

# **Сравнительный анализ классических тестов творческого мышления и цифровой среды Creo Datum по оценке творческого мышления школьников**

## **A comparative analysis of classical creative thinking tests and the Creo Datum digital environment for assessing the creative thinking of schoolchildren**

УДК 37.013.75, 37.015.3

DOI: 10.12737/2500-3305-2025-10-4-145-151

### **Коваленко И.В.**

Аспирант, ФГБОУ ВО «Московский педагогический государственный университет», учитель химии ГБОУ города Москвы «Школа имени В.В. Маяковского», г. Москва  
e-mail: iljakov@yandex.ru

### **Kovalenko I.V.**

Postgraduate student, Moscow State Pedagogical University, chemistry teacher, V.V. Mayakovsky School, Moscow  
e-mail: iljakov@yandex.ru

### **Оржековский П.А.**

Д-р пед. наук, член-корреспондент РАО, профессор кафедры естественнонаучного образования и коммуникативных технологий ФГБОУ ВО «Московский педагогический государственный университет», г. Москва  
e-mail: p.a.orzhekovskiy@gmail.com

### **Orzhekovsky P.A.**

Doctor of Pedagogical Sciences, Corresponding Member of the Russian Academy of Education, Professor of the Department of Natural Science Education and Communication Technologies, Moscow State Pedagogical University, Moscow  
e-mail: p.a.orzhekovskiy@gmail.com

### **Аннотация**

В статье рассматривается проблема эффективности и достоверности результатов оценки творческого мышления школьников 8-9-х классов путем решения классических тестов творческого мышления Вильямса и Торренса и выполнения заданий в цифровой среде Creo Datum.

**Ключевые слова:** творческое мышление, Creo Datum, творческие задания, цифровизация в образовании, тест Торренса, тест Вильямса.

### **Abstract**

The article examines the problem of the effectiveness and reliability of the results of assessing the creative thinking of schoolchildren in grades 8-9 by solving the classic Williams and Torrence creative thinking tests and completing tasks in the Creo Datum digital environment.

**Keywords:** creative thinking, Creo Datum, creative tasks, digitalization in education, Torrence test, William's test.

## **Введение**

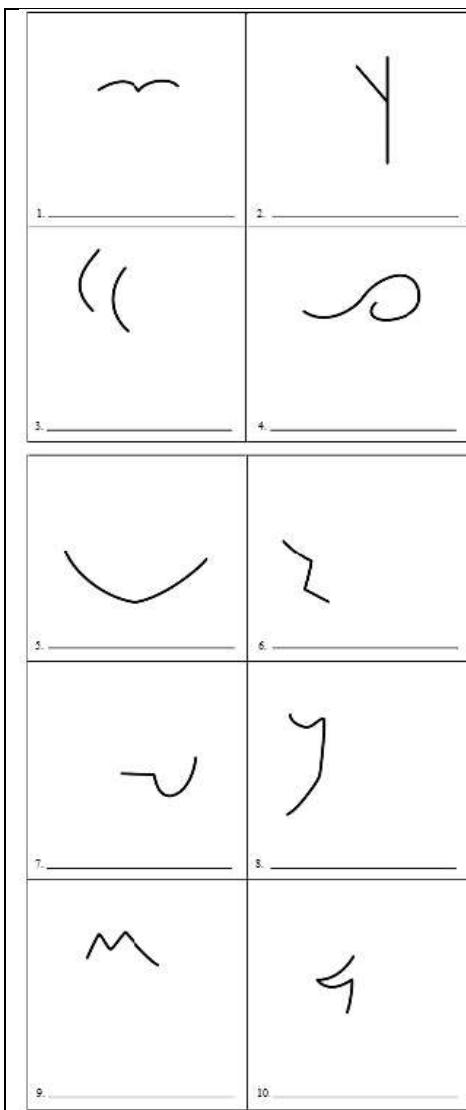
В современном мире цифровизация проникла практически во все сферы деятельности, в том числе и в образование. Новые системы и сервисы позволяют учителю современной школы гораздо быстрее и эффективнее выдавать и проверять задания, работать с библиотеками данных, систематизировать рутинные процессы. Именно под автоматизацию повседневной деятельности и заточены современные цифровые системы, однако для творческой деятельности таких разработок до сих пор нет. Связано это с тем, что не существует четких границ и рамок для работы такой системы в творческой среде, слишком много параметров приходится одновременно учитывать.

Дополнительную трудность вносит и то, что классические тесты, позволяющие оценивать творческое мышление, основаны на оценке завершенности предложенного рисунка [1, 2], который испытуемый должен закончить по своему усмотрению. Автоматизация такого процесса требует анализа картинок по набору ответов этой же группы учащихся, что даже с учетом современного уровня развития систем искусственного интеллекта представляет значительную трудность и существенные временные затраты.

Данная статья предлагает рассмотреть новую цифровую систему Creo Datum [3] в качестве перспективного метода оценки творческих способностей учащихся, не уступающую классическим творческим тестам.

### **Преимущества и недостатки классических систем оценивания творческого мышления.**

Рассмотрим более подробно систему оценивания креативного теста Торренса на примере его основной части [2]. Учащемуся предлагается завершить фрагменты 10 рисунков (рис. 1) по своему усмотрению, при этом обязательным условием является включение нарисованной части в структуру рисунка.



**Рис. 1.** Фрагменты рисунков, предлагаемые учащимся для дополнения в рамках тестового задания

		
Название: птица	Синие Красные Серые Темный цвет	Название: птица
Пример 1 Разработанность ответа: 0	Пример 2 Разработанность ответа: 4	Пример 3 Разработанность ответа: 12
Название: бабочка с двумя глазками		
Приведены три примера подсчета баллов за разработанность ответов. Следует их внимательно изучить.		
СПИСОК № 1(Напомним, что ответы, не указанные в списке № 1, получают оценку по оригинальности 2 балла как нестандартные и встречающиеся реже, чем в 2% случаев. Категория этих ответов определяется по списку № 2). Ответы на задание с указанием номеров категорий и оценок по оригинальности		
<b>Фигура 1</b> 0 баллов (5% и более ответов) (24) Абстрактный узор. (37) Лицо, голова человека. (1) Очки. (38) Птица (летящая), чайка. 1 балл (от 2 до 4,99%) (10) Брови, глаза человека. (33) Волна, море. (4) Животное (морда). (4) Кот, кошка. (21) Облако, туча; (58) Сверхъестественные существа. (10) Сердце («любовь»). (4) Собака. (8) Сова. (28) Цветок. (37) Человек, мужчина. (31) Яблоко.		
<b>Фигура 2</b> 0 баллов (5% и более ответов) (24) Абстрактный узор. (64) Дерево и его детали. (67) Рогатка. (28) Цветок. 1 балл (от 2% до 4,99%) (41) Буква: Ж, У и др. (13) Дом, строение. (42) Знак, символ, указатель. (8) Птица, слепыш, ноги. (45) Цифра. (37) Человек.		
<b>Фигура 3</b> 0 баллов (5% и более ответов) (24) Абстрактный узор. (53) Звуковые и радиоволны. (37) Лицо человека. (9) Парусный корабль, подка. (31) Фрукты, ягоды. 1 балл (от 2 до 4,99%) (21) Ветер, облака, дождь. (7) Воздушные шары. (64) Дерево и его детали. (49) Дорога, мост. (4) Животное или его морда. (48) Карусели, качели. (68) Колеса. (67) Лук и стрелы. (35) Луна. (27) Рыба, рыбы. (48) Санки. (28) Цветок.		
<b>Фигура 4</b> 0 баллов (5% и более ответов) (24) Абстрактный узор. (33) Волна, море. (41) Вопросительный знак. (4) Змея. (37) Лицо человека. (4) Хвост животного, хобот слона. 1 балл (от 2 до 4,99%) (4) Кот, кошка. (32) Кресло, стул. (36) Ложка, половник. (4) Мышь. (38) Насекомое, гусеница, червь. (1) Очки. (8) Птица: гусь, лебедь. (27) Ракушка. (58) Сверхъестественные существа. (1) Трубка для курения. (28) Цветок.		
<b>Фигура 5</b> 0 баллов (5% и более ответов) (24) Абстрактный узор. (36) Блюдо, ваза, чаша. (9) Корабль, подка. (37) Лицо человека. (65) Зонт. 1 балл (от 2 до 4,99%)		

**Рис. 2.** Фрагмент таблицы оценивания каждого рисунка с критериями оценивания и выставляемыми баллами

Проверка теста производится с помощью специальных листов оценивания (рис. 2), на которых указаны категории возможных ответов с присвоенными баллами по каждой из категорий. При этом изначально в систему оценивания заложено несколько значимых для проверки факторов:

- требуется сравнение предложенных ответов внутри работ группы, что, при большом числе участников, достаточно трудоемко;
- категории ответов, за которые начисляются максимальные баллы, вынесены в отдельный список, при этом требуется соотносить рисунок, если его нет в списке, с близкой по смыслу категорией;
- наиболее значимый аспект проверки: желательно, чтобы была возможность проверки экспертом [2], имеющим опыт оценивания таких работ на других группах, что еще более затруднительно в рамках отдельно взятой школы;
- оценивание в достаточной степени субъективно, то есть зависит от проверяющего и его понимания детализации рисунка;
- степень оригинальности установлена на очень низком уровне – 2%, т.е. от каждого испытуемого, фактически, зависит, сколько баллов можно получить за работу в целом.

Например, если из 20 чел. у 2 работ одинаковые оригинальные рисунки, отсутствующие в критериальной таблице, то оба участника получат 0 баллов за оригинальность, тогда как если одну из работ убрать, то балл одного из учеников сразу же повышается на 2.

Как видно из вышеприведенных факторов оценивания, креативные тесты в большой степени субъективны. С одной стороны, это позволяет проявить индивидуальный подход к каждому при оценивании, но, с другой стороны, возникают значительные расхождения в выставляемой оценке.

Аналогичные проблемы при оценивании возникают и при выполнении теста Вильямса, так как структура и оценивание схожи с тестом Торренса. Дополнительным значимым фактором для тестов подобной структуры является временной ресурс. Сравнение более 30 работ одновременно между собой и по таблице категорий требует значительных временных затрат, не говоря уже о поточной проверке таких тестов на регулярной основе.

#### **Преимущества и недостатки цифровой системы оценивания творческого мышления Creo Datum.**

Использование автоматизированной цифровой платформы для оценки творческого мышления позволило бы разрешить ряд проблем, возникающих при выполнении классических тестов. Во-первых, это позволило бы избежать субъективного оценивания на критериальной основе. Во-вторых, в разы бы сократило время проверки, что позволило бы одному учителю проводить оценку параллельно в нескольких группах. Авторами предложено альтернативное решение в оценке творческого мышления в цифровой системе Creo Datum.

Работа системы построена на анализе словарного набора тестируемого ученика [4] и дальнейшей алгоритмической обработке [5] этого ответа отдельно от тестируемой группы или внутри нее. Система распознает и считает только уникальные слова ответа на поставленную задачу, преобразуя полученные баллы в итоговый результат с соотнесением по конвергентности (т.е. числу предложенных решений одной задачи) и по дивергентности (числу предложенных недостатков этих решений).

Основное отличие от классических тестов состоит в том, что связно формулировать ответ на задание значительно сложнее, чем дорисовывать необходимые элементы и детали к образцу, в связи с чем для корректной оценки творческого мышления в Creo Datum требуется некоторое накопление данных, что в определенной степени является недостатком системы. Ученик должен привыкнуть к интерфейсу, модели и форме сбора данных, чтобы в дальнейшем не отвлекаться от выполнения задания.

Отправка, обработка и оценивание ответа ведется в почти полностью автоматическом режиме, учитель проверяет только содержание ответа на самом начальном этапе во избежание не соответствующих заданию формулировок (например, «не хочу решать эту задачу»).

#### **Численные характеристики методов оценки творческого мышления.**

Рассмотрим численные показатели каждого метода по отдельности и сравним между собой. В тестировании приняли участие 14, 11 и 19 учащихся, соответствующих следующим категориям:

- Группа 1 (12 чел.) – 8 класс, начинающие работать с цифровой средой, ученики профильного естественно-научного класса;
- Группа 2 (11 чел.) – 9 класс, имеющие опыт ученики профильного естественно-научного класса, проходящие тестирования с 8 класса;
- Группа 3 (20 чел.) – 9 класс, ученики социально-экономического класса, начинающие работать с цифровой средой (контрольная группа).

Каждой группе учеников предлагалось в разные промежутки времени выполнить тестирование каждым способом, при этом о взаимосвязи тестов друг с другом не сообщалось. Группы 1 и 3 проходили тестирование в Creo Datum два раза, группа 2 выполняла задания более 10 раз.

Необходимо отметить, что при выполнении заданий в цифровой среде учащиеся одного класса разбиваются системой на разноуровневые группы, которые затем решают задачи разной сложности, поэтому подсчет баллов немного отличается, что будет описано ниже.

Сравнение проводилось по следующей процедуре:

– Баллы за тест Торренса и Вильямса суммировались, всего каждый испытуемый мог набрать 311 баллов, затем рассчитывался процент, который учащийся набрал по итогу двух тестов (см. столбец 1, табл. 1).

– Баллы за решение задач в системе Creo Datum суммировались в зависимости от группы, в которой учащийся проходил тестирование, поэтому для каждого ученика он мог отличаться, но для него являлся максимально возможным. Затем рассчитывался процент, который был набран по итогу решения всех задач (см. столбец 2, табл. 1).

– Рассчитывалась разность значений двух относительных показателей (см. столбец 3, табл. 1).

– За доверительное значение стандартного отклонения выбрано минимальное значение стандартного отклонения теста Вильямса, составляющее 28% [2] (у теста Торренса оно составляет порядка 40-47%), которое и являлось маркером, определяющим, попадает ли получаемое значение в заданный диапазон или нет.

– Рассчитывалось частное от деления разности значений долей тестов и долей системы: если доля оказывалась больше 28%, то значение помечалось красным (не совпадает), если меньше или равно – зеленым (совпадает).

Результаты представлены в табл. 1 и 2. Приведен общий вид таблицы расчетов и полученных значений, а также полученные результаты по трем группам и сводный по 43 ученикам (табл. 2)

Таблица 1

**Расчетные значения выполнения креативных тестов, тестов в системе Creo Datum и параметры корреляции значений между собой для учеников Группы 1**

Ученик группы 1	Доля за тесты	Доля в Creo Datum	Δ	Система
1.1	0,73	0,53	0,20	0,38
1.2	0,78	0,65	0,13	0,21
1.3	0,72	0,46	0,26	0,58
1.4	0,70	0,92	0,22	0,24
1.5	0,68	0,63	0,05	0,08
1.6	0,60	0,81	0,21	0,26
1.7	0,81	0,62	0,19	0,31
1.8	0,70	0,67	0,03	0,05
1.9	0,72	0,72	0,00	0,01
1.10	0,80	0,64	0,16	0,25
1.11	0,72	0,64	0,08	0,12
1.12	0,58	0,17	0,41	2,45

Таблица 2

**Процентные соотношения числа учеников, попадающих и не попадающих в доверительный интервал креативных тестов после прохождения тестирования в цифровой системе**

Категория	Группа 3	Группа 1	Группа 2	Объединенная группа
Не попадают в доверительный интервал	45%	33%	18%	33%
Попадают в доверительный интервал	55%	67%	82%	67%
Стандартное отклонение	45%	29,9%	21%	34,7%

На основании полученных результатов можно отметить несколько важных деталей. Контрольная группа 3 демонстрирует, что ученики с практически равной вероятностью попадают и не попадают в заявленный доверительный интервал с большим стандартным отклонением их показателей от среднего значения.

Это говорит о том, что с поверхностными знаниями о предмете, а тестирование происходит на задачах, связанных с химией (но не требующих каких-то глубоких знаний), учащиеся непрофильных классов демонстрируют отсутствие корреляции с креативными тестами.

Группа 1, напротив, имея представление о предмете, но с малым числом попыток решения задач в цифровой системе, демонстрирует уже значительные изменения в показателях, причем стандартное отклонение практически совпадает с рассчитанным для теста Вильямса, т.е. можно говорить о взаимной корреляции результатов тестирований – результаты в пределах погрешности совпадают между собой.

Группа 2, состоящая из наиболее опытных пользователей цифровой системы демонстрирует показатели значительно меньше стандартного отклонения теста Вильямса, что говорит о высокой степени совпадения результатов креативных тестов и системы Creo Datum между собой, т.е. о высокой и значимой корреляции результатов. Можно говорить о том, что результаты могут быть взаимозаменяемыми и оценка творческого мышления путем прохождения теста в цифровой системе эквивалентна результату креативного теста.

### **Выводы**

Продемонстрированные результаты тестирования учеников 8-9-х классов позволяют говорить о возможности замены классических креативных тестов на цифровую систему оценки творческого мышления Creo Datum, что значительно облегчит работу по формированию заданий, их проверке и объективному оцениванию, а также значительно сократит общее время выполнения и проверки. Ограничениями системы на текущий момент является наличие широкого набора заданий только по химической тематике, однако масштабирование системы на другие предметы и компетенции не составляет большого труда.

Небольшое время выполнения тестов в электронной системе также нивелирует проблему накопляемости данных для достоверной оценки показателя творческого мышления, тесты можно проводить от нескольких раз в месяц до ежедневного использования.

Рассчитанные значения стандартных отклонений по полученным данным позволяют говорить о корректности и достоверности получаемых значений, а следовательно, и о взаимозаменяемости двух методов оценки творческого развития школьников.

### **Литература**

1. Коваленко И.В., Оржековский П.А. Влияние используемого словарного набора при решении творческих задач по химии на оценку креативного мышления школьника. // Учебный эксперимент в образовании. 2024. № 2 (110). С. 77-86.
2. Оржековский П.А., Степанов С. Ю., Мишина И.Б. О непрерывности оценки развития у обучающихся репродуктивных и креативных мыслительных действий // Непрерывное образование: XXI век. 2019. № 3 (27). С. 28-39.
3. Туник Е.Е. Диагностика креативности. Тест Торренса. Методическое руководство. // Санкт-Петербург. ГП «ИМАТОН». 1998. 170 с. ISBN 5-7822-0021-9.
4. Туник Е.Е. Модифицированные креативные тесты Вильямса. // Санкт-Петербург. Речь. 2003. 96 с. ISBN 5-9268-0164-8.