировать уровень подготовленности и предоставить этому необходимые доказательства.

Для иллюстрации и облегчения доказательной аргументации С. Тулмин создал

«графическую диаграмму рассуждений, которая соединяет шесть основополагающих элементов, которые устанавливают достоверность, актуальность и силу доказательств в отношении целевых выводов по оценке»⁸. В общем виде *C* – утверждение (claim), *D* – данные (data), *W* – основание (warrant), *B* – поддержка (backing), *R* – опровержение (rebuttal) и *Q* – квалификатор (qualifier) связаны схемой



Допущения и предположения служат основанием для заключения, как и почему из имеющихся данных следует именно сделанное утверждение об оценке. Поддержка усиливает основания за счет дополнительных данных (контрольные вопросы, результаты анкетирования и др.). Опровержение или контраргумент указывают на ограниченность утверждения. Квалификатор как набор слов (вероятно, возможно, поскольку, невозможно, безусловно, предположительно, всегда и др.) ограничивает достоверность утверждения и подчеркивает, при каких условиях утверждение может быть верным.

На основе шаблона последовательно выстраивается дизайн оценочного средства и описание сценария оценочного процесса.

⁸ The Uses of Argument. By Stephen E. Toulmin (2003, Paperback, Revised). Cambridge, UK: University Press. 259 p. – С. 106.

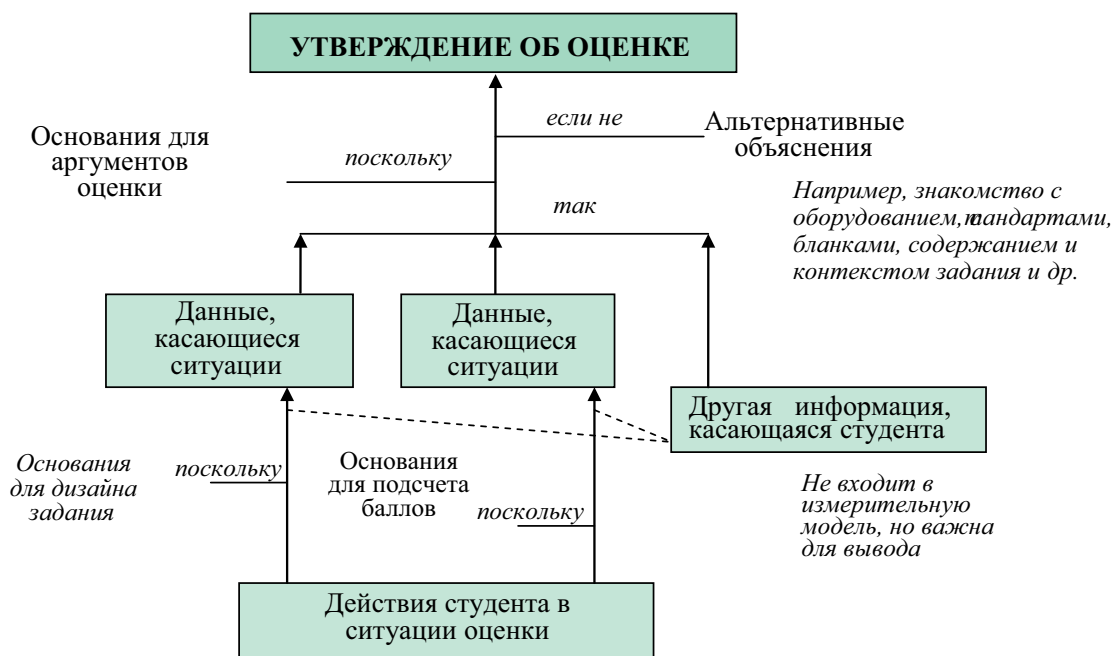


Рис. 4. Структура аргументов в оценке ECD по шаблону С. Тулмина

Модель доказательств объединяет модель компетенций и модель задачи (подзадач), описывает, что должно быть выполнено обучающимся, как это выявить и как совместить наблюдения и свидетельства с баллами и уровневыми шкалами для окончательного вывода об оценке сформированности компетенций (рис. 4).

Таким образом, шаблоны проектирования представляют собой инструменты для описания структуры и содержания заданий, которые могут поддерживать разработку и реализацию мер по изучению стратегий и организации поведения испытуемых при решении заданий. При необходимости шаблоны могут пересматриваться для улучшения психометрического качества и повышения надежности разрабатываемых на их основе оценочных средств. Они представляют собой компоненты аргумента, необходимые для обоснования выводов об оценке когнитивной и компетентностной подготовленности обучающихся.

Многомерный конструкт комплексного задания для оценки знаний, умений, компетенций и других атрибутов обеспечивает выбор конкретных задач на основе реальных или квазиреальных ситуаций, интересных для решения, разработку дизайна и спецификации оценочного средства (каких и сколько заданий необходимо) (рис. 5).

Рассуждения указывают на высокую степень взаимосвязи между множеством решений, которые должны быть приняты в процессе разработки, осуществления, оценки и отчетности с целью удовлетворения доказательственных требований заинтересованных сторон. Одни и те же данные могут быть хорошим доказательством для одного вывода, но плохими доказательствами для другого (Schum, 1994). Согласно этому подходу, считается, что данные становятся свидетельством только тогда, когда установлено их значение для соответствующих выводов.

Это особенно важно для проектирования инструмента оценки, потому что на каждом этапе выполнения задания аргумент оценки необходим для получения ответа на ряд вопросов. Что именно требуется узнать о подготовленности обучающихся? Какие действия должны выполнять испытуемые, как они это делают и в каких ситуациях?

Конкретизация модели С. Тулмина в этой статье автором представлена шаблоном рассуждений для получения сведений о сформированности информационной компетенции обучающегося (рис. 6).

В более формальных терминах, основанных на общей структуре аргументаций С. Тулмина, Р. Мислеви (2006) утверждает, что именно вопросы формируют основу дизайна для получения выводов об оценке, а

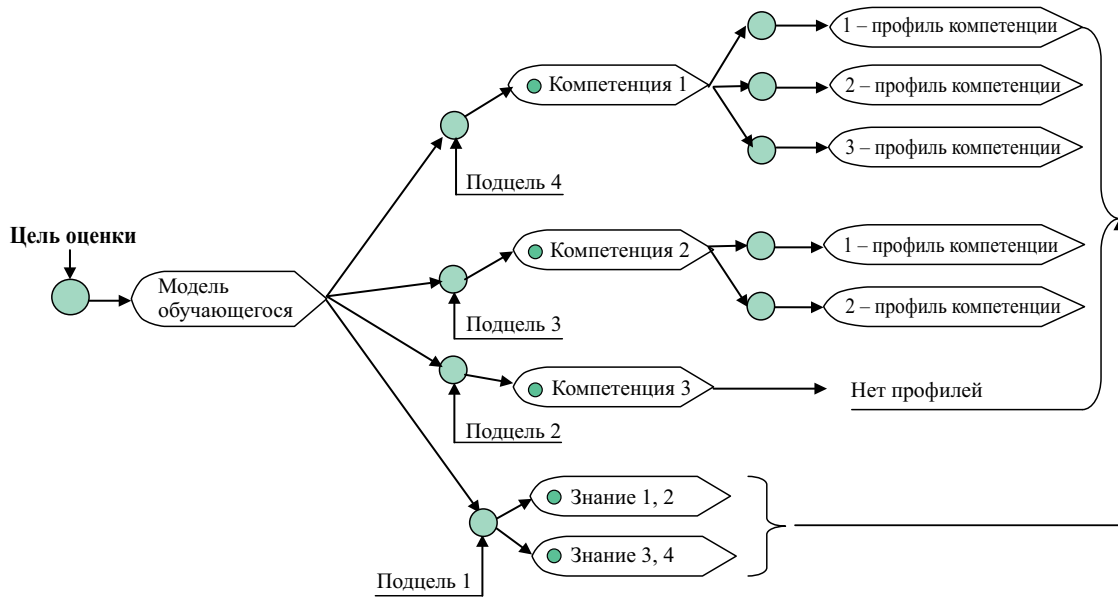


Рис. 5. Шаблон многомерного конструкта

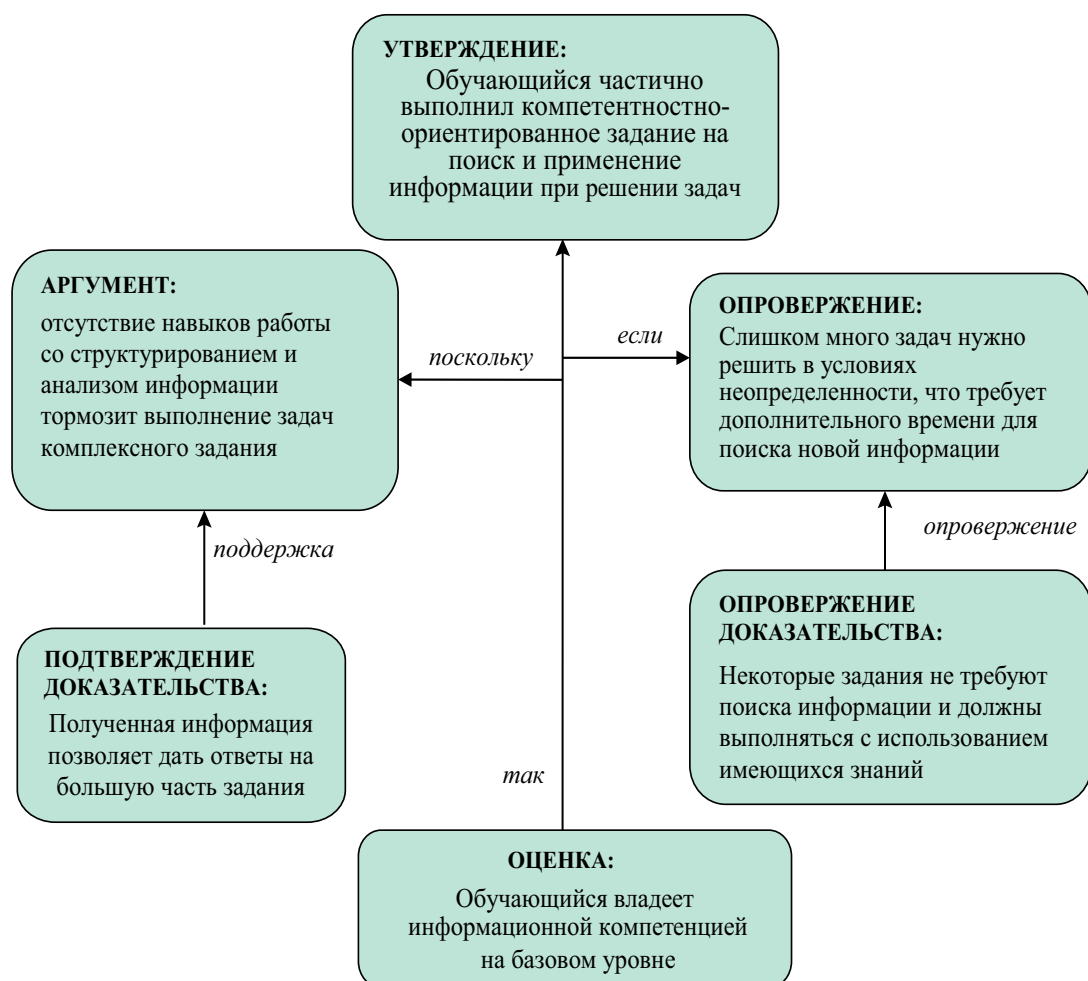


Рис. 6. Шаблон доказательной аргументации сформированности информационной компетенции

также того, какие доказательства или интерпретации необходимы или какая психологическая перспектива оценки используется для формирования аргумента⁹.

Структура ECD позволяет помочь разработчикам оценки структурировать свое мышление и свои действия для решения сложной задачи оценивания комплекса характеристик обучающихся. Весь процесс проектирования оценки управляется контролируемым содержанием предметной области и отбором элементов задания, в котором указывается, сколько и какие типы задач могут быть использованы для оценки каждого утверждения, а в пределах каждого утверждения – сколько, какие и как бу-

дут отбираться и распределяться оценочные суждения.

Использование ECD эффективно решает валидационные исследования с точки зрения процедурного развития процесса планирования и конструирования комплексного задания. Однако следует отметить, проектирование комплексной оценки – очень сложная работа. Поэтому команда, разрабатывающая такие оценочные средства, должна включать проектировщиков оценки, дизайнеров, учителей, инженеров баз данных, разработчиков программного обеспечения, когнитивных психологов и психометриков. Это требует достаточно высокой квалификации разработчиков и немалых затрат, но единоразово правильно разработанный дизайн оценочного средства может быть использован многократно, а главное, давать надежную информацию, использоваться широкой аудиторией педагогов, что вполне оправдывает первоначальные затраты на его разработку.

⁹ Mislevy, R. J., Corrigan, S., Oranje, A., Dicerbo, K., John, M., Bauer, M. I., Hao, J. (2014). Psychometric Considerations in Game-based Assessment. New York, NY: Institute of Play.

Riconscente, M.M., Mislevy, R. J., Corrigan, S. Evidence-Centered Design from: Handbook of Test Development Routledge. 2015. London. Pp. 40-63.