

УДК 159.9.072

ИЗМЕРЕНИЕ КРЕАТИВНОСТИ И КРИТИЧЕСКОГО МЫШЛЕНИЯ В НАЧАЛЬНОЙ ШКОЛЕ

© 2020 г. И. Л. Угланова*, Е. А. Орел**, И. В. Брун***

*Национальный исследовательский университет “Высшая школа экономики”,
Институт образования, Центр психометрики и измерений в образовании;
101000, г. Москва, Потаповский пер., д. 16, стр. 10, Россия.*

**Аспирант, стажер-исследователь.*

E-mail: iuglanova@hse.ru

***Кандидат психологических наук, научный сотрудник.*

E-mail: eorel@hse.ru

****Младший научный сотрудник.*

E-mail: ibrun@hse.ru

Поступила 02.03.2020

Аннотация. В статье описывается процедура создания инструмента оценки креативности и критического мышления учащихся начальной школы, разработанного в логике “систематического подхода к разработке измерительных инструментов” (Evidence-centered Design, ECD). Компьютеризированный тест представляет собой набор сценариев, где учащийся должен в игровой форме решить задачу из своего повседневного опыта (оформить домашний аквариум, выдумать персонажа для фантастической истории и пр.). В соответствии с методологией ECD проведен теоретический анализ измеряемых конструктов, предложена их структура, описаны разработанные сценарии, а также проведенное эмпирическое исследование ($N = 909$ учащихся четвертых классов). Была подтверждена валидность разработанного инструмента, в том числе его соответствие теоретическим ожиданиям о структуре креативности и критического мышления в раннем школьном возрасте.

Ключевые слова: Evidence-centered Design, валидизация инструментов измерения, начальная школа.

DOI: 10.31857/S020595920011124-2

Критическое мышление и креативность считаются одними из самых важных навыков современного человека. Так, в докладе Всемирного экономического форума¹ креативность, критическое мышление, коммуникация и кооперация образуют группу ключевых компетенций в составе общей рамки навыков XXI в. Их же называет в своем аналитическом докладе о будущем образования ЮНЕСКО [29], включая их в группу навыков “учиться делать”. Естественно, включаются они и в список навыков XXI в., подготовленный под руководством П. Гриффина [15].

В российских образовательных стандартах креативность и критическое мышление можно найти в разделе “метапредметные образовательные результаты”. Так, в начальной школе ФГОС требует освоить “способы решения проблем творческого и поискового характера” или овладеть “логическими

действиями сравнения, анализа, синтеза, обобщения, классификации”².

Таким образом, значимость этих навыков для успешной жизни в современном мире признается образовательным сообществом как на уровне политики (через наиболее важные программные документы о развитии образования), так и педагогической практики (через образовательные стандарты). Однако, помимо признания и развития, система образования предполагает также этап оценивания формируемых компетенций и знаний. Следовательно, логичным и востребованным шагом является разработка инструментов оценки креативности и критического мышления.

В своей работе мы поставили *цель* описать процесс разработки инновационного инструмента оценки креативности и критического мышления

¹ The Future of Jobs Report 2018. World Economic Forum, Geneva, 2018.

² Федеральный государственный образовательный стандарт начального общего образования. 2009. С. 41.

в конце начальной школы и привести доказательства его валидности. В качестве основы для создания инструмента была выбрана методология Evidence-centered Design, или “систематический подход к разработке измерительных инструментов”. В рамках этой методологии разработчики создают задания-сценарии, максимально приближенные к реальной жизни, однако позволяющие валидно и надежно оценить критическое мышление и креативность, основываясь на системе свидетельств — том, как эти навыки выражаются в поведении учащихся при решении конкретных задач.

МЕТОДОЛОГИИ РАЗРАБОТКИ: EVIDENCE-CENTERED DESIGN

В работе мы придерживались методологии разработки инструментов измерения Evidence-centered Design (ECD) [24]. В отечественной литературе можно встретить перевод этого термина как “систематический подход к разработке измерительных инструментов” [7] и “метод проектирования тестов при помощи доказательной аргументации” [3]. В рамках ECD разработка инструмента измерения включает несколько последовательных этапов:

1. Определение содержательной области: анализ и моделирование области компетенций. Разработка инструмента начинается с анализа литературы, проведения экспертных панелей и интервью с целевой группой для получения полного представления об изучаемом конструкте.

2. Формирование модели конструкта: устанавливается предположительная внутренняя структура конструкта, его описание в измеряемой форме — в терминах наблюдаемого поведения. Определяется подход к измерению: оценка продукта деятельности или самоотчет.

3. Формирование модели задания: определение конкретных действий тестируемого. Она складывается из ключевых паттернов, которые будут характеризовать все элементы теста, и включает инструкции, ключевые характеристики стимульного материала и типа задания.

Модель задания — это скелет, который обрastaет деталями, специфичными для сценария. Для каждого сценария выбирается определенный контекст (учебный или бытовой), в котором свидетельства конструкта выражаются в наблюдаемых действиях тестируемого. Действия тестируемого, которые потенциально отражают ненаблюдаемый конструкт, называются *индикаторами*.

4. Система начисления баллов и модель измерений. Система начисления баллов продумывается для каждого индикатора: определяются правила, за что тестируемый получает “0” или “1” балл (на дихотомическом примере). Систему начисления баллов и модель конструкта связывает модель измерений. Модель измерений призвана определить, какая математическая (психометрическая) модель выступает наиболее подходящей для анализа определенного типа данных в рамках определенной цели тестирования.

В целом преимущество методологии ECD в том, что она выстраивает логику разработки инструментов от представлений о конструкте до наблюдаемого поведения. С более подробным описанием методологии можно ознакомиться в англоязычных [24] и русскоязычных источниках [3; 7].

АНАЛИЗ И МОДЕЛИРОВАНИЕ ОБЛАСТИ КОМПЕТЕНЦИЙ: СОВРЕМЕННЫЕ ПОДХОДЫ К ОПИСАНИЮ И ИЗМЕРЕНИЮ КРИТИЧЕСКОГО МЫШЛЕНИЯ И КРЕАТИВНОСТИ

Задача этого раздела — привести обзор современных подходов к определению креативности и критического мышления, чтобы на следующем шаге определиться с конкретным содержанием, которое войдет в наш инструмент. Поскольку теорий, описывающих эти конструкты, существует огромное множество, мы остановимся на тех, которые ложатся в основу современных измерительных инструментов.

Креативность

Обзору исследований креативности и ее роли в образовательном процессе можно посвятить отдельную статью. Здесь мы вынуждены остановиться только на принципиальных моментах, определяющих содержание измеряемого нами конструкта.

Началом массовых исследований креативности в психологии и образовании, по-видимому, можно считать работы Дж. Гилфорда по дивергентному мышлению — “продуцированию большого количества разнообразных идей в ответ на поставленную задачу или проблему” [16]. С тех пор появилось множество уточнений и определений креативности [9; 10; 18; 19; 21], однако за последние 25 лет большинство ученых пришли к общему пониманию, что креативность — это “продуцирование идей или вещей, которые а) являются новыми или оригинальными, с точки зрения их создателя,

и таким же важным считается б) их полезность для решения задачи или хотя бы релевантность” [11]. Это определение стало основой для создания моделей креативности, некоторые из них противоречат друг другу.

Так, дискуссионным выступает вопрос, являются ли креативность универсальной или ее проявления зависят от области, в которой возникает задача. Сторонники первого подхода полагают, что составляющие креативности когнитивные навыки, особенности мотивации или личностные характеристики будут одними и теми же в любой сфере деятельности [9; 20]. Однако сегодня преобладает позиция, согласно которой проявления креативности специфичны по отношению к проблемной области [14; 17]. Иначе говоря, можно быть креативным ученым и делать остроумные шаги в исследовательской работе, но при этом не уметь сочинять рассказы.

Другая модель, предложенная М. Роудсом [26], описывает креативность в четырех измерениях: личность (качества креативного субъекта), процесс создания нового, характеристики креативного продукта и особенности среды. Еще один подход [18] разделяет креативность на “большую” (big C) — крупные новые идеи, меняющие мир, возникающие нечасто, но имеющие существенное влияние на человека, и “малую” (little C) — оригинальные решения задач повседневной жизни каждого человека.

В отечественной психологии близким к креативности является понятие “творческое мышление” — интегральное психическое образование, характеризующееся “...мерой умственной инициативности, ее наиболее значимой качественной характеристикой, может служить интеллектуальная инициатива, понимаемая как продолжение мыслительной деятельности за границами ситуативной заданности, никак не предопределенное ни практическими нуждами, ни внешней или субъективной отрицательной оценкой работы” [2]. “Преодоление границ ситуативной занятости” в повседневных задачах соотносится с представлениями о “малой” креативности.

Различные подходы помогли нам определиться с контекстом задач, используемых в сценариях. Во-первых, из-за того, что креативность специфична по отношению к сфере деятельности, мы выбираем максимально нейтральный контекст: если бы мы взяли только задачи в одной сфере (скажем, в естественных науках), то мы не смогли бы оценить учащихся, которые не обладают креативностью именно в этой сфере, и существенно сузили фокус оценки. Поэтому мы остановились на внеучебных, повседневных задачах (“малой” креативности), в которых часто совмещаются различные сферы

деятельности. Во-вторых, модель Роудса позволяет описать сценарий — какое поведение мы ожидаем от учащегося, какой продукт получится в результате решения, какие шаги он (она) должны предпринять, чтобы его получить, и каким должен быть интерфейс задания, чтобы стимулировать процесс решения задачи.

Критическое мышление

В этом разделе мы остановимся на современных подходах к критическому мышлению в рамках философской, психологической и педагогической традиций.

В философской традиции признаками критического мышления выступают обоснованность суждений и следование правилам формальной логики [25]. Большинство исследователей сходятся во мнении, что оно определяется через рациональное, осознанное решение о том, что делать или чему верить [14].

Примером такого подхода является перечень составляющих критического мышления, предложенный Э. Лай [22]. Она выделяет четыре компонента, общих для различных описаний этого конструкта: 1) анализ аргументов, высказываний, свидетельств; 2) способность сделать вывод, используя индукцию или дедукцию; 3) оценка и вынесение суждения; 4) принятие решений или решение проблем.

К этой же группе относится и модель, предложенная Л. Лью и ее коллегами [23], с важным дополнением в виде эмпирических результатов. Модель была специально создана для разработки инструмента оценки критического мышления студентов вузов. В ней выделяются три группы размерностей: аналитический компонент (оценка надежности источников, релевантности аргументов и пр.), синтетический компонент (логический вывод, оценка последствий и т.д.) и общий (построение причинно-следственных связей и оценка альтернативных объяснений).

Внедрение критического мышления в образовательный процесс основывается на идеях Дж. Дьюи, который определял его как “активное, настойчивое и аккуратное рассмотрение обоснованности убеждения или априорных форм знания, в свете тех оснований, которые их поддерживают, а также выводов, которые из них следуют” [13]. В таксономии учебных целей Б. Блума [28] выделяется шесть уровней образовательных целей, связанных с глубиной освоения материала. Высшие (анализ, синтез и оценка) являются составляющими критического мышления.

Отечественная педагогика также не обошла вниманием проблемы развития критического

мышления в школе. Так, П.П. Блонский пишет о необходимости развития критического отношения уже к концу “центральной” части школьного курса [1]. Учитель обучает ученика не спешить с выводами (относиться критично даже к своим мыслям), собирать как можно больше фактов, подтверждающих или опровергающих полученную информацию, объясняет основы каузального анализа.

Психологическая традиция определения критического мышления отличается от философской и образовательной: она выявляет, как в действительности думает и поступает критически мыслящий субъект, в отличие от того, каким должно быть критическое мышление в идеальных условиях [23]. Пожалуй, самая известная операционализация критического мышления в психологии была предложена Стернбергом [30]. Он приводит следующее определение: “критическое мышление включает в себя мыслительные процессы, стратегии и репрезентации, которые используются людьми для того, чтобы принимать решения, решать проблемы и узнавать новые концепты”.

В отечественной психологии существует несколько понятий, близких к термину “критическое мышление”, как мы описывали его выше, однако не идентичных ему. Одно из таких понятий — “критичность мышления”. Это одно из свойств нормальной психической деятельности, способность осознавать свои ошибки, умение оценивать свои мысли, взвешивать доводы “за” и “против” и подвергать свои гипотезы всесторонней проверке [4–6].

С.Л. Рубинштейн приводит примеры развития критического отношения к поступающей информации у детей 4 лет и отмечает, что именно опыт дает им для этого опорные точки. Но только у подростков и юношей можно говорить о полном развитии критического мышления [5]. Д.Б. Эльконин отмечал, что к концу начальной школы формируется словесно-логический тип мышления, а также теоретический (если в обучении были использованы структурные единицы теоретического обобщения по В.В. Давыдову) [8]. Следовательно, уже в этом возрасте возможно проводить диагностические методики, чтобы определить, как идет развитие критического мышления у школьников.

МОДЕЛЬ КОНСТРУКТА

По результатам анализа литературы с учетом специфики выбранного подхода к измерению креативности мы предлагаем следующее рабочее определение.

Креативность — способность представить и разработать принципиально новые подходы к решению проблем, ответы на вопросы, стоящие перед субъектом, или способы выражения идей для решения задач в повседневной жизни.

На первом этапе мы выделили следующие составляющие: 1) оригинальность — способность продуцировать новые идеи и решения задач, устанавливать неочевидные связи между идеями; 2) детальность — глубина проработки идеи, выделение и осмысление различных аспектов решения задачи; 3) беглость — способность продуцировать несколько разных идей и/или вариантов решения задачи в ограниченное время; 4) любознательность — проявление интереса к различным аспектам задачи, демонстрация желания узнать о них больше. Чтобы сценарии выглядели более естественно, было принято решение отказаться от компонента “беглость”: не все задачи, требующие креативного решения, предполагают генерацию большого количества идей. В перспективе этот компонент может быть добавлен в сценарий. Также, поскольку решение задачи происходит в рамках заданного сценария, оценить любознательность оказывается крайне сложно. Наконец, в современных подходах продукт деятельности не может считаться креативным, если он не соответствует условиям задачи [11]. В нашем случае соответствие возникает автоматически: сценарий “ведет” респондента к решению, а значит, конечный результат будет отвечать условиям. Мы не рассматриваем этот аспект как существенное ограничение, так как в реальной жизни требования задачи также играют решающую роль в создании оригинального продукта. Итоговая модель конструкта “Креативность” представлена в табл. 1.

Второй конструкт, который мы измеряем, — это **критическое мышление**. Под ним мы понимаем **способность оценивать аргументы и данные, допущения, абстрактные понятия, чтобы вынести суждение, и сформулировать соответствующие вопросы для достижения решения** (табл. 2).

МЕТОДИКА

Модель заданий

Мы разработали три задания сценарного типа, предъявляемые тестируемому через браузер в онлайн-режиме. Респондент становится героем истории, в которую заложено несколько проблем, требующих последовательного решения.

Модель задания была разработана таким образом, что действия тестируемого, характеризующие одни

Таблица 1. Модель конструкта “Креативность”

Субкомпетенция	Описание и обоснование	Ожидаемое наблюдаемое поведение
Оригинальность	Способность продуцировать новые идеи и решения задач, которая может выражаться в появлении новых идей, в установлении новых, ранее не фиксировавшихся связей между существующими идеями	<i>Предлагает новые, оригинальные идеи</i> — придумывает решение, не существовавшее ранее в его опыте. <i>Устанавливает новые связи</i> — находит неочевидные связи между идеями
Детальность	Способность глубоко, с большой степенью подробности проработать предложенную идею	<i>Глубоко и детально прорабатывает идеи</i> — степень подробности описания, количество использованных деталей, в том числе отвечающих за то, чтобы идея продолжала работать в разных ситуациях

Таблица 2. Модель конструкта “Критическое мышление”

Субкомпетенция	Описание и обоснование	Ожидаемое наблюдаемое поведение
Анализ	Навыки работы с информацией в соответствии с целями и условиями поставленной задачи	<i>Выделяет надежные источники информации</i> — выносит суждения о достоверности источника и использует информацию из него. <i>Выделяет релевантную информацию для решения задачи</i> — выделяет в источнике только нужную информацию. <i>Выявляет недостаток информации и собирает полную информацию</i> — оценивает полноту информации, ищет дополнительные источники. <i>Исследует информацию</i> — категоризирует информацию на факты, мнения, предпосылки и пр. <i>Выявляет взаимоотношения между элементами текста</i> — распознает в тексте логические связи, противоречия и пр.
Вывод и аргументация	Построение собственного вывода (решение проблемы) и аргументов к нему на основе проведенного анализа	<i>Формирует собственную позицию</i> — вырабатывает решение, соответствующее условиям задачи. <i>Прорабатывает аргументы</i> — подбирает аргументы, подкрепляющие решение на основе результатов этапа анализа. <i>Оценивает вывод</i> — выносит суждения о сильных и слабых сторонах своего решения, формулирует альтернативные выводы

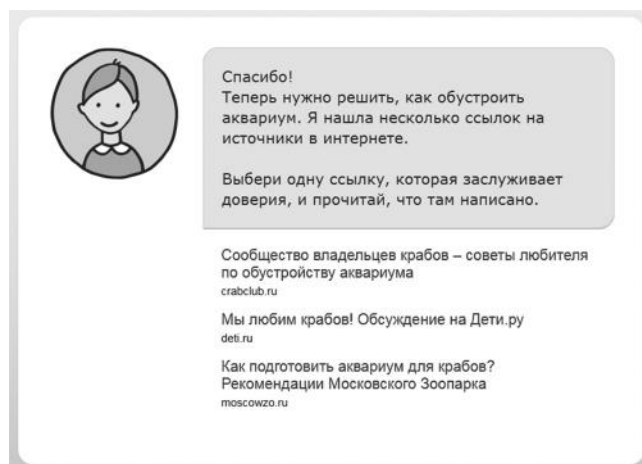
свидетельства конструктов, были схожи в различных сценариях. Так, для оценивания составляющей критического мышления Анализ тестируемому предлагали выбрать наиболее надежный и релевантный источник информации, а также найти в тексте ключевые предложения и классифицировать их. Для оценивания Вывода тестируемого просили принять решение, которое должно оказаться верным на основании проанализированной информации.

В задании Аквариум основная задача — построить аквариум для крабов. Для этого тестируемому необходимо выбрать релевантные источники информации, ознакомиться с ними, сконструировать аквариум из предложенных элементов. На этапе оценки Анализа тестируемый выбирает источник информации из нескольких ссылок. После выбора источника информации мы просим прочитать текст и выделить наиболее важные предложения для постройки аквариума (рис. 1).

Далее для оценки Вывода тестируемому предлагают собрать аквариум из предложенных элементов в конструкторе-симуляторе. На этом этапе

сравнивается согласованность ключевой информации в тексте и объектов, добавленных в аквариум.

Схожим образом оценка критического мышления происходит в сценарии Динозавр: нужно определить, сколько конечностей использовал для

**Рис. 1.** Интерфейс задания Аквариум. Этап оценивания Анализа

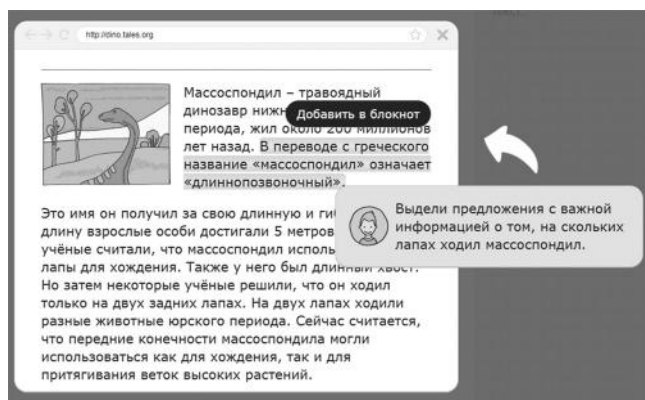


Рис. 2. Интерфейс задания Динозавр. Этап оценивания Анализа

ходьбы массоспондил. Респондент выбирает важную, с его (ее) точки зрения, информацию из разных источников, принимает решение и оценивает надежность вывода (рис. 2).

Для оценивания Детальности учитывалось количество элементов, которые использовались в конструкторе-симуляторе. В Аквариуме на этапе работы с конструктором оцениваются как креативность, так и критическое мышление (см. рис. 3, б). Если в тексте было указано на необходимость поставить определённый элемент и ученик ставит его, такое поведение служит индикатором Вывода; а добавление необязательных элементов, например, предметов декора, выступает индикатором Детальности.

Основная задача сценария Монстр — составить изображения монстров для разных целей (рис. 3, а). На первом этапе нужно составить фоторобот монстра-преступника, который скрылся с места преступления, улетев. Если ученик добавляет монстру крылья, то такое действие трактуется как проявление критического мышления.

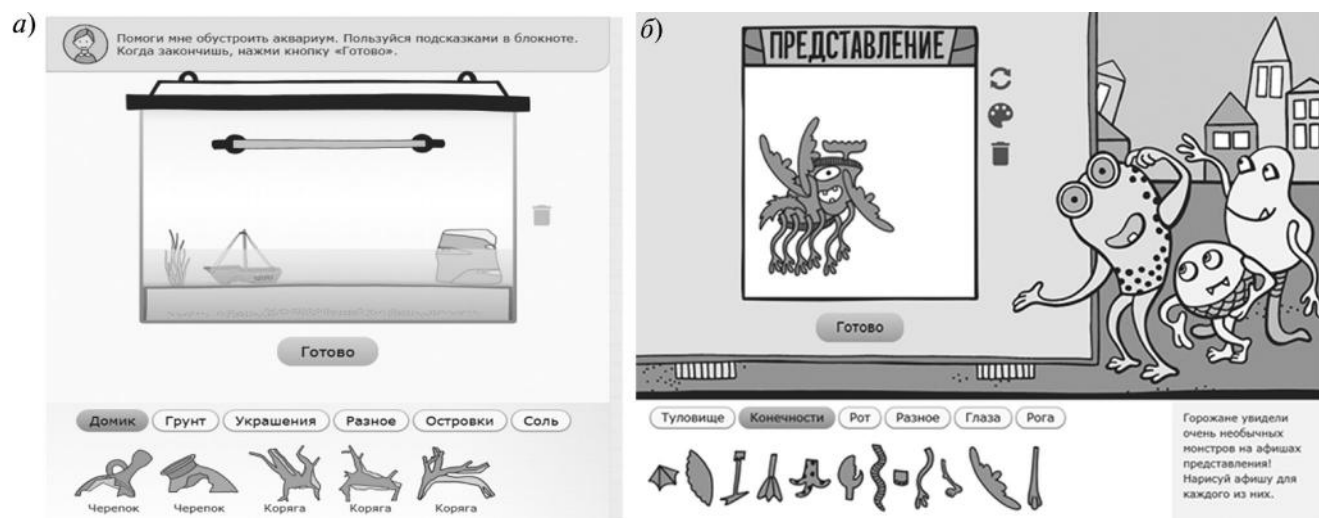


Рис. 3. Интерфейс задания Аквариум (а) и Монстр (б). Этап оценивания Вывода, Детальности и Оригинальности

В свою очередь, креативность оценивается по облику монстра: ученика просят нарисовать афишу с изображением монстра, который удивил бы всех жителей Города. Как свидетельства Детальности рассматриваются количество использованных элементов, изменение цвета или вращение элементов.

Для оценивания Оригинальности нами была разработана система “референсов”. Эта идея была введена для возможности начислять баллы автоматически, без помощи экспертов. Измерение креативности — это нетривиальная задача, которая становится только сложнее в ситуации с автоматическим анализом данных тестирования. В традиционных тестах на измерение креативности используется подход, основанный на экспертном оценивании и “нормировке”: после сбора ответов на вопрос или рисунков в заданной целевой группе эксперты сравнивают ответы тестируемых с наиболее “типичными” внутри данной группы [31]. Мы отошли от этапа сбора дополнительных данных и выстроили систему оценивания, не предполагающую работу экспертов. В сценарий вводятся “референсы” — недетализированные, упрощенные изображения. Оригинальность оценивается по отличиям от заданных референсов. В задании Монстр на фоне рабочего поля, где тестируемый делает афишу с изображением “нетипичного” монстра, изображены другие монстры этого города. Они имеют простую антропоморфную структуру: одно туловище, две симметрично расположенные “руки”, две симметрично расположенные “ноги”, рот ниже глаз. Когда ребенок строит своего монстра, мы сохраняем расположение элементов относительно друг друга и в дальнейшем количественном анализе можем выяснить, напоминает ли построенный ребенком монстр референсов. Каждое

заданное отличие фиксируется как проявление оригинальности.

Система начисления баллов и описание индикаторов

В фокусе этого и трех последующих разделов находятся результаты количественного анализа данных тестирования. Для удобства и краткости в этих разделах мы будем использовать два условных обозначения: КМ — критическое мышление, КР — креативность.

Для оценки компетенций было разработано 28 индикаторов, отражающих КМ, и 32 индикатора, отражающих КР, большая часть которых оценивалась дихотомически. В ходе предварительного анализа 4 индикатора, разработанные для оценки КМ, и 10 индикаторов, потенциально отражающие КР, показали свою неэффективность — в недостаточной мере отражали заявленный конструкт или оказались неудачными для оценивания компетенции в текущей версии программирования сценариев. Они были удалены из последующего анализа. Мы проанализировали проблемы удаленных индикаторов и учтем их в последующих итерациях разработки инструмента.

Таким образом, дальнейший анализ строится с использованием 24 индикаторов для оценки КМ и 22 индикаторов для оценки КР. Распределение индикаторов разных сценариев по компетенциям представлено в табл. 3.

Более подробное описание функционирования индикаторов выходит за рамки фокуса статьи, но доступна от авторов по запросу. В этой работе функционирование индикаторов описывается для подтверждения выводов о внутренней структуре инструмента.

Модель измерений и стратегия анализа данных

В работе как доказательство валидности рассматривается анализ согласованности теоретически ожидаемой структуры и полученных данных. Для этих целей в качестве модели измерений нами используется методологический аппарат конфирматорного факторного анализа (КФА).

Согласно модели конструкта, мы ожидаем, что действия испытуемых при прохождении сценария будут отражать две компетенции, каждая из которых представлена двумя составляющими. КР включает в себя составляющие Оригинальность и Детальность; КМ — Анализ и Вывод.

При использовании конфирматорного факторного анализа как модели измерений исследователь сталкивается с некоторыми ограничениями. Так, модель, которая состоит из латентного фактора, нагруженного двумя другими латентными факторами не может быть определена. Для максимального приближения модели измерений к модели конструкта построим модель, в которой будут представлены только составляющие компетенций. Модель будет включать четыре латентных фактора, связанные между собой. Мы будем считать, что модель удовлетворительна, если она, во-первых, отражает теоретические ожидания; во-вторых, встречает следующие условия относительно статистик согласия: $RMSEA < 0.5$; $CFI > 0.9$; $TLI > 0.9$ [12].

Задания сценарного типа отличаются от более типичных заданий, например заданий с множественным выбором, более нагруженной внутренней средой тестирования — контекстом заданий. Если несколько индикаторов связаны общим контекстом достаточно сильно, эту связь необходимо учитывать в модели измерений. В противном случае оценка параметров модели будет искажена из-за нарушения допущения о локальной независимости. Это приводит к невалидным выводам о качестве модели [32]. Для преодоления угроз локальной независимости в этом исследовании мы вводим дополнительные факторы — факторы контекста [27]. Эти факторы учитывают только принадлежность индикаторов к определенному аспекту сценария и не отражают компетенции. Факторы контекста ортогональны по отношению друг к другу и ко всем другим факторам. С такой же целью преодоления локальной зависимости в модель вводится несколько ковариаций ошибок индикаторов.

Инструменты измерения, разработанные в логике ECD, чаще направлены на измерение конструктов со сложной структурой, поэтому некоторые методы проверки психометрических характеристик,

Таблица 3. Распределение индикаторов сценариев по компетенциям

Компетенция	Составляющая компетенции	Сценарий		
		Аквариум	Монстр	Динозавр
Критическое мышление	Анализ	10	—	2
	Вывод	9	2	1
Креативность	Детальность	3	12	—
	Оригинальность	—	7	—

разработанные для более просто организованных конструкторов и полученных одномерных данных (например, оценка надежности как внутренней согласованности), становятся малоинформативными.

Доказательство других аспектов валидности (например, конвергентной) выступает важным направлением дальнейших исследований.

Выборка и процедура сбора данных

Выборку составили учащиеся четвертых классов общеобразовательных школ одного региона центральной части России.

На первом этапе было проведено качественное исследование характеристик инструмента — когнитивная лаборатория (интервью с представителем целевой группы для выявления слабых сторон инструмента [33]). Когнитивные лаборатории были проведены по всем сценариям на 11 учениках третьего и четвертого классов. В результате были выявлены слабые места программирования, инструкции и логики сценария, которые устранены к началу проведения количественного исследования.

Выборку количественного исследования составили 995 учеников четвертых классов. Предварительный анализ показал, что для части респондентов (8.6%) некоторые сценарии не предъявлялись из-за технических проблем. Анализ проводился на полных профилях — 909 учеников (48.3% — девочки).

РЕЗУЛЬТАТЫ

Психометрический анализ в рамках модели КФА показал, что модель, которая включает в себя четыре латентных фактора, отражающие компетенции, и пять факторов, отражающих принадлежность

индикаторов к общему контексту, хорошо подходит данным: RMSEA = 0.037 (CI 90% 0.035–0.039); CFI = 0.942; TLI = 0.937 (χ^2 при методе оценки параметров WLSMV = 2137.171, df = 955, p < 0.001). Схема внутренней структуры инструмента представлена на рис. 4.

В рамках этой модели, составляющие КМ Вывод и Анализ связаны на уровне 0.88 (p < 0.001). Связь составляющих КР Детальность и Оригинальность оказывается на близком значении 0.89 (p < 0.001). Отметим низкую связь между составляющими КР и КМ. Оригинальность незначимо связана (p > 0.5) с как с Выводом, так и с Анализом. Детальность связана с Выводом и Анализом значимо (p < 0.001), но слабо (до 0.29).

В модель было добавлено пять факторов контекста (см. нижнюю часть рис. 4). Три фактора отражали особенности построения монстра: вид туловища монстра, вращение элементов и изменение цвета монстра. Кроме того, были дополнительно введены ковариации ошибок двух индикаторов, отражающих количество используемых элементов для построения монстра. Другие два фактора контекста отражали разные части задания Аквариум: один фактор отражал работу с текстом; другой — работу с обустройством аквариума. Кроме того, были введены ковариации ошибок для двух пар индикаторов, которые принадлежали к одной части сценария.

Стандартизированные факторные нагрузки индикаторов, отражающих КМ, варьируются от 0.26 до 0.77. Стандартизированные факторные нагрузки индикаторов, отражающих КР, находятся в схожем диапазоне от 0.21 до 0.77. Распределение факторных нагрузок по всем индикаторам представлено на рис. 5.

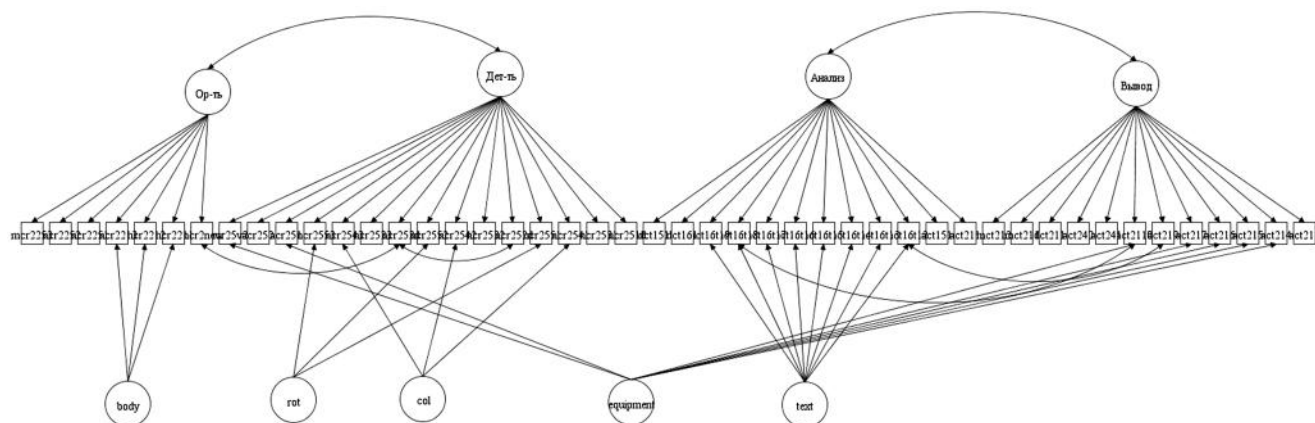


Рис. 4. Внутренняя структура инструмента 2К. В верхней части — факторы, отражающие оцениваемые конструкты, в нижней части — факторы, отражающие принадлежность к контексту

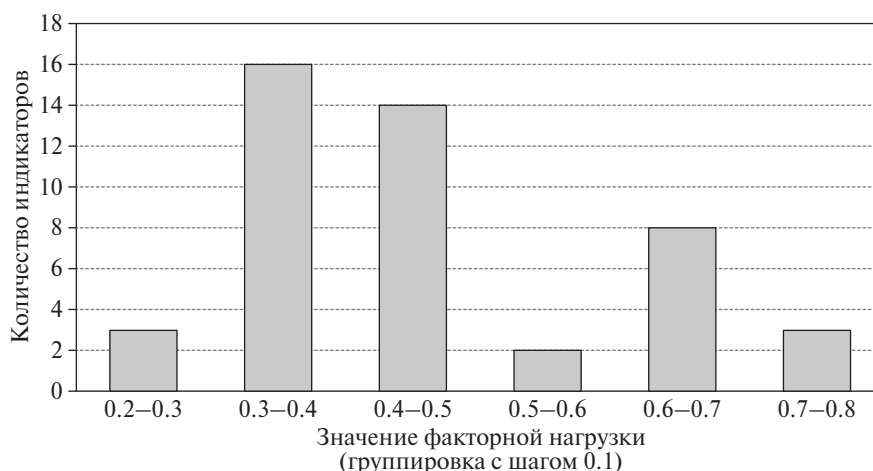


Рис. 5. Распределение стандартизованных факторных нагрузок индикаторов

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Цель статьи заключалась в том, чтобы представить методологию разработки инструмента креативности и критического мышления в начальной школе, представить особенности разработки заданий сценарного типа и привести доказательства валидности измерений через доказательство соответствия эмпирических данных и теоретически ожидаемой структуры конструкторов.

В нашей работе критическое мышление и креативность рассматриваются как компетенции, включающие в себя несколько составляющих, а в фокусе измерения — продукт деятельности тестируемого, а не самоотчет. На основе анализа литературы и технических ограничений, накладываемых выбранным способом измерения, была предложена концептуальная рамка, в которой критическое мышление включало две непосредственно связанные составляющие: Анализ (навыки работы с информацией в соответствии с поставленной задачей) и Вывод (построение собственного вывода на основе проведенного анализа). Количественный анализ подтвердил, что данные содержат два фактора, образованные индикаторами, которые разработчики закладывали как свидетельства соответствующих конструкторов. Однако важно отметить, что концептуальная рамка предполагает, что Анализ и Вывод связаны более сложным характером отношений: для того чтобы продемонстрировать высокий уровень Вывода, на первом этапе требуется достаточный уровень Анализа. Учет такого рода отношений выступает приоритетной задачей дальнейших исследований³.

На наш взгляд, больше концептуальных и психометрических вопросов вызывает часть инструмента 2К, связанная с креативностью. Согласно проведенному анализу литературы, было обнаружено больше важных составляющих этой компетенции, чем нам удалось облечь в измеряемую форму в рамках даже гибкого инструмента сценарного типа. Так, Любознательность не может быть измерена в рамках заданного формата тестирования. Также пока мы не нашли способа встроить в сценарии ситуацию, в которой тестируемый предлагал бы большое количество идей, а разработчики могли бы оценивать их связь между собой. Также на данном этапе Полезность (соответствие результата условию задачи) была изначально внедрена в сценарий, так что у респондента не было возможности создавать что-то непредусмотренное разработчиками. Так, в задании Аквариум Полезность задана тем, что ученик находится в конструкторе аквариума, так что его (ее) действия будут отнесены к внутреннему обустройству дома для крабов. В задании Монстр введена система “референсов”, чтобы мотивировать учащихся собирать именно монстра. В дальнейших исследованиях мы планируем обратить большее внимание на Полезность и разнообразить сценарии так, чтобы у респондентов появились более широкие возможности действовать в среде задания.

Согласно результатам количественного анализа, были выявлены два фактора, отражающие составляющие креативности: Детальность и Оригинальность, отражающие количество использованных элементов и новизну идеи через их расположение, что полностью соответствует ожиданиям. Также составляющие креативности показывают невысокую

³ Uglanova I. Measuring critical thinking through innovative assessment: An investigation of the dimensionality // The 20th Annual Association for Educational Assessment (AEA) — Europe

Conference Assessment for transformation: Teaching, learning and improving educational outcomes. Portugal, 2019.

корреляцию с составляющими критического мышления. Задания устроены таким образом, что конструкты тесно переплетены между собой в силу общего контекста, однако они остаются хорошо различимыми.

В инструменте используется автоматический подход к обработке результатов, без привлечения экспертов. С одной стороны, мы экономим ресурсы на проверку, снижаем субъективность оценки, имеем возможность быстро дать обратную связь. С другой — ограничены техническими возможностями фиксации действий респондентов и их малой вариативностью. Для оценки критического мышления мы не используем задания с открытым вводом текста, а ограничиваемся работой с выбором из готовых решений. Это определено в том числе и возрастом тестируемых: не все четвероклассники бегло печатают.

Для оценки креативности автоматический анализ результатов становится еще более сложной задачей. На данном этапе мы используем конструктор с заданным набором элементов, который позволяет ученикам строить уникальные объекты. В модели измерений эти объекты анализируются как сочетание отдельных частей. Автоматический анализ изображений с учетом целостности объекта — важное направление для наших дальнейших исследований.

ВЫВОДЫ

В целом эмпирически подтвержденная структура конструктов соотносится с современными подходами к изучению как креативности [21], так и критического мышления [22], а выбранный формат измерения, хотя и обладает перечисленными выше ограничениями, позволяет преодолеть недостатки методов самоотчета, субъективность и затратность экспертных оценок. Развитие технологий же позволяет надеяться на то, что описанные ограничения будут успешно преодолены в следующих поколениях инструмента. Другой линией развития представленного исследования может выступить разработка аналогичного инструмента для более старшего возраста.

СПИСОК ЛИТЕРАТУРЫ

1. Блонский П.П., Смирнов А.А. Избранные психологические произведения. М.: Просвещение, 1964. С. 546.
2. Богоявленская Д.Б. Психология творческих способностей. М.: Федоров, 2009. С. 342.
3. Ефремова Н.Ф. Оценка надежности компетенций студентов с помощью доказательной аргументации // *Problems of the development of modern science: theory and practice*. 2018. С. 197–200.
4. Зейгарник Б.В. Патопсихология. М.: Изд-во Моск. ун-та, 1986. С. 287.
5. Рубинштейн Л.С. Основы общей психологии. 2-е изд. СПб.: Питер, 2002. С. 720.
6. Теплов Б.М. Психология: Учебник для средней школы. М.: Учпедгиз, 1946. С. 223.
7. Уланова И.Л., Брун И.В., Васин Г.М. Методология Evidence-Centered Design для измерения комплексных психологических конструктов // *Современная зарубежная психология*. 2018. Т. 7. № 3. С. 18–27.
8. Эльконин Д.Б. К проблеме периодизации психического развития в детском возрасте // *Вопросы психологии*. 1971. № 4. С. 6–20.
9. Amabile T. *The Social Psychology of Creativity*. N.Y.: Springer Verlag, 1983.
10. Baer J. *Creative teachers, creative students*. Boston, MA: Allyn & Bacon, 1997. P. 173.
11. Baer J. *Creativity* // *The Oxford Handbook of Educational Psychology* / A. O'Donnell. Oxford University Press, 2018.
12. Brown T.A. *Confirmatory Factor Analysis for Applied Research*. 1st ed. N.Y.: Guilford, 2006. P. 462.
13. Dewey J. *Moral principles in education*. Houghton Mifflin, 1909. P. 60.
14. Feist G.J. The Evolved Fluid Specificity of Human Creative Talent // *Creativity: From potential to realization* / R.J. Sternberg, E.L. Grigorenko, J.L. Sibger. Washington, DC: American Psychological Association, 2004. P. 57–82.
15. Griffin P., Care E. (eds). *Assessment and teaching of 21st century skills: Methods and approach*. Springer, 2014. P. 348.
16. Guilford J.P. *Intelligence, creativity, and their educational implications*. Edits Pub., 1968. P. 229.
17. Ivcevic Z. Artistic and everyday creativity: An act-frequency approach // *The Journal of Creative Behavior*. 2007. V. 41. № 4. P. 271–290.
18. Kaufman J.C., Beghetto R.A. Beyond big and little: The four c model of creativity // *Review of General Psychology*. 2009. V. 13. № 1. P. 1–12.
19. Kaufman J.C. Beyond new and appropriate: Who decides what is creative? // *Creativity Research Journal*. 2012. V. 24. № 1. P. 83–91.
20. Kent P. Fluid intelligence: A brief history // *Applied Neuropsychology: Child*. 2017. V. 6. № 3. P. 193–203.
21. Kupers E., Lehmann-Wermser A., McPherson G., van Geert P. Children's creativity: A theoretical framework and systematic review // *Review of Educational Research*. 2019. V. 89 (1). P. 93–124.

22. *Lai E.R.* Critical thinking: A literature review research report // *Pearson's Research Reports*. 2011. V. 6. P. 40–41.
23. *Liu O.L.* Assessing critical thinking in higher education: Current state and directions for next-generation assessment // *ETS Research Report Series*. 2014. № 1. P. 1–23.
24. *Mislevy R.J.* A brief introduction to evidence-centered design // *ETS Research Report Series*. 2003. № 1. P. 1–29.
25. *Paul R.* Critical thinking: What Every Person Needs to Survive in a Rapidly Changing World. Rohnert Park, Calif.: Center for Critical Thinking and Moral Critique, Sonoma State University, 1990. P. 575.
26. *Rhodes M.* An Analysis of Creativity // *The Phi Delta Kappan*. 1961. V. 42. № 7. P. 305–310.
27. *Rosenbaum P.R.* Items bundles // *Psychometrika*. 1988. V. 53. № 3. P. 349–359.
28. *Seaman M.* Bloom's Taxonomy: its evolution, revision, and use in the field of education // *Curriculum & Teaching Dialogue*. 2011. V. 13. P. 29–35.
29. *Spaulding S.* UNESCO's world education report: Its evolution, strengths and possible futures // *International Journal of Educational Development*. 1999. V. 19. № 1. P. 53–63.
30. *Sternberg R.J.* Critical Thinking: Its Nature, Measurement, and Improvement. Washington, DC: National Inst. of Education, 1986. P. 37.
31. *Torrance E.P.* The nature of creativity as manifest in its testing // *The nature of creativity* / R.J. Sternberg. Cambridge University Press, 1988. P. 43–75.
32. *Yen W.M.* Scaling performance assessments: Strategies for managing local item dependence // *Journal of Educational Measurement*. 1993. V. 30. № 3. P. 187–213.
33. *Zucker S., Sassman C., Case B.J.* Cognitive labs. San Antonio, TX: Harcourt Assessment, 2004. V. 15. P. 10.

MEASURING CREATIVITY AND CRITICAL THINKING IN PRIMARY SCHOOL

I. L. Uglanova*, E. A. Orel, I. V. Brun*****

*National Research University Higher School of Economics, Institute of Education,
Center for Psychometrics and Measurements in Education;
101000, Moscow, Potapovskiy lane, 16, bld 10, Russia.*

**Postgraduate student, research assistant. E-mail: iuglanova@hse.ru*

***PhD in Psychology, Researcher. E-mail: eorel@hse.ru*

****Junior Researcher. E-mail: ibrun@hse.ru*

Received 02.03.2020

Abstract. The procedure for developing of a tool for evaluating creativity and critical thinking of primary school pupils performed in the logic of “systematic approach to the development of measuring tools” (Evidence-centered Design; ECD) is described in the article. The computerized test is a set of scenarios where the pupil in a playful way has to solve a problem from his/her everyday experience (to design a home aquarium, to invent a character for a fantastic story, etc.). Theoretical analysis of the measured constructs is given according to the ECD methodology, their structure is proposed, the developed scenarios as well as an empirical study ($N = 909$ fourth-grade students) are described. As a result, the validity of the developed tool including compliance with theoretical expectations about the structure of creativity and critical thinking in early school age has been confirmed.

Keywords: Evidence-centered Design, validation of measurement tools, primary school.

REFERENCES

1. *Blonskij P.P., Smirnov A.A.* Izbrannye psihologicheskie proizvedeniya. Moscow: Prosveshchenie, 1964. P. 546. (In Russian)
2. *Bogoyavlenskaya D.B.* Psihologiya tvorcheskikh sposobnostej. Moscow: Fedorov, 2009. P. 342. (In Russian)
3. *Efremova N.F.* Ocenka nadezhnosti kompetencij studentov s pomoshch'yu dokazatel'noj argumentacii. Problems of the development of modern science: theory and practice. 2018. P. 197–200. (In Russian)
4. *Zeigarnik B.V.* Patopsihologiya. Moscow: Izdatel'stvo Moskovskogo Universiteta, 1986. P. 287. (In Russian)

5. *Rubinshtejn L.S.* Osnovy obshchej psihologii: Seriya "Mastera psihologii". 2-e izd. St. Petersburg: Piter, 2002. P. 720. (In Russian)
6. *Teplov B.M.* Psihologiya: Uchebnik dlya srednej shkoly. Moscow: Uchpedgiz, 1946. P. 223. (In Russian)
7. *Uglanova I.L., Brun I.V., Vasin G.M.* Metodologiya Evidence-Centered Design dlya izmereniya kompleksnyh psihologicheskikh konstruktov. Sovremennaya zarubezhnaya psihologiya. 2018. V. 7. № 3. P. 18–27. (In Russian)
8. *El'konin D.B.* K probleme periodizacii psihicheskogo razvitiya v detskom vozraste. Voprosy psihologii. 1971. № 4. P. 6–20. (In Russian)
9. *Amabile T.* The Social Psychology of Creativity. New York: Springer Verlag, 1983.
10. *Baer J.* Creative teachers, creative students. Boston, MA: Allyn & Bacon, 1997, P. 173.
11. *Baer J.* Creativity. The Oxford Handbook of Educational Psychology. A. O'Donnell. Oxford University Press: 2018.
12. *Brown T.A.* Confirmatory Factor Analysis for Applied Research. 1st ed. New York: Guilford, 2006. P. 462.
13. *Dewey J.* Moral principles in education. Houghton Mifflin, 1909. P. 60.
14. *Feist G.J.* The Evolved Fluid Specificity of Human Creative Talent. Creativity: From potential to realization. R.J. Sternberg, E.L. Grigorenko, J.L. Sibger. Washington, DC: American Psychological Association, 2004. P. 57–82.
15. *Griffin P., Care E., editors.* Assessment and teaching of 21st century skills: Methods and approach. Springer, 2014. P. 348.
16. *Guilford J.P.* Intelligence, creativity, and their educational implications. Edits Pub., 1968. P. 229.
17. *Ivcevic Z.* Artistic and everyday creativity: An act-frequency approach. The Journal of Creative Behavior. 2007. V. 41. № 4. P. 271–290.
18. *Kaufman J.C., Beghetto R.A.* Beyond big and little: The four c model of creativity. Review of General Psychology. 2009. V. 13. № 1. P. 1–12.
19. *Kaufman J.C.* Beyond new and appropriate: Who decides what is creative? Creativity Research Journal. 2012. V. 24. № 1. P. 83–91.
20. *Kent P.* Fluid intelligence: A brief history. Applied Neuropsychology: Child. 2017. V. 6. № 3. P. 193–203.
21. *Kupers E., Lehmann-Wermser A., McPherson G., van Geert P.* Children's creativity: A theoretical framework and systematic review. Review of Educational Research. 2019. V. 89 (1). P. 93–124.
22. *Lai E.R.* Critical thinking: A literature review research report. Pearson's Research Reports. 2011. V. 6. P. 40–41.
23. *Liu O.L.* Assessing critical thinking in higher education: Current state and directions for next-generation assessment. ETS Research Report Series. 2014. № 1. P. 1–23.
24. *Mislevy R.J.* A brief introduction to evidence-centered design. ETS Research Report Series. 2003. № 1. P. 1–29.
25. *Paul R.* Critical thinking: What Every Person Needs to Survive in a Rapidly Changing World. Rohnert Park, Calif.: Center for Critical Thinking and Moral Critique, Sonoma State University, 1990. P. 575.
26. *Rhodes M.* An Analysis of Creativity. The Phi Delta Kappan. 1961. V. 42. № 7. P. 305–310.
27. *Rosenbaum P.R.* Items bundles. Psychometrika. 1988. V. 53. № 3. P. 349–359.
28. *Seaman M.* Bloom's Taxonomy: its evolution, revision, and use in the field of education. Curriculum & Teaching Dialogue. 2011. V. 13. P. 29–35.
29. *Spaulding S.* UNESCO's world education report: Its evolution, strengths and possible futures. International Journal of Educational Development. 1999. V. 19. № 1. P. 53–63.
30. *Sternberg R.J.* Critical Thinking: Its Nature, Measurement, and Improvement. Washington, DC: National Inst. of Education, 1986. P. 37.
31. *Torrance E.P.* The nature of creativity as manifest in its testing. The nature of creativity. R.J. Sternberg. Cambridge University Press, 1988. P. 43–75.
32. *Yen W.M.* Scaling performance assessments: Strategies for managing local item dependence. Journal of Educational Measurement. 1993. V. 30. № 3. P. 187–213.
33. *Zucker S., Sassman C., Case B.J.* Cognitive labs. San Antonio, TX: Harcourt Assessment. 2004. V. 15. P. 10.