

Н.Б. Печатнова  
Контрольно-измерительные материалы ЕГЭ  
как механизм внешнего оценивания.

Актуальность исследования.

С начала 90-х годов XX века в российском образовании идет процесс обновления традиционной контрольно-оценочной системы, обусловленный необходимостью поиска новых подходов к управлению качеством образования, мониторинга процессов модернизации российского образования. Одним из путей внедрения научных методов оценки учебных достижений, в частности, является введение единого государственного экзамена (ЕГЭ). В качестве целей, которых предполагается достичь в процессе эксперимента по введению единого государственного экзамена, обычно выделяют следующие:

- расширение доступности высшего образования,
- повышение качества среднего школьного образования, путем создания механизма объективного внешнего оценивания.

Созданный в настоящее время механизм внешнего оценивания выпускников направлен на оценку содержательной стороны интеллекта:

- в контрольно-измерительных материалах ЕГЭ по гуманитарным дисциплинам, в основном, присутствуют вопросы фактологического («кто – что – где – когда») и терминологического («как называется») типов,
- контрольно-измерительные материалы ЕГЭ по математическим дисциплинам в основном состоят из заданий направленных на воспроизведение ранее изученного алгоритма (вычислите, найдите, решите).

Однако для объективного внешнего оценивания выпускников в контрольно-измерительных материалах ЕГЭ явно не хватает заданий, направленных на выявление не только знаний и умений, но и способностей учащихся, обеспечивающих эффективность отбора будущих студентов на основе контрольно-измерительных материалов ЕГЭ. Таким образом, в рамках исследования, ставилась цель сравнительного исследования: определить корреляции успешности сданного по форме и материалам ЕГЭ экзамена и/или психологического тестирования умственного развития с успешностью дальнейшего обучения в вузе и выработать на основе полученных материалов рекомендаций по совершенствованию контрольно-измерительных материалов ЕГЭ.

С этой целью было проведено психологического тестирования умственного развития студентов 1 курса математики и информатики и педагогического факультета и на основании данных, представленных Региональным центром обработки единого государственного экзамена по результатам ЕГЭ и деканатами факультетов по итогам зимней сессии проведен сравнительный анализ полученных результатов.

Гипотеза: можно предполагать, что контрольно-измерительные материалы ЕГЭ будут содействовать улучшению качества объективного внешнего оценивания, если будут оценивать не только знания и умения выпускников, но и уровень развития способностей.

Проверка достоверности сформулированной гипотезы требует решения следующих задач:

1. Выявить, насколько выявленный в результате тестирования уровень способностей учащихся соответствует результатам первой зимней сессии студентов 1 курса факультета математики и информатики и педагогического факультета.
2. Выявить, насколько результаты ЕГЭ по математике и русскому языку соответствуют результатам дальнейшего обучения студентов 1 курса факультета математики и информатики и педагогического факультета.

На основе проведенного анализа:

1. Выявить необходимость включения в контрольно-измерительные материалы ЕГЭ заданий, направленных на диагностику способностей выпускников.
2. Разработать методические рекомендации по внедрению в контрольно-измерительные материалы ЕГЭ заданий, направленных на диагностику способностей выпускников, обеспечивающих прогнозирование успешности их дальнейшего обучения в вузе.

#### **Проведение экспериментального исследования.**

Для участия в эксперименте были выбраны студенты 1 курса факультета математики и информатики БГПУ и студенты 1 курса педагогического факультета БГПУ, сдававшие в 2003 году Единый государственный экзамен по математике и русскому языку. Из 89 студентов, обучающихся по специальности «Информатика и математика» в психологическом тестировании умственного развития участвовало 54 человека, что составляет чуть больше 60 % от общего количества. На педагогическом факультете в психологическом тестировании участвовало 35 человек из групп, обучающихся по специальности «Педагогика и психология» и по специальности «Дошкольная педагогика и психология» с дополнительной специальностью «Иностранный язык», количество студентов участвовавших в психологическом тестировании составляет 58,3% от общего количества.

Для диагностики умственного развития студентов были выбраны психологические тесты ШТУР, Амтхауэра и Равена.

Первым в экспериментальном исследовании был использован тест ШТУР (школьный тест умственного развития), разработанный авторским коллективом К.М. Гуревич, М.К. Акимовой, Е.М. Борисовой, В.Г. Зархиным, В.Т. Козловой, Г.П. Логиновой. В задания ШТУР включены понятия, подлежащие обязательному усвоению в учебных предметах трех циклов: математического, гуманитарного и естественно-научного. Помимо этого, определялась осведомленность в некоторых понятиях общественно-политического и научно-культурного содержания.

ШТУР состоит из 6 субтестов, каждый из которых включает от 15 до 25 однородных заданий возрастающей сложности. Два первых субтеста направлены на выявление общей осведомленности студентов, в этих субтестах исследуется индуктивное мышление, третий субтест - на проверку умения устанавливать аналогии, исследует комбинаторные способности, четвертый – на классификацию, исследует способность к абстрагированию, пятый - на обобщение, исследует способность выносить суждения, шестой – на установление закономерностей в числовых рядах, исследует способность оперировать числами.

Выделим несколько принципиальных отличий теста ШТУР от традиционных тестов:

- 1) Особое содержание заданий, используются не житейские, а школьные понятия, обязательные для усвоения.
- 2) Способы репрезентации и обработки диагностических результатов, ориентированы не на статистическую норму, а на социально-психологический норматив.
- 3) Коррекционные возможности методики, предусматривают на ее основе специальные способы исправления замеченных дефектов развития.

ШТУР соответствует высоким статистическим критериям, которым должен соответствовать любой диагностический тест. Он апробирован на больших выборках и доказал свою эффективность при решении таких задач, как:

- 1) анализ успешности развития по мере перехода из класса в класс;
- 2) сравнение групп учащихся, отобранных в соответствии с какими-либо целями;
- 3) сравнение эффективности различных программ и методов обучения;
- 4) отбор в спецклассы и спецшколы;
- 5) диагноз причин школьной неуспеваемости;
- 6) отбор учащихся, нуждающихся в коррекции умственного развития;
- 7) оценка эффективности работы разных педагогов и педагогических коллективов;

- 8) выявление учащихся, нуждающихся в индивидуальных программах обучения;
- 9) профориентационная работа с учащимися.

Тест ШТУР удовлетворяет социально-психологическим нормативам, которые предъявляет общество к каждому своему члену и может выступать критерием оценки индивидуального и возрастного развития.

Вторым в экспериментальном исследовании был использован американский тест Амтхауэра, обладающий высокой прогностической валидностью до 72%.

Тест Амтхауэра предназначен для оценки уровня интеллектуального развития лиц в возрасте от 13 до 61 года. Тест Амтхауэра более объемный по сравнению с тестом ШТУР, он состоит из 9 субтестов. Помимо, субтестов содержащихся в тесте ШТУР, в тесте Амтхауэра содержится субтест, исследующий уровень развития практического мышления, субтест, проверяющий способность сосредоточить внимание и сохранить в памяти усвоенное, и 2 субтеста, исследующие пространственное воображение студентов.

Третий тест, использованный в экспериментальном исследовании – тест Равена. В тесте Равена содержится набор заданий, с помощью которых выявляются интеллектуальные способности детей с 5 лет и взрослых до 65 лет. В процессе решения составляющих теста проявляются три психологических процесса: внимание, восприятие, мышление. С помощью теста Равена исследователи судят не столько об интеллекте в целом, сколько о способностях к систематизированной интеллектуальной деятельности, о логичности мышления.

При решении задач по таблицам Равена большое значение имеет концентрация активного внимания, его достаточный уровень распределения. Отсутствие вербальных заданий в тесте Равена уменьшает влияние образования и жизненного опыта.

Шкала состоит из 60 задач, разделённых на 5 серий по 12 задач в каждой. Во всех наборах решение первой задачи самоочевидно. Задачи, которые следуют в каждом ряду, постепенно усложняются. Также усложняются задачи и от серии к серии.

Выполнение заданий требует от испытуемого тщательного анализа структуры основного изображения и обнаружения этих же особенностей в одном из нескольких фрагментов.

### **Обработка результатов исследования для гипотезы <sup>1</sup>.**

---

<sup>1</sup> Безусловно, результаты на основе столь простых статистических методов не являются достаточно достоверными.

Более строгий и полный статистический анализ полученных результатов (на основе критерия хи - квадрат, распределения Стьюдента и пр.) будет приведен в следующей работе.

Величина коэффициента корреляции по Пирсону позволяет выяснить, в какой связи находятся результаты одного тестирования с результатами других тестов.

В нашем случае, используя формулу вычисления коэффициента корреляции по Пирсону, мы будем сравнивать:

I. Результаты ЕГЭ по математике и русскому языку с результатами с результатами первой зимней сессии студентов 1 курса факультета математики и информатики и педагогического факультета;

II. Результаты первой зимней сессии с результатами каждого субтеста теста ШТУР студентов 1 курса факультета математики и информатики и педагогического факультета.

III. Результаты первой зимней сессии с результатами каждого субтеста теста Амтхауэра студентов 1 курса факультета математики и информатики и педагогического факультета.

IV. Результаты субтестов, проверяющих одну и ту же способность в тестах ШТУР и Амтхауэра студентов 1 курса факультета математики и информатики и педагогического факультета.

V. Результаты первой зимней сессии с результатами теста Равена студентов 1 курса факультета математики и информатики и педагогического факультета.

Формула Пирсона имеет много различных видов, представляющих собой ее преобразование. Для проводимого исследования мы выбрали следующую формулу:

$$\text{Формула } r_{xy} = \frac{\sum_{i=1}^N X_i Y_i - \frac{\sum_{i=1}^N X_i \sum_{i=1}^N Y_i}{N}}{\sqrt{(\sum_{i=1}^N X_i^2 - (\sum_{i=1}^N X_i)^2 / N)(\sum_{i=1}^N Y_i^2 - (\sum_{i=1}^N Y_i)^2 / N)}} .$$

В результате вычисления корреляционной зависимости между результатами первой зимней сессии (алгебра, геометрия, математический анализ) и результатами ШТУР, Амтхауэра и ЕГЭ по математике на факультете математики и информатики были получены следующие результаты:

Таблица 1.

## Сводная таблица по тесту ШТУР ФМиИ

	Результаты первой зимней сессии		
	Алгебра	Геометрия	Математический анализ
<b>Субтест 1</b> (на выявление общей осведомленности)	0,33	0,26	0,27
<b>Субтест 2</b> (на выявление общей осведомленности)	0,23	0,36	0,22
<b>Субтест 3</b> (на установление аналогий)	0,36	0,35	0,21
<b>Субтест 4</b> (на классификации)	<b>0,47</b>	<b>0,34</b>	<b>0,33</b>
<b>Субтест 5</b> (на обобщение)	0,21	0,21	0,3
<b>Субтест 6</b> (на установление закономерностей в числовых рядах)	0,19	0,29	0,27
<b>Результаты ЕГЭ по математике</b>	<b>0,3</b>	<b>0,35</b>	<b>0,32</b>

Таблица 2.

## Сводная таблица по тесту Амтхауэра ФМиИ

	Алгебра	Геометрия	Матем. анализ
<b>Субтест 1 «Логический отбор»</b> Исследует индуктивное мышление, чутье языка, задача – закончить предложение одним из приведенных в качестве ответа слов.	<b>0,48</b>	0,16	-0,08
<b>Субтест 2 «Определение черт»</b> Исследует способность к абстрагированию, оперированию вербальными понятиями, Задача – в заданиях предлагаются пять слов, четыре из которых имеют общую смысловую связь, а одно слово связи не имеет, в качестве ответа необходимо указать это слово. (Субтест 4 ШТУР)	0,145	-0,085	-0,16
<b>Субтест 3 «Аналогия».</b> Исследует комбинаторные способности. Задача – в заданиях необходимо установить связь, аналогичную той, которая существует между парой предложенных слов.	0,19	0,07	0,08
<b>Субтест 4 «Классификация»</b> Исследует способность выносить суждения. Задача – обозначить два слова общим понятием.	-0,002	-0,03	-0,05

<b>Субтест 5 «Задания на счет»</b> Исследует уровень развития практического мышления. Задача – решить математическую задачу.	-0,008	-0,097	-0,2
<b>Субтест 6 «Ряды чисел»</b> Исследует индуктивное мышление, способность оперировать с числами. Задача – установить закономерность числового ряда и продолжить его.	0,03	-0,012	-0,22
<b>Субтест 7 «Выбор фигуры»</b> Исследует пространственное воображение, комбинаторные способности. Задача – указать фигуру, которая соответствует фигуре – заданию.	0,01	0,089	-0,1
<b>Субтест 8 «Задания с кубиками»</b> Исследует пространственное воображение, при оперировании пространственными фигурами. Задание идентифицировать кубики с кубиками-ответами.	-0,03	-0,1	-0,25
<b>Субтест 9 «Задание на способность сосредоточить внимание и сохранить в памяти усвоенное»</b> Задача запомнить слова, соединенные в группы и вспомнить, в какую из групп они входят.	0,2	<b>0,39</b>	-0,07
<b>Результаты ЕГЭ по математике</b>	<b>0,24</b>	<b>0,25</b>	<b>0,27</b>

**Из таблицы 1 видно, что на факультете математики и информатики существует высокая корреляционная зависимость между результатами субтеста 4 ШТУР, направленного на выявление умения устанавливать логические классификации (содержащего задания «на исключение лишнего», в задании предложены пять слов, четыре из которых имеют общую смысловую связь, а одно слово связи не имеет, в качестве ответа необходимо указать это слово), и результатами первой зимней сессии, особенно по алгебре.**

Ниже приведены результаты вычисления корреляционной зависимости между результатами первой зимней сессии (русский язык, введение в псих.-пед. деятельность, психологизм русской культуры) и результатами ШТУР, Амтхауэра и ЕГЭ по русскому языку на педагогическом факультете:

Сводная таблица по тесту ШТУР ПФ

	Результаты первой зимней сессии		
	Русский язык	Введение в псих.-пед. деятельность	Психологизм русской литературы
<b>Субтест 1</b> (на выявление общей осведомленности)	<b>0,33</b>	<b>0,28</b>	<b>0,56</b>
<b>Субтест 2</b> (на выявление общей осведомленности)	0,18	0,1	0,02
<b>Субтест 3</b> (на установление аналогий)	0,22	-0,05	0,11
<b>Субтест 4</b> (на классификации)	0,2	-0,52	0,17
<b>Субтест 5</b> (на обобщение)	<b>0,38</b>	<b>0,21</b>	<b>0,24</b>
<b>Субтест 6</b> (на установление закономерностей в числовых рядах)	0,21	0,43	-0,21
<b>Результаты ЕГЭ по русскому языку</b>	<b>0,28</b>	<b>0,42</b>	<b>0,47</b>

Таблица 4.

## Сводная таблица по тесту Амтхауэра ПФ

	Русский язык	Введение в псих.-пед. деятельность	Психологизм русской литературы
<b>Субтест 1</b> «Логический отбор» Исследует индуктивное мышление, чутье языка, задача – закончить предложение одним из приведенных в качестве ответа слов.	<b>0,49</b>	<b>0,44</b>	0,16
<b>Субтест 2</b> «Определение черт» Исследует способность к абстрагированию, оперированию вербальными понятиями, задача – в заданиях предлагаются пять слов, четыре из которых имеют общую смысловую связь, а одно слово связи не имеет, в качестве ответа необходимо указать это слово.	0,25	0,28	-0,22
<b>Субтест 3</b> «Аналогия». Исследует комбинаторные способности. задача – в заданиях необходимо установить связь, аналогичную той, которая существует между парой предложенных слов.	<b>0,31</b>	<b>0,54</b>	<b>0,52</b>
<b>Субтест 4</b> «Классификация» Исследует способность выносить суждения. задача – обозначить два слова общим понятием.	<b>0,52</b>	<b>0,74</b>	<b>0,18</b>
<b>Субтест 5</b> «Задания на счет» Исследует уровень развития практического мышления. задача – решить математическую задачу.	0,2	0,27	-0,32
<b>Субтест 6</b> «Ряды чисел» Исследует индуктивное мышление, способность оперировать с числами. задача – установить закономерность числового ряда и продолжить его.	-0,002	0,21	-0,3
<b>Субтест 7</b> «Выбор фигуры» Исследует пространственное воображение, комбинаторные способности. задача – указать фигуру, которая соответствует фигуре – заданию.	0,44	0,2	0,1
<b>Субтест 8</b> «Задания с кубиками» Исследует пространственное воображение, при оперировании пространственными фигурами. задание идентифицировать кубики с кубиками-ответами.	0,01	-0,15	-0,05
<b>Субтест 9</b> «Задание на способность сосредоточить внимание и сохранить в памяти усвоенное» Задача запомнить слова, соединенные в группы и вспомнить, в какую из групп они входят.	0,015	0,12	0,34
<b>Результаты ЕГЭ по русскому языку</b>	<b>0,49</b>	<b>0,59</b>	<b>0,51</b>

**Из таблиц 3 и 4 видно, что на педагогическом факультете существует высокая корреляционная зависимость между результатами субтестов, содержащих задания, в которых нужно обозначить два слова общим понятием (субтест 5 теста ШТУР и субтест 4 теста Амтхауэра) и результатами первой зимней сессии.**

Далее рассмотрим, какова корреляционная зависимость между субтестами, проверяющими развитие одной и той же способности в тестах ШТУР и Амтхауэра студентов 1 курса факультета математики и информатики и педагогического факультета.

Таблица 5.

**Сравнительная таблица  
связи результатов каждого субтеста теста ШТУР и теста Амтхауэра**

	<b>Субтест</b>	<b>ФМиИ</b>	<b>ПФ</b>
1.	Общая осведомленность, нужно закончить предложение (выбрать правильный вариант)	0,076	0,028
2.	Аналогии	0,137	-0,199
3.	На классификацию (ШТУР), на определение черт (Амтхауэр) задания на исключение лишнего	<b>0,333</b>	0,081
4.	На обобщение (ШТУР), на классификации (Амтхауэр) задание – обозначьте два слова общим понятием.	0,149	<b>0,89</b>
5.	На установление правильной последовательности (ШТУР), ряды чисел (Амтхауэр) задание – установите закономерность числового ряда и продолжите его.	-0,035	0,566

**Из таблицы 5 видно, что высокая корреляционная зависимость существует:**

- 1) на факультете математики и информатики между субтестами (субтест 4 ШТУР и субтест 2 теста Амтхауэра), содержащими задания, в которых предлагается пять слов, четыре из которых имеют общую смысловую связь, а одно слово связи не имеет, в качестве ответа необходимо указать это слово;**
- 2) на педагогическом факультете между субтестами (субтест 5 ШТУР и субтест 4 теста Амтхауэра), содержащими задания, в которых нужно**

**обозначить два слова общим понятием.**

В результате вычисления корреляционной зависимости между результатами ЕГЭ и результатами теста Равена, между результатами первой зимней сессии и результатами теста Равена студентов 1 курса факультета математики и информатики и педагогического факультета были получены следующие результаты, отраженные в таблицах 6 и 7:

Таблица 6.

**Сводная таблица по тесту Равена ФМИ**

	Результаты ЕГЭ		Результаты первой зимней сессии		
	Математика	Русский язык	Алгебра	Геометрия	Мат. анализ
Результаты теста Равена	0,19	0,08	0,26	0,23	-0,26

Таблица 7.

**Сводная таблица по тесту Равена ПФ**

	Результаты ЕГЭ		Результаты первой зимней сессии		
	Математика	Русский язык	Русский язык	Введение в псих-пед. деятельность	Психологизм русской литературы
Результаты теста Равена	0,6	0,32	0,25	0,07	0,01

Из таблиц 6 и 7 видно, что корреляционная зависимость между результатами ЕГЭ и результатами теста Равена на факультете математики и информатики ниже, чем между результатами первой зимней сессии и результатами теста Равена; на педагогическом факультете наоборот, корреляционная зависимость между результатами ЕГЭ и результатами теста Равена выше, чем между результатами первой зимней сессии и результатами теста Равена.

В соответствии с поставленными задачами проведенное экспериментальное исследование по выдвинутой гипотезе способствовало получению следующих выводов:

1. Для студентов педагогического факультета способность, обозначить два слова общим понятием, является прогностической способностью, т.к. результаты субтеста, проверяющего развитие этой способности, находятся в высокой корреляционной зависимости с результатами первой зимней сессии;
2. Для студентов факультета математики и информатики способность, проверяющая умение находить лишнее слово, из пяти предложенных, которое не связано общей смысловой связью, является прогностической способностью,

т.к. результаты субтеста, проверяющего развитие этой способности, находятся в высокой корреляционной зависимости с результатами первой зимней сессии;

3. Корреляционная зависимость между результатами ЕГЭ по русскому языку и математике является средней прямой на факультете математики и информатики и сильной прямой на педагогическом факультете.

Таким образом, для дальнейшего анализа на основе более точных статистических методов выдвигается следующая гипотеза:

Можно предполагать, что включение в контрольно-измерительные материалы ЕГЭ заданий, оценивающих уровень развития прогностической способности:

- 1) по математике заданий «на исключение лишнего» (в задании предложены пять слов, четыре из которых имеют общую смысловую связь, а одно слово связи не имеет, в качестве ответа необходимо указать это слово);
- 2) по русскому языку, заданий, в которых нужно обозначить два слова общим понятием

повысит качество объективного внешнего оценивания и обеспечит успешность дальнейшего обучения выпускников школ в вузах.

Литература:

1. Психологическая коррекция умственного развития учащихся.-М., 1990.
2. Тест структуры интеллекта Амтхауэра Приложение к книге "Практикум по возрастной психологии" Сост. Ж. А. Балакшина и др.; Под ред. Л. А. Головей.-СПб. 2002.
3. А.Я. Психология >> Описания тестов: <http://azps.ru/tests/>