

# Использование данных о процессе выполнения задания при оценке креативного мышления

**Тарасов С.В.**

ФГАОУ ВО «Национальный исследовательский университет  
«Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»),  
г. Москва, Российская Федерация  
ORCID: <https://orcid.org/0000-0003-4151-115X>, e-mail: [svtarasov@hse.ru](mailto:svtarasov@hse.ru)

Креативное мышление является важным навыком современного мира, а его оценка с помощью современных цифровых инструментов становится все более сложной методологической задачей. Включение в модель оценки креативного мышления данных о процессе выполнения заданий является перспективным направлением, которое становится возможным в компьютерном тестировании. Применение таких данных позволяет учитывать процессы креативного мышления в динамике, что делает оценку уровня креативности учеников более точной и многогранной. В исследовании представлен анализ работ 823 учеников 4 класса, которые в ходе выполнения задания создавали изображения в закрытой симуляционной среде для оценки креативного и критического мышления. В эмпирической части с помощью критерия хи-квадрат были сравнены последовательности от одного до трех действий учеников с разным уровнем сформированности креативного мышления, а также при выполнении задания на креативное и критическое мышление. В результате было показано, что данные о процессе выполнения задания могут быть использованы при проверке качества заданий и валидации инструментов измерения, а также расширяют систему подсчета баллов и обратную связь по результатам тестирования.

**Ключевые слова:** креативное мышление; процессные данные; компьютерное тестирование; N-grams.

---

**Финансирование.** Статья подготовлена в рамках гранта Министерства науки и высшего образования Российской Федерации от 25.04.2022 № 075-15-2022-325.

**Для цитаты:** Тарасов С.В. Использование данных о процессе выполнения задания при оценке креативного мышления // Психологическая наука и образование. 2023. Том 28. № 4. С. 63—80. DOI: <https://doi.org/10.17759/pse.2023280404>

# Using Process Data from Completing a Task in Creative Thinking Assessment

**Sergei V. Tarasov**

HSE University, Moscow, Russia

ORCID: <https://orcid.org/0000-0003-4151-115X>, e-mail: [svtarasov@hse.ru](mailto:svtarasov@hse.ru)

Creative thinking is an important skill in the modern world, and assessing its efficiency using modern digital tools is an increasingly complex methodological task. Using the information about the process of solving the tasks into the assessment model using the digital testing mechanisms became a promising trend. The use of such data allows us to consider the processes of creative thinking in dynamics, which makes the assessment of the level of creativity more accurate and diverse. The paper presents an analysis of the work of 823 students in the 4th grade who were asked to create images in a closed simulation environment. In this way, we analyzed their creative and critical thinking. Then the sequences of actions of students at different levels of creative thinking were compared, and various strategies they used to complete creative thinking tasks were compared with strategies used for critical thinking tasks. The information on the process helps to understand how these tasks work and to perform validation study. It also enhance the scoring rules and the feedback that can be received after the test.

**Keywords:** creative thinking; process data; computer testing; N-grams.

---

**Funding.** The reported study was funded by The Ministry of Education and Science of the Russian Federation (project number 075-15-2022-325 from 25.04.2022).

**For citation:** Tarasov S.V. Using Process Data from Completing a Task in Creative Thinking Assessment. *Psikhologicheskaya nauka i obrazovanie = Psychological Science and Education*, 2023. Vol. 28, no. 4, pp. 63—80. DOI: <https://doi.org/10.17759/pse.2023280404> (In Russ.).

## Введение

В области образования были определены некоторые важные способности, названные ключевыми компетенциями или навыками 21-го века [29], овладение которыми необходимо для успешной реализации себя в жизни. Креативность или креативное мышление — один из таких навыков. Поскольку компьютеры и различные искусственные интеллектуальные системы в настоящее время заменяют работников для выполнения многих стандартных задач [3], способности к креативному мышлению при решении задач и сложные коммуникативные и социальные навыки становятся все более ценными на рынке труда. Такая потребность бросает вызов для учителей относительно того, как эти

навыки развивать, и для психометриков относительно того, как эти навыки оценивать наилучшим образом.

В области оценки креативности в образовательной сфере широко распространен подход 4Р, представляющий четыре основные категории определений и направлений исследований креативности: продукт, процесс, личность и среда (product, process, person and press) [20]. Далее мы рассмотрим инструменты измерения когнитивных процессов, связанных с креативностью (процесс), а также творческих продуктов или результатов (продукт).

Процессный подход к измерению креативности фокусируется на специфических когнитивных процессах, которые способствуют

ют творческому развитию. Традиционные тесты, разработанные в рамках этого подхода, включают открытые или плохо структурированные задачи, требующие выработки как можно большего числа ответов, которые затем оцениваются для определения различных факторов креативности, среди которых обычно отмечают беглость (количество ответов), оригинальность (статистическая редкость), гибкость (количество различных категорий) и проработанность (количество деталей). Основной идеей оценки является не только рассмотрение количества ответов, но и их качества.

Большой вклад в оценку творческих способностей внес Гилфорд [10], разработав тесты дивергентного мышления (Structure of the Intellect Divergent Production Tests, SOI), которое считается важным элементом креативности. Также стоит отметить, что наиболее используемые тесты на креативность — это тесты творческого мышления Торренса (The Torrance Tests of Creative Thinking, ТТСТ) [27], которые были переведены более чем на 40 языков. Торренс разработал 12 тестов для разных возрастов, сгруппированных в вербальную, изобразительную и звуковую батареи. Однако данные тесты являются трудозатратными в плане обучения экспертов и администрирования тестирования. Кроме того, тесты не измеряют все аспекты творчества, хотя и не претендуют на это. В области оценки креативности также существует позиция, что никакой когнитивный тест не является предиктором творчества, если не включает аффективные и мотивационные факторы [26].

Среди новых тестов креативности в отношении школьников можно отметить оценку потенциального творчества (The Evaluation of Potential Creativity, EPoC), которая разрабатывается как процедуры, оценивающие «область перекрытия». Тест состоит из ряда субтестов, направленных для измерения как общих, так и конкретных творческих способностей в двух областях — вербальной/литературной и графической [4]. Субтесты EPoC измеряют две ключевые модели творческого познания — дивергентное и конвергентное

мышление. Задания дивергентного мышления включают создание как можно большего количества рисунков, используя простую абстрактную форму или знакомый предмет. Вербальные задания включают создание нескольких простых окончаний истории в ответ на уникальное начало или, наоборот, несколько завязок в ответ на уникальную развязку. В заданиях конвергентного мышления в графической области тестируемые создают законченный оригинальный рисунок, используя по крайней мере четыре из восьми абстрактных форм или знакомых объектов, служащих основой для их композиции. В словесно-литературной области задания конвергентного мышления необходимо создать законченный рассказ либо на основе предоставленного названия, либо на основе заданных вымышленных персонажей.

Однако традиционные подходы к тестированию креативности, такие как тестирование на бумаге с использованием карандаша, уже устаревают и обладают рядом недостатков. В первую очередь это необходимость экспертов для оценивания, что создает высокую нагрузку при проведении тестирования из-за сложности администрирования, а также вносит эффект субъективности экспертов в итоговые оценки. Кроме этого, в традиционных тестах отмечаются и психометрические проблемы, такие как непостоянные свидетельства надежности и валидности инструментов на разных выборках, влияние тестовой среды на результаты тестирования, устаревший психометрический анализ [23].

В связи с этим для оценивания сложных навыков требуется прибегать не только к традиционным форматам тестирования, но и к более современным форматам. Одним из таких форматов являются задания в цифровой среде, содержащие интерактивные элементы, которые зачастую похожи на игры по своей форме представления [15]. В качестве примеров можно привести инструменты для оценки конвергентного мышления — *BuzzWords* [13], дивергентного мышления — *Immune Defense* [16], конвергентного и дивергентного мышления одновременно — *Crea.blender* [19]. Зачастую инструменты в цифровой среде обла-

дают не только автоматическими правилами подсчета баллов за тестирование без участия экспертов, но и собирают данные о процессе выполнения задания.

Данные о процессе выполнения задания относятся к данным, полученным от тестируемых, взаимодействующих с компьютерным элементом оценки. Это записывается в лог-файлах компьютера и часто представляется в виде последовательностей событий (определенных действий в тестовой среде) с отметками времени [9]. Такие данные собираются и анализируются как в рамках оценивания сложных навыков, таких как совместное решение проблем [30], так и в современных форматах заданий, имитирующих видеоигры [14].

Например, в рамках Международной программы по оценке образовательных достижений учащихся (PISA) в 2012 году учащиеся оценивались на предмет сформированности навыков решения проблем с помощью набора интерактивных тестовых заданий [17]. Тест предназначен для измерения способности тестируемых к когнитивной обработке информации для понимания и разрешения проблемных ситуаций, когда метод решения не очевиден сразу. Изучая проблемную ситуацию и взаимодействуя с компьютерной средой, учащиеся находили фрагменты информации, которые были полезны для решения рассматриваемой проблемы. В дополнение к ответам на тестовые вопросы также записываются данные о поведении учащихся, чтобы иметь представление об их стратегиях решения проблем. Для заданий PISA по решению задач явные ответы учащегося на каждый вопрос являются данными о продукте, а серии нажатий и вводов, выполненных во время его или ее взаимодействия с вопросами, а также временные метки, связанные с каждым действием, являются данными процесса выполнения задания. Различные методы анализа данных о процессе применительно к одной из задач PISA 2012 года были показаны в работе [18], выводы которой свидетельствуют о том, что выбор методов анализа данных для анализа про-

цесса выполнения задания в насыщенной цифровой среде зависит от цели анализа и структуры данных.

Цель исследования — определить возможности использования данных о процессе выполнения задания в рамках оценки креативного мышления с помощью инструмента в цифровой среде. На основе заложенных в инструмент показателей мы получаем информацию об уровне сформированности навыка креативности, поэтому можем сравнивать между собой процессы решения задания учениками с разным уровнем. В данном инструменте, описанном подробно в следующем разделе, с помощью заданий одного формата оценивается не только креативность, но и критическое мышление.

Таким образом, в исследовании мы хотим ответить на следующие исследовательские вопросы относительно данных о процессе выполнения задания:

1. Какие действия отличают процесс выполнения задания на креативное мышление учениками с высоким уровнем креативного мышления от выполнения учениками с низким уровнем?
2. Какие действия отличают процесс выполнения задания на креативное мышление от процесса выполнения задания на критическое мышление?

## Методы

### *Инструмент*

В исследовании проводится анализ одного из заданий инструмента для оценки навыков 21-го века 4К среди учеников 4 класса, разработанного сотрудниками Центра психометрики и измерений в образовании (Лаборатория измерения новых конструкторов и дизайна тестов) Института образования НИУ ВШЭ. Инструмент состоит из нескольких заданий сценарного типа, оценивающих четыре навыка: креативное и критическое мышление, коммуникацию и кооперацию; причем большинство заданий измеряют несколько навыков.

Инструмент предьявляется тестируемому в компьютерной форме. Экран заданий интерактивен: тестируемый нажимает на

выбранную им область, после чего видит заранее подготовленную специфическую реакцию системы тестирования. Такой формат позволяет проявить сложные навыки, а также поддерживает мотивацию тестируемых и снижает тестовую тревожность. В то же время имитация реальной среды позволяет более точно фиксировать наблюдаемое поведение, то есть свидетельство того, что тестируемый обладает конкретным навыком.

В рамках данной работы рассматривалось задание «Монстр», которое направлено на оценку креативного и критического мышления. Общий контекст задания вымышленный и заключается в том, что тестируемый оказался на праздничном представлении в Городе монстров. Задание состоит из экранов с описанием и конструктора (рис. 1—2), с помощью которого тестируемый создает изображения.

В первой части этого задания, направленного на оценку креативного мышления, каждый тестируемый для афиши представления создает трех монстров, которые должны быть удивительными и необычными, а также отличаться от местных жителей. Интерфейс данной части задания представлен на рис. 1. Конструктор состоит из холста в центре, на который могут быть помещены элементы из нижней панели. Все элементы разделены на категории (туловище, руки, разное и т.д.) и могут быть использованы неограниченное количество раз (кроме туловища, которое обязательно должно быть одно и фиксируется в центре холста). Справа на экране был представлен местный житель, который для цели оценки оригинальности считался референсом. Также на экране над референсом и справа на нижней панели содержатся инструкции для выполнения задания.

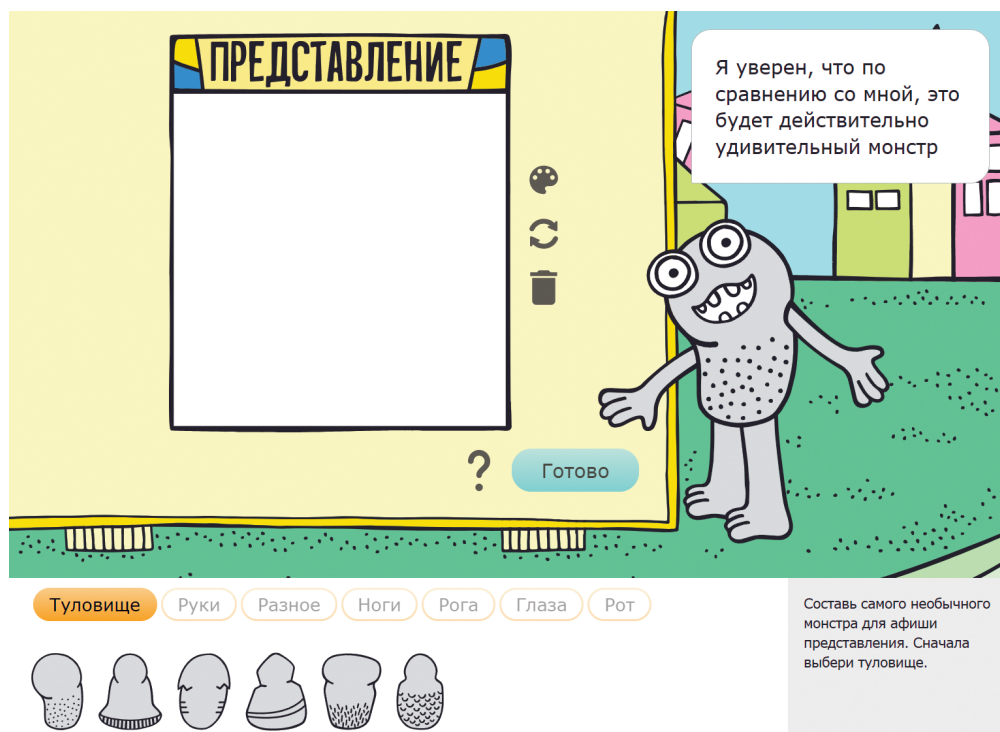


Рис. 1. Интерфейс задания «Монстр» для оценки креативного мышления

Во второй части задания оказывается, что во время представления произошло чрезвычайное происшествие и тестируемый по короткому описанию преступника должен составить его фоторобот. Интерфейс этой части задания не сильно отличается от интерфейса предыдущей части и представлен на рис. 2. Основное отличие заключается в инструкции, согласно которой тестируемому надо составить фоторобот монстра-преступника, а не необычного монстра.

**Операционализация конструкта**

Для создания инструмента измерения работчиками была определена теоретическая рамка конструкта креативности [1], которая основывается как на когнитивном подходе [10; 27], так и на концепции структурированного воображения [28]. Креативность в рамках данного инструмента включает два субконструкта:

1. **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.

Для оценки оригинальности построенного тестируемым изображения мы сравниваем его с референсом. Референс представляет из себя прокси-изображение, отражающее образ наиболее часто встречающегося в выборке. При создании референса мы выявляем наиболее типичные элементы и их количество с помощью когнитивных интервью и количественных исследований на пилотном этапе. Направленность задания на креативность определяется тем, что в рамках задания тестируемых просят создать новый образ, который должен отличаться от исходного.

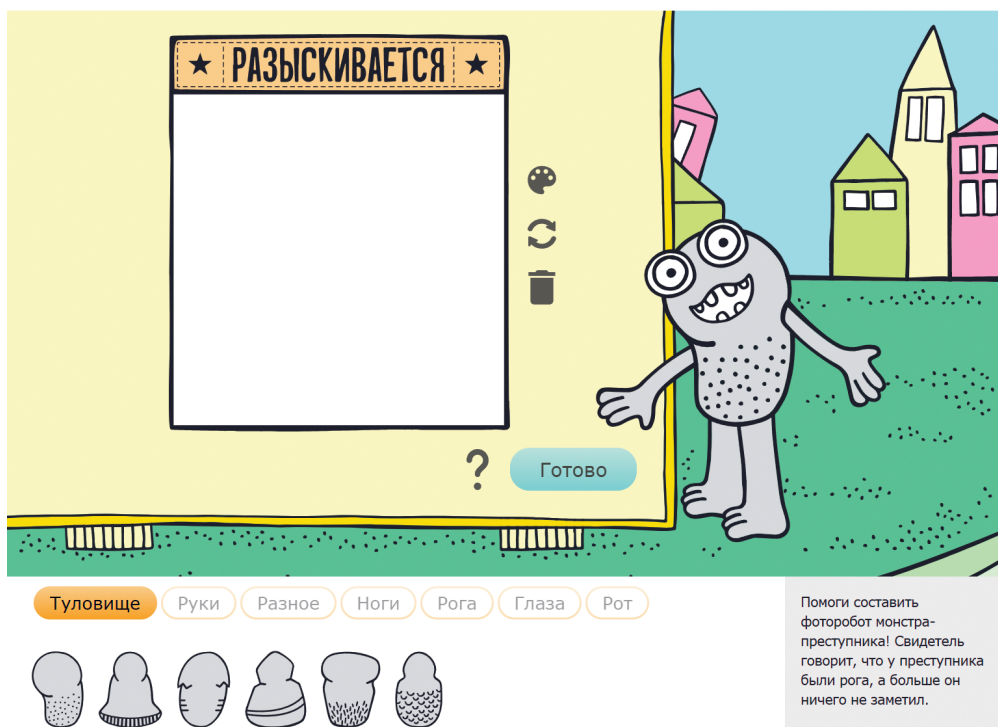


Рис. 2. Интерфейс задания «Монстр» для оценки критического мышления

2. *Детальность* — способность глубоко, с большой степенью подробности проработать предложенную идею.

Аналогично измерению проработанности в рамках когнитивного подхода в нашем исследовании детализация измеряется как количество элементов, которые используются в решении, и их функциональное разнообразие.

### **Выборка и процедура**

Тестирование проводилось весной 2022 года в нескольких городах России. Из всей выборки были отобраны ученики, которые полностью выполнили задание «Монстр» (каждое изображение содержит хотя бы 2 элемента). Таким образом, база для анализа состояла из 823 учащихся четвертых классов.

Среднее время выполнения задания составляет 15 минут. Тестирование проходило в школах под наблюдением учителя с согласия родителей в соответствии с исследовательской этикой. Результаты учеников были анонимизированными для целей анализа.

### **Показатели**

В сценарных заданиях инструмента 4К оценивание навыков происходит с помощью индикаторов. Под индикаторами понимаются изначально заложенные в задания поведенческие проявления навыка, например, выбрали или не выбрали правильное поведение в возникшей ситуации. В качестве поведенческого проявления креативности и критического мышления в задании «Монстр» мы рассматриваем особенности изображений, созданных тестируемыми. Эти особенности изображений (индикаторы) имеют предопределенные правила подсчета баллов, представленные в табл. 1. Правила подсчета баллов применялись к каждому изображению отдельно. Ниже приведены правила начисления баллов по показателям оригинальности и детальности.

Кроме значений индикаторов для каждого тестируемого фиксировался процесс построения изображения в виде записи всех совершенных тестируемым действий через

Таблица 1

### **Правила подсчета баллов для оценки креативности в задании «Монстр»**

Индикатор	Правила подсчета баллов	
	Оригинальность	
origin1	1 — количество элементов из категории Ноги отлично от 2 0 — количество элементов из категории Ноги равно 2	
origin2	1 — количество элементов из категории Руки отлично от 2 0 — количество элементов из категории Руки равно 2	
origin3	1 — количество элементов из категории Рот отлично от 1 0 — количество элементов из категории Рот равно 1	
origin4	1 — количество элементов из категории Глаза отлично от 2 0 — количество элементов из категории Глаза равно 2	
symmetLegs1	1 — отсутствие симметрии для элементов из категории Ноги 0 — есть симметрия	
symmetHands1	1 — отсутствие симметрии для элементов из категории Руки 0 — есть симметрия	
position1	1 — хоть один элемент из категории Ноги стоит не в слотах ног 0 — все элементы из категории Ноги стоят в слотах ног NA — элементы из категории Ноги не использованы вовсе	
position2	1 — хоть один элемент из категории Руки стоит не в слотах рук 0 — все элементы из категории Руки стоят в слотах рук NA — элементы из категории Руки не использованы вовсе	
position3	1 — хоть один элемент из категории Рот стоит не в верхней части туловища 0 — все элементы из категории Рот стоят в верхней части туловища NA — элементы из категории Рот не использованы вовсе	

position4	1 — хоть один элемент из категории Глаза стоит не в верхней части туловища 0 — все элементы из категории Глаза стоят в верхней части туловища NA — элементы из категории Глаза не использованы вовсе
<i>Детальность</i>	
nElements	2 = 9 и более элементов, кроме категории Разное 1 = от 5 до 8 элементов, кроме категории Разное 0 = от 0 до 4 элементов, кроме категории Разное
nVarious	2 = 2 и более элементов из категории Разное 1 = 1 элемент из категории Разное 0 = 0 элементов из категории Разное
color	1 — менял цвет монстра хотя бы один раз 0 — не менял цвет
turn	1 — поворачивал элемент хотя бы 1 раз 0 — не поворачивал элементы

запятую в виде «Start, Add\_Torso, Add\_Hands, Add\_Hands, Add\_Legs, Add\_Legs, Add\_Mouth, Color, Add\_Eyes, Finish». Описание всех возможных действий представлено в табл. 2.

### **Методология анализа данных**

Для оценки уровня креативности использовалась методология конфирматорного факторного анализа (КФА) на порядковых переменных [6]. Мы используем метод взвешенных наименьших квадратов с использованием матрицы полихорических корреляций (WLSMV) для надежной оценки из-за категориального порядка данных [7]. Соответствие модели

данным определялось на основе индекса сравнительного соответствия (CFI), индекса Такера-Льюиса (TLI) и среднеквадратичной ошибки аппроксимации (RMSEA). Первые два индекса показывают расстояние построенной модели от нулевой модели, в которой ни одна переменная не связана с другой [5], а последний основан на анализе остатков модели [25]. Мы опирались на следующие критические значения коэффициентов, следуя общепринятым правилам (Yu, 2002): CFI>0,95; TLI>0,95; RMSEA≤0,06. Для улучшения качества модели использовались индексы модификации, основанные на анализе модельных остатков.

Таблица 2

### **Действия в журнале действий**

<b>Запись в журнале действий</b>	<b>Действие</b>
Start	Начало задания
Finish	Окончание задания
Color	Изменение цвета монстра
Add_X, где X может быть: Torso Hands Legs Horns Eyes Mouth Various	Добавление элемента из категории: Туловище Руки Ноги Рога Глаза Рот Разное
Tur_X, где X аналогичен случаю с Add, кроме Torso, т.к. туловище всегда расположено по центру холста и не поворачивается	Поворот элемента из разных категорий
Del_X, где X аналогичен случаю с Add	Удаление элемента из разных категорий
Que	Обращение к справке нажатием на знак вопроса



Единицей анализа в моделях выступает индикатор. О взаимосвязи между индикатором и фактором можно говорить на основе факторной нагрузки индикатора в модели КФА. Статистически значимая и положительная факторная нагрузка указывает на наличие этой взаимосвязи, а более высокое значение факторной нагрузки указывает на более высокую связь индикатора с изучаемым фактором.

В качестве оценки уровня сформированности способности используются факторные баллы из модели КФА, рассчитанные с помощью регрессионного метода (подробнее в [8]). Факторные баллы представляют из себя стандартизованную непрерывную шкалу со средним 0 и стандартным отклонением, равным корню из квадрата множественной корреляции между всеми индикаторами и фактором. Таким образом, полученные баллы можно использовать для дальнейшего ранжирования тестируемых.

Для работы с данными процесса выполнения задания записанные переменные были разбиты на N-граммы — небольшие последовательности определенных действий, где N — количество действий в этой последовательности. Самыми короткими последовательностями являются уни-граммы, то есть последовательности из одного действия, также в работе используются би-граммы и три-граммы — последовательности из двух и трех действий соответственно. Если тестируемый в процессе выполнения задания совершает 16 действий, то эта последовательность содержит 16 уни-грамм, 15 би-грамм и 14 три-грамм. N-граммы могут повторяться как между тестируемыми, так и в рамках выполнения задания одним тестируемым. На основе табл. 2 можно предположить, что в задании будет встречаться 24 уникальных уни-грамм, а уни-граммы «Start» и «Finish» должны встречаться во всех заданиях. В связи с различной частотой встречаемости определенные последовательности вносят разный вклад в различение групп тестируемых, поэтому принятой практикой при работе с N-граммами является их взвешивание [12] по формуле:

$$w(i, j) = \begin{cases} (1 + \log(n_{i,j})) * \log\left(\frac{N}{n_i}\right), & \text{если } n_{i,j} > 0 \\ 0, & \text{если } n_{i,j} = 0 \end{cases}$$

где  $i, j$  — определенное действие  $i$  (N-грамм) в определенной последовательности  $j$ ,

$N$  — общее количество последовательностей,

$n_{i,j}$  — частота действия  $i$  в последовательности  $j$ ,

$n_i$  — частота действия  $i$  во всех последовательностях.

Для сравнения процесса выполнения задания мы рассчитываем взвешенную частоту N-грамм в разных подгруппах, чтобы определить, насколько одна группа отличается от другой с точки зрения частот действий. Согласно нулевой гипотезе два набора случайным образом эквивалентны, поэтому распределение их действий пропорционально друг другу. Для оценки отклонения от этой нулевой гипотезы вычисляется критерий хи-квадрат ( $\chi^2$ ) (подробнее в [2]). N-граммы с более высокими значениями  $\chi^2$  являются теми последовательностями действий, которые отличают процесс выполнения задания в выделенных подгруппах. На уровне значимости 0,05 критическое значение  $\chi^2$  равно 3,84, то есть если наблюдаемое значение превышает его, то мы можем быть на 95% уверены, что действие происходило чаще в одной из двух подгрупп. Чем больше значение  $\chi^2$ , тем большие различия между подгруппами.

Все расчеты были произведены с помощью языка программирования для статистической обработки данных R версии 4.3.1 с использованием пакета lavaan [21] для оценки моделей КФА и пакета ngram [24] для расчета N-грамм.

## Результаты

### *Расчет уровня креативности*

Сначала были построены модели конфирматорного факторного анализа для определения уровня креативности согласно заложенным в задание индикаторам. Были построены две отдельные модели для субконструктов Оригинальность и Детальность. Показатели качества обеих моделей представлены в табл. 3 и говорят о хорошем соответствии модели данным.

Таблица 3  
**Показатели соответствия модели**

Индекс	Оригинальность	Детальность
CFI	0,980	0,973
TLI	0,979	0,961
RMSEA	0,039	0,045

На рис. 3 показана факторная структура модели для субконструкта Оригинальность. В модели для лучшего соответствия данным для всех трех изображений были добавлены корреляции между индикаторами симметрии и позиции рук, симметрии и оригинального количества рук, симметрии и оригинального количества ног.

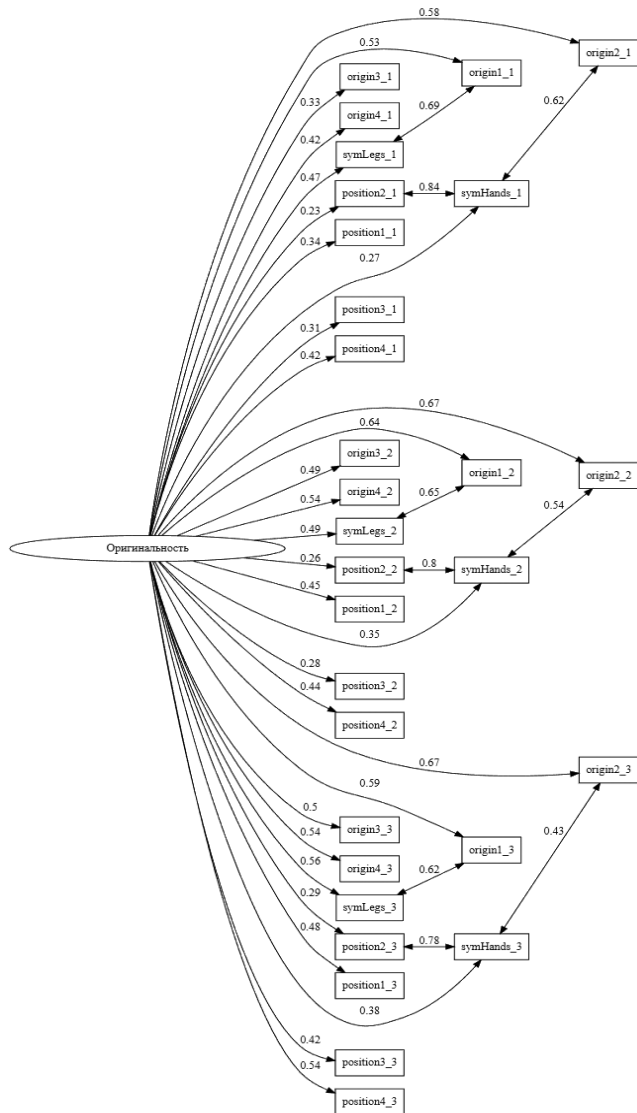


Рис. 3. Факторная структура субконструкта Оригинальность

Все стандартизированные факторные нагрузки значимы ( $p < 0,05$ ) и находятся в диапазоне от 0,22 до 0,67 при среднем значении 0,45, что является приемлемым значением.

На рис. 4 показана факторная структура модели для субконструкта Детальность. В данной модели были также добавлены корреляции между связанными между собой индикаторами, а именно — поворот, количество элементов без категории Разное и количество элементов из категории Разное.

Все стандартизированные факторные нагрузки значимы ( $p < 0,05$ ) и находятся в диапазоне от 0,16 до 0,83 при среднем значении 0,41.

**Сравнение процесса выполнения задания на креативное мышление учениками с высоким и низким уровнем креативного мышления**

Для ответа на первый исследовательский вопрос мы рассмотрели первые изобра-

жения, которые создавались тестируемыми при решении задания, направленного на оценку креативного мышления. Были построены уни-граммы, би-граммы и три-граммы, так как более длинные последовательности не обладают высокой частотой в данных и сложны в интерпретации. В табл. 4—5 приведены значения хи-квадрат при сравнении двух групп тестируемых по уровню сформированности двух субконструктов креативности (оригинальность и детальность), причем в группу с высоким уровнем были отобраны 200 учеников с наивысшим факторным баллом по соответствующему субконструкту, а в группу с низким уровнем — 200 учеников с наименьшим факторным баллом.

Среди уни-грамм нет существенных различий между учениками с высоким и низким уровнем оригинальности. На основе би-грамм можно заключить, что ученики с высоким уровнем чаще используют элементы из ка-

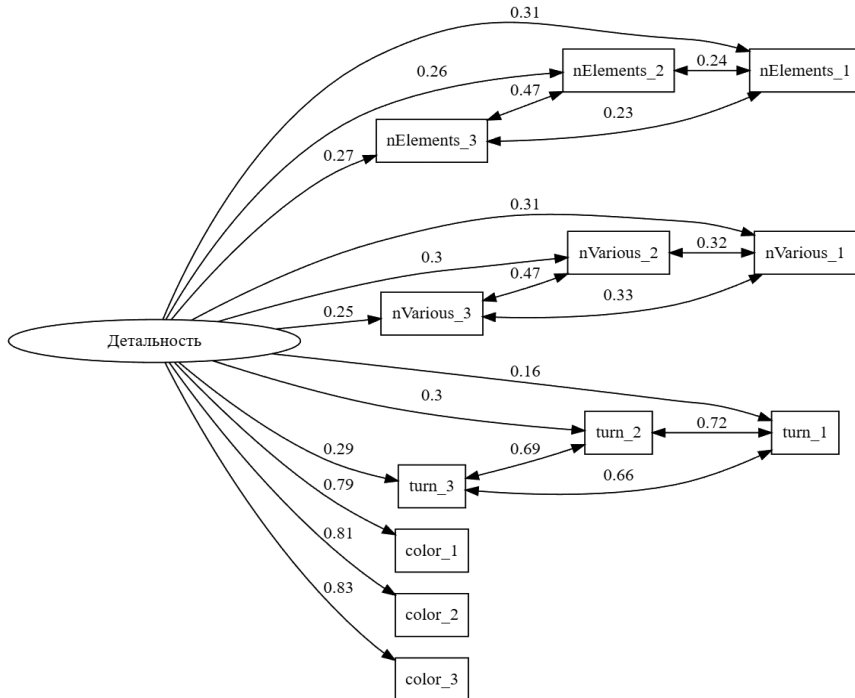


Рис. 4. Факторная структура субконструкта Детальность

Таблица 4

**Различие частоты N-грам в процессе выполнения заданий на креативное мышление у учеников с высоким и низким уровнем оригинальности**

Высокий уровень		Низкий уровень	
Действие	$\chi^2$	Действие	$\chi^2$
<i>Уни-граммы (<math>\chi^2 &gt; 5</math>)</i>			
Tur_Various	5,3	Del_Mouth	8
		Tur_Legs	6,7
		Tur_Mouth	6,5
<i>Би-граммы (<math>\chi^2 &gt; 10</math>)</i>			
Add_Various Add_Eyes	26,1	Tur_Hands Color	17
Add_Various Add_Various	17,1	Add_Eyes Add_Eyes	14,6
Add_Mouth Add_Mouth	15,1	Add_Mouth Tur_Mouth	12,9
Del_Legs Add_Eyes	15,1	Tur_Mouth Tur_Mouth	12,9
Add_Horns Add_Legs	14,8	Color Add_Horns	11,3
Add_Legs Finish	14,2	Tur_Horns Add_Various	10,8
Add_Various Del_Hands	10,3	Tur_Hands Add_Mouth	10,2
Color Del_Legs	10,3		
Del_Horns Add_Various	10,3		
Del_Horns Finish	10,3		
Del_Legs Color	10,3		
Tur_Horns Add_Legs	10,3		
<i>Три-граммы (<math>\chi^2 &gt; 20</math>)</i>			
Add_Hands Add_Hands Add_Hands	33	Add_Mouth Add_Eyes Add_Eyes	24,5
Add_Eyes Add_Eyes Add_Eyes	27,1	Add_Eyes Add_Eyes Finish	20
Add_Hands Add_Legs Finish	27,1	Del_Torso Add_Torso Add_Legs	20
Add_Legs Add_Various Add_Hands	27,1		
Add_Mouth Add_Legs Finish	23,3		

теории Разное, что является логичным, так как в данной категории находятся уникальные элементы, которые создают оригинальность итогового образа. Также у учеников с высоким уровнем оригинальности встречаются би-граммы, включающие действие удаления элемента, что может говорить о творческом процессе, когда ученик пробует разные элементы, чтобы создать оригинальный образ. Среди три-грам наибольшая разница между учениками с разным уровнем оригинальности наблюдается у последовательностей добавления трех элементов из категории (руки и глаза). Это подтверждает существующую систему подсчета баллов за задание, так как для этих категорий количество элементов на референсе равно двум, и если тестируемый

добавляет сразу три элемента, то это ассоциируется с большим баллом за оригинальность.

Ученики с высоким уровнем детальности значительно чаще меняют цвет монстра и поворачивают элементы из разных категорий. Это согласовывается с системой подсчета баллов за детальность, так как эти действия ассоциированы с более высокими баллами за индикаторы детальности. В то же время ученики с низким уровнем детальности удаляют элементы из разных категорий, что уменьшает общее количество элементов монстра и говорит о низком уровне детальности тестируемого. Кроме этого, наблюдается, что уни-грам в виде обращения к справке чаще встречается у учеников с низким уровнем детальности. Это можно

Таблица 5

**Различие частоты N-грам в процессе выполнения заданий на креативное мышление у учеников с высоким и низким уровнем детальности**

Высокий уровень		Низкий уровень	
Действие	$\chi^2$	Действие	$\chi^2$
<i>Уни-граммы (<math>\chi^2 &gt; 4</math>)</i>			
Color	64,4	Del_Torso	18,1
Tur_Horns	27	Que	18
Tur_Eyes	9,4	Add_Mouth	5,9
Tur_Various	8,6	Del_Eyes	4,8
Tur_Legs	7,2	Del_Hands	4,6
<i>Би-граммы (<math>\chi^2 &gt; 35</math>)</i>			
Add_Torso Color	75,8	Add_Torso Add_Mouth	75,5
Color Color	59,6	Add_Torso Add_Eyes	73
Color Finish	47,1	Add_Torso Add_Hands	70,7
Color Add_Eyes	44,9	Add_Legs Finish	65
Color Add_Legs	43,4	Add_Hands Finish	62,2
Color Add_Hands	40,9	Add_Hands Add_Legs	58,4
Add_Horns Tur_Horns	39,5	Add_Eyes Finish	57
Color Add_Various	39,2	Add_Torso Add_Legs	47,9
Color Add_Mouth	35,6		
<i>Три-граммы (<math>\chi^2 &gt; 25</math> для высокого уровня и <math>\chi^2 &gt; 45</math> для низкого уровня)</i>			
Start Add_Torso Color	72,8	Start Add_Torso Add_Mouth	88,9
Add_Torso Color Color	42,6	Start Add_Torso Add_Hands	83,6
Add_Horns Tur_Horns Tur_Horns	37,2	Add_Hands Add_Hands Finish	74,3
Color Add_Legs Add_Legs	35,9	Add_Torso Add_Hands Add_Hands	72,4
Color Add_Hands Add_Hands	34	Start Add_Torso Add_Legs	66,7
Tur_Horns Tur_Horns Tur_Horns	32,5	Start Add_Torso Add_Eyes	61,9
Add_Horns Add_Horns Tur_Horns	30	Add_Torso Add_Legs Add_Legs	59
Add_Torso Color Add_Legs	30	Add_Hands Add_Hands Add_Legs	55,8
Color Add_Eyes Add_Eyes	30	Add_Mouth Add_Eyes Finish	51,2
Add_Torso Color Add_Eyes	26,7	Add_Torso Add_Mouth Add_Hands	51,2
Color Color Color	26,7	Add_Legs Add_Legs Finish	50,6
		Add_Torso Add_Eyes Add_Eyes	48,9

объяснить тем, что такие ученики, возможно, не поняли, как работать в конструкторе и поэтому не смогли построить монстра и проявить свою креативность.

**Сравнение процесса выполнения заданий на креативное и критическое мышление**

Для ответа на второй исследовательский вопрос были взяты первые изображения,

которые создавались тестируемыми при решении двух частей задания: для оценки креативного и критического мышления. Как и в предыдущем пункте, были построены уни-граммы, би-граммы и три-граммы, но среди три-грамм были отобраны те, которые встречаются хотя бы 10 раз во всех решениях, чтобы убрать очень редкие последовательности, даже если они позволяют различать решения. В табл. 6 приведены значения хи-

квадрат при сравнении двух групп монстров (были отобраны N-граммы с наибольшим значением хи-квадрат в каждой из групп).

Среди уни-грамм прослеживаются четкие отличия между решениями. В задании на креативное мышление тестируемые значимо чаще удаляют элементы, причем во всех 7 категориях. В задании на критическое мышление тестируемые значительно чаще добавляют элементы из категории Рога, что объясняется условиями задачи, т.к. при составлении фоторобота преступника, который, по словам очевидцев, имел рога, добавление элементов из этой категории

показывает, что тестируемый правильно понял условие задания и проявил критическое мышление. Но в то же время в задании на критическое мышление тестируемые чаще добавляют и элементы из других категорий, а также поворачивают руки и рога. На основе би-грамм и три-грамм можно говорить о том, что в задании на креативное мышление ученики чаще меняют цвет монстра, а в задании на критическое мышление ученики оставляют его цвет изначальным (серым). Также они удаляют и поворачивают элементы из разных категорий, что свидетельствует о творческом процессе, в то время как при

Таблица 6

### Различие частоты N-грамм в процессе выполнения заданий на креативное и критическое мышление

Креативное мышление		Критическое мышление	
Действие	$\chi^2$	Действие	$\chi^2$
<i>Уни-граммы (<math>\chi^2 &gt; 15</math>)</i>			
Del_Legs	101,8	Add_Horns	73,8
Del_Hands	52,5	Tur_Hands	50,6
Del_Mouth	38,5	Tur_Horns	23,1
Del_Eyes	34,2	Add_Mouth	21,9
Del_Torso	30,3	Add_Eyes	20
Del_Horns	27,6	Add_Hands	18,3
Del_Various	24	Add_Legs	17,2
<i>Би-граммы (<math>\chi^2 &gt; 60</math>)</i>			
Del_Hands Add_Hands	111,1	Add_Torso Add_Horns	238,4
Del_Legs Add_Legs	96,6	Add_Torso Add_Eyes	88,7
Add_Legs Del_Legs	88,2	Add_Eyes Finish	83,1
Color Color	79,4	Add_Horns Finish	72,1
Del_Legs Del_Legs	78,2	Add_Legs Finish	70,6
Add_Hands Del_Hands	64,7		
<i>Три-граммы (<math>\chi^2 &gt; 60</math>)</i>			
Add_Torso Color Color	106,1	Start Add_Torso Add_Horns	187,8
Del_Legs Add_Legs Add_Legs	103,9	Add_Torso Add_Horns Add_Horns	109,9
Add_Hands Del_Hands Add_Hands	99,4	Add_Torso Add_Horns Add_Mouth	100
Add_Legs Add_Legs Del_Legs	85,5	Start Add_Torso Add_Eyes	79
Del_Hands Add_Hands Add_Hands	78,6	Add_Torso Add_Horns Add_Legs	74,6
Add_Legs Del_Legs Add_Legs	77	Add_Legs Add_Eyes Finish	68,4
Del_Hands Add_Hands Tur_Hands	66,2		
Add_Torso Color Add_Legs	63		
Tur_Legs Tur_Legs Tur_Legs	60,9		
Color Add_Legs Add_Legs	60,8		

составлении фоторобота ученики стараются просто собрать монстра из деталей, поэтому после взвешивания действия, связанные с добавлением элементов, преобладают в процессе выполнения задания на критическое мышление.

### **Обсуждение результатов и заключение**

Целью данной работы являлось определение возможностей использования данных о процессе выполнения задания в рамках оценки креативного мышления. Был рассмотрен инструмент тестирования, реализуемый в цифровой среде и состоящий из двух частей: для оценки креативного и критического мышления.

На первом этапе анализа были построены модели КФА для получения оценок тестируемых по субконструктам креативного мышления. Далее эти оценки были использованы для выделения групп тестируемых с высоким и низким уровнем оригинальности и детальности.

На последующем этапе для анализа данных о процессе выполнения задания был выбран метод N-грам, небольших последовательностей действий, извлеченных из журнала действий тестируемого. Были использованы последовательности от одного до трех действий.

Мы сравнили с помощью критерия  $\chi^2$  частоту встречаемости различных N-грам у учеников с высоким и низким уровнем креативного мышления. В результате был выделен ряд N-грам, которые ассоциируются с разным уровнем креативного мышления, например, использование элемента из категории Разное свидетельствует о высоком уровне оригинальности, а удаление элемента из категории Туловище — о низком уровне детальности.

В рамках второго исследовательского вопроса мы сравнили выполнение заданий на креативное и критическое мышление, реализованных в одном и том же конструкторе. На основе частоты встречаемости определенных N-грам были получены доказательства того, что разные задания работают со-

ответствующим образом. Так, в задании на креативное мышление ученики чаще удаляют и поворачивают элементы, меняют цвет изображения, а в задании на критическое мышление ученики чаще добавляют те элементы, которые необходимы для правильного решения задания.

Таким образом, в качестве вывода можно отметить, что выбранный метод анализа данных о процессе выполнения задания оказался полезен и может быть использован для понимания поведения тестируемых с разным уровнем развития навыка (в данном исследовании — креативного мышления) и при проверке качества заданий.

Использование данных о процессе выполнения задания также может обогатить систему подсчета баллов и обратную связь по результатам тестирования. Например, при анализе учеников с высоким и низким уровнем оригинальности была обнаружена связь наличия в профиле тестируемого биграмм, состоящих из удаления и последующего добавления элементов, с высоким уровнем оригинальности. Данная связь никак не учитывается в текущей системе начисления баллов. Однако необходимо уделять особое внимание предполагаемому использованию таких данных. Это важно, поскольку расширение в сторону накопления более сложных данных может бросить вызов традиционным подходам к масштабированию результатов оценки в рамках образовательного тестирования и может быть обработано неадекватно пользователями тестов [22].

На основе анализа процесса выполнения задания мы получили, что на результат выполнения задания на креативное мышление могут влиять и другие характеристики тестируемых, например, их уровень цифровой грамотности. Это можно подкрепить тем, что ученики с низким уровнем детальности значимо чаще учеников с высоким уровнем обращались к справке по заданию, в которой содержится обучение по работе с конструктором. Можно предположить, что из-за низкого уровня цифровой грамотности они не разобрались с тем, как выполнять задание на оценку креативного мышления в цифро-

вой среде. Данное наблюдение описывает одно из ограничений современных инструментов оценивания, а именно — возможное вмешательство других факторов.

К ограничениям данного исследования можно отнести небольшую выборку для анализа, из-за чего часть N-грамм были очень редкими в совокупности решений. Также по этой причине на этапе анализа мы объединили действия по категориям (например, добавление или поворот руки), а не рассматривали конкретные элементы внутри категории (добавление руки № 1 или поворот руки № 2). Анализ на уровне отдельных элементов находится в наших планах по уточнению работы самого конструктора, с помощью которого тестируемые создают изображения. Есть гипотеза, что определенные элементы внутри категорий могут быть связаны с уровнем креативности, например, тестируемые с высоким уровнем оригинальности будут склонны добавлять руку № 3, а не другие руки.

Среди дальнейших направлений исследования следует отметить анализ данных о

времени выполнения задания, а именно — времени между определенными действиями в последовательности выполнения задания, а также использование других подходов для анализа данных о процессе решения заданий в цифровой среде. Одним из таких подходов является метод самой длинной общей последовательности (Longest Common Subsequence) [11] — метод анализа последовательностей, используемый в обработке естественного языка и биостатистике для понимания стратегии тестируемых при решении цифровых задач. Данный подход может быть применен к рассматриваемому в статье заданию «Монстр», так как среди решений можно выделить неоригинальные решения (повторяющие референс) и считать, насколько решения учеников будут отличаться от такого решения. Также в рамках этого подхода возможен расчет близости решений разных учеников, что позволит выделять некие кластеры решений, которые можно использовать для выводов о креативном мышлении тестируемых.

### Литература

1. Углова И.Л., Орел Е.А., Брун И.В. Измерение креативности и критического мышления в начальной школе // Психологический Журнал. 2020. № 6(41). С. 96—107.
2. Agresti A. Categorical data analysis. New York, NY: John Wiley & Sons, Inc, 1990.
3. Autor D.H., Levy F., Murnane R.J. The skill content of recent technological change: an empirical exploration // Quarterly Journal of Economics. 2003. Vol. 118. P. 1279—1333. DOI:10.1162/003355303322552801
4. Barbot B., Besançon M., Lubart T. The generality-specificity of creativity: Exploring the structure of creative potential with EPoC // Learning and Individual Differences. 2016. Vol. 52. P. 178—187.
5. Bentler P.M. Comparative fit indexes in structural models // Psychological Bulletin. 1990. Vol. 107(2). P. 238—246. DOI:10.1037/0033-2909.107.2.238
6. Bock R.D., Gibbons R., Muraki E. Full-information item factor analysis // Applied psychological measurement. 1988. Vol. 12(3). P. 261—280.
7. Brown T.A., Moore M.T. Confirmatory factor analysis // Handbook of structural equation modeling. 2012. P. 361—379.
8. DiStefano C., Zhu M., Mindrila D. Understanding and Using Factor Scores: Considerations for the Applied Researcher // Practical Assessment, Research, and Evaluation. 2009. Vol. 14, Article 20. DOI:10.7275/da8t-4g52
9. Griffin P., Care E. Assessment and teaching of 21st century skills: Methods and approach. Springer, 2014.
10. Guilford J.P. The nature of human intelligence. McGraw-Hill, 1967.
11. He Q., Borgonovi F., Paccagnella M. Leveraging process data to assess adults' problem-solving skills: Identifying generalized behavioral patterns with sequence mining // Computers and Education. 2021. Vol. 166, 104170. DOI:10.1016/j.compedu.2021.104170
12. He Q., von Davier M. Analyzing process data from problem-solving items with N-grams: Insights from a computer-based large-scale assessment / In R. Yigal, F. Steve, M. Maryam (Eds.). // Handbook of research on technology tools for real-world skill development. Hershey, PA: Information Science Reference, 2016. P. 749—776.
13. Howell S., Veale T. Designing serious games with linguistic resources // Proceedings of the 4th International Conference on Foundations of Digital Games. 2009. P. 291—298. DOI:10.1145/1536513.1536564
14. Kerr D., Chung G.K.W.K., Iseli M.R. The feasibility of using cluster analysis to examine log data from educational video games // CRESST Report 790. 2011.



15. Kim Y.J., Shute V.J. Opportunities and challenges in assessing and supporting creativity in video games // In G. P. Green, & J. C. Kaufman (Eds.) // *Video games and creativity*. Elsevier Academic Press. 2015. P. 99—117. DOI:10.1016/B978-0-12-801462-2.00005-9
16. Krebs E., Jaschek C., von Thienen J., Borchart K.-P., Meinel C., Kolodny O. Designing a Video Game to Measure Creativity // 2020 IEEE Conference on Games (CoG). 2020. P. 407—414. DOI:10.1109/CoG47356.2020.9231672
17. OECD. "Problem-Solving Framework" // In PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy, OECD Publishing, Paris. 2013. DOI:10.1787/9789264190511-6-en
18. Qiao X., Jiao H. Data Mining Techniques in Analyzing Process Data: A Didactic // *Frontiers in Psychology*. 2018. № 9:2231. DOI:10.3389/fpsyg.2018.02231
19. Rafner J., Hjorth A., Risi S., Philippsen L., Dumas C., Biskjær M.M., Noy L., Tylén K., Bergenholtz C., Lynch J., Zana B., Sherson J. crea.blender: A Neural Network-Based Image Generation Game to Assess Creativity // In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. Association for Computing Machinery. 2020. P. 340—344. DOI:10.1145/3383668.3419907
20. Rhodes M. An Analysis of Creativity // *The Phi Delta Kappan*. 1961. Vol. 42(7). P. 305—310.
21. Rosseel Y. lavaan: An R Package for Structural Equation Modeling // *Journal of Statistical Software*. 2012. Vol. 48(2). P. 1—36. DOI:10.18637/jss.v048.i02
22. Rupp A. Feature selection for choosing and assembling measurement models: A building-block-based organisation // *International Journal of Testing*. 2002. Vol. 2(3/4). P. 311—360.
23. Said-Metwaly S., Van den Noortgate W., Kyndt E. Methodological issues in measuring creativity: A systematic literature review // *Creativity. Theories-Research-Applications*. 2017. Vol. 4(2). P. 276—301.
24. Schmidt D., Heckendorf C. "ngram: Fast n-Gram Tokenization." R package version 3.2.2. 2022.
25. Shi D., Maydeu-Olivares A., Rosseel Y. Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA // *Structural Equation Modeling: A Multidisciplinary Journal*. 2020. Vol. 27(1). P. 1—15.
26. Sternberg R.J., Lubart T.I. *Defying the crowd: Cultivating creativity in a culture of conformity* // New York: Free Press, 1995
27. Torrance E.P. *Guiding creative talent* // Prentice-Hall, Inc. 1962. DOI:10.1037/13134-000
28. Ward T.B. Structured Imagination: the Role of Category Structure in Exemplar Generation // *Cognitive Psychology*. 1994. Vol. 27(1). P. 1—40. DOI:10.1006/cogp.1994.1010
29. Schwab K. *The Global Competitiveness Report* // World Economic Forum. 2018.
30. Yuan J., Xiao Y., Liu H. Assessment of Collaborative Problem Solving Based on Process Stream Data: A New Paradigm for Extracting Indicators and Modeling Dyad Data // *Frontiers in Psychology*. 2019. № 10:369. DOI:10.3389/fpsyg.2019.00369

## References

1. Uglanova I.L., Orel E.A., Brun I.V. Izmerenie kreativnosti i kriticheskogo myshlenija v nachal'noj shkole [Measuring creativity and critical thinking in primary school]. *Psihologicheskij Zhurnal [Psychological Journal]*. 2020, no. 6(41), pp. 96—107. (In Russ.)
2. Agresti A. *Categorical data analysis*. New York, NY: John Wiley & Sons, Inc, 1990.
3. Autor D.H., Levy F., Murnane R.J. The skill content of recent technological change: an empirical exploration. *Quarterly Journal of Economics*, 2003. Vol. 118, pp. 1279—1333. DOI:10.1162/003355303322552801
4. Barbot B., Besançon M., Lubart T. The generality-specificity of creativity: Exploring the structure of creative potential with EPoC. *Learning and Individual Differences*, 2016. Vol. 52, pp. 178—187.
5. Bentler P.M. Comparative fit indexes in structural models. *Psychological Bulletin*, 1990. Vol. 107(2), pp. 238—246. DOI:10.1037/0033-2909.107.2.238
6. Bock R.D., Gibbons R., Muraki E. Full-information item factor analysis. *Applied psychological measurement*, 1988. Vol. 12(3), pp. 261—280.
7. Brown T.A., Moore M.T. Confirmatory factor analysis. *Handbook of structural equation modeling*, 2012, pp. 361—379.
8. DiStefano C., Zhu M., Mindrila D. Understanding and Using Factor Scores: Considerations for the Applied Researcher. *Practical Assessment, Research, and Evaluation*, 2009. Vol. 14, no. 20. DOI:10.7275/da8t-4g52
9. Griffin P., Care E. *Assessment and teaching of 21st century skills: Methods and approach*. Springer, 2014.
10. Guilford J.P. *The nature of human intelligence*. McGraw-Hill, 1967.
11. He Q., Borgonovi F., Paccagnella M. Leveraging process data to assess adults' problem-solving skills: Identifying generalized behavioral patterns with sequence mining. *Computers and Education*, 2021. Vol. 166, 104170. DOI:10.1016/j.compedu.2021.104170
12. He Q., von Davier M. Analyzing process data from problem-solving items with N-grams: Insights from a computer-based large-scale assessment. In R. Yigal, F. Steve, M. Maryam (Eds.). *Handbook of research on technology tools for real-world skill development*. Hershey, PA: Information Science Reference, 2016, pp. 749—776.
13. Howell S., Veale T. Designing serious games with linguistic resources. *Proceedings of the 4th International*

- Conference on Foundations of Digital Games*, 2009, pp. 291—298. DOI:10.1145/1536513.1536564
14. Kerr D., Chung G.K.W.K., Iseli M.R. The feasibility of using cluster analysis to examine log data from educational video games. *CRESST Report 790*, 2011.
15. Kim Y.J., Shute V.J. Opportunities and challenges in assessing and supporting creativity in video games. In G.P. Green, & J.C. Kaufman (Eds.). *Video games and creativity*. Elsevier Academic Press, 2015, pp. 99—117. DOI:10.1016/B978-0-12-801462-2.00005-9
16. Krebs E., Jaschek C., von Thienen J., Borchart K.-P., Meinel C., Kolodny O. Designing a Video Game to Measure Creativity. *2020 IEEE Conference on Games (CoG)*, 2020, pp. 407—414. DOI:10.1109/CoG47356.2020.9231672
17. OECD. “Problem-Solving Framework”. In *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, OECD Publishing, Paris, 2013. DOI:10.1787/9789264190511-6-en
18. Qiao X., Jiao H. Data Mining Techniques in Analyzing Process Data: A Didactic. *Frontiers in Psychology*, 2018, no. 9:2231. DOI:10.3389/fpsyg.2018.02231
19. Rafner J., Hjorth A., Risi S., Philipsen L., Dumas C., Biskjær M.M., Noy L., Tylén K., Bergenholtz C., Lynch J., Zana B., Sherson J. crea.blender: A Neural Network-Based Image Generation Game to Assess Creativity. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. Association for Computing Machinery, 2020, pp. 340—344. DOI:10.1145/3383668.3419907
20. Rhodes M. An Analysis of Creativity. *The Phi Delta Kappan*, 1961. Vol. 42(7), pp. 305—310.
21. Rosseel Y. IJG: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 2012. Vol. 48(2), pp. 1—36. DOI:10.18637/jss.v048.i02
22. Rupp A. Feature selection for choosing and assembling measurement models: A building-block-based organisation. *International Journal of Testing*, 2002. Vol. 2(3/4), pp. 311—360.
23. Said-Metwaly S., Van den Noortgate W., Kyndt E. Methodological issues in measuring creativity: A systematic literature review. *Creativity. Theories-Research-Applications*, 2017. Vol. 4(2), pp. 276—301.
24. Schmidt D., Heckendorf C. “ngram: Fast n-Gram Tokenization.” R package version 3.2.2, 2022.
25. Shi D., Maydeu-Olivares A., Rosseel Y. Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. *Structural Equation Modeling: A Multidisciplinary Journal*, 2020. Vol. 27(1), pp. 1—15.
26. Sternberg R.J., Lubart T.I. *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press, 1995
27. Torrance E.P. *Guiding creative talent*. Prentice-Hall, Inc., 1962. DOI:10.1037/13134-000
28. Ward T.B. Structured Imagination: the Role of Category Structure in Exemplar Generation. *Cognitive Psychology*, 1994. Vol. 27(1), pp. 1—40. DOI:10.1006/cogp.1994.1010
29. Schwab K. *The Global Competitiveness Report*. World Economic Forum, 2018.
30. Yuan J., Xiao Y., Liu H. Assessment of Collaborative Problem Solving Based on Process Stream Data: A New Paradigm for Extracting Indicators and Modeling Dyad Data. *Frontiers in Psychology*, 2019, no. 10:369. DOI:10.3389/fpsyg.2019.00369

### Информация об авторах

Тарасов Сергей Владимирович, аспирант, стажер-исследователь Центра психометрики и измерений в образовании, Институт образования, ФГАОУ ВО «Национальный исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0003-4151-115X>, e-mail: [svtarasov@hse.ru](mailto:svtarasov@hse.ru)

### Information about the authors

Sergei V. Tarasov, PhD Student, Research Assistant at Center for Psychometrics and Measurement in Education, Institute of Education, HSE University, Moscow, Russia, ORCID: <https://orcid.org/0000-0003-4151-115X>, e-mail: [svtarasov@hse.ru](mailto:svtarasov@hse.ru)

Получена 15.06.2023

Принята в печать 28.07.2023

Received 15.06.2023

Accepted 28.07.2023