

# Апробация опросника Г.А. Цукерман «Всегда — иногда — никогда» для диагностики метапредметных навыков учащихся пятых классов

**М.Г. Сорокова**

доктор педагогических наук, кандидат физико-математических наук, профессор кафедры математики и естественнонаучных дисциплин факультета информационных технологий  
Московского городского психолого-педагогического университета

Переход на новые Федеральные государственные образовательные стандарты начального и основного общего образования придает особую актуальность проблеме разработки адекватного инструментария для психолого-педагогических измерений. В 2011/12 учебном году ПИ РАО совместно с МГППУ было проведено масштабное научное исследование на тему «Комплексный подход в диагностике метапредметных результатов начального школьного образования в соответствии с новыми образовательными стандартами» в ЦАО Москвы (руководитель — ведущий научный сотрудник И.М. Улановская). Одной из задач исследования стала апробация ряда методик оценки различных аспектов компетентности учащихся, разработанных психологами ПИ РАО на материале школьных предметов. Настоящая статья посвящена статистическому анализу данных апробации опросника Г.А.Цукерман «Всегда — иногда — никогда» на выборках пяти классников ЦАО Москвы (1398 учащихся из 40 школ, из них 702 мальчика и 696 девочек) и гимназии № 1567 ЗАО Москвы (74 учащихся, из них 46 мальчиков, 27 девочек, пол одного испытуемого не указан). Статистически доказана надежность шкал опросника и влияние факторов «пол» и «категория учащихся» на показатели по итоговым шкалам и отдельным пунктам. Определены репрезентативные нормы для итоговых шкал. Установлено, что математика лучше, чем русский язык, развивает умения классифицировать, доказывать и общую компетентность учащихся. Тем самым экспериментально подтверждена возможность использования данного опросника для оценки заявленных метапредметных навыков в генеральной совокупности пятиклассников.

**Ключевые слова:** надежность, валидность, стандартизационная выборка, критерий Хи-квадрат, критерий Колмогорова—Смирнова, критерий Манна—Уитни, критерий Уилкоксона, двухфакторный дисперсионный анализ.

В настоящее время в отечественном образовании на всех его уровнях происходят интенсивные процессы обновления. В 2011/12 учебном году начальная школа перешла на новые Федеральные государственные образовательные стандарты, призванные предоставить школам более широкие возможности для самостоятельного формирования учебных программ, значительно модернизировать учебный процесс и демократизировать характер взаимоотношений педагогов и учащихся, школы и родителей. Новые образовательные стандарты основного общего образования для V—IX классов стартовали в режиме апробации с 1 сентября 2012 г. в ряде школ страны. Поможет ли переход на новые стандарты обеспечить высокое качество образования? Каков реальный уровень знаний, умений и навыков учащихся? Какие требования к ним можно предъявить на разных возрастных ступенях? В связи с необходимостью найти ответы на эти и другие подобные вопросы особенно актуальной стала проблема разработки адекватного инструментария для психолого-педагогических измерений.

Первые стандартизированные тесты для оценки результатов школьного обучения появились на рубеже

XX в. Так, Г. Эббингауз предъявлял учащимся тесты арифметического счета, объема памяти и завершения предложений, а в 1920-е гг. появились батареи достижений, позволяющие сравнить выполнение заданий по разным школьным предметам относительно единой системы нормативов. Развитию тестирования в США и Европе способствовало также создание региональных и национальных программ. В нашей стране начало широкому использованию тестов достижений в школах положило введение ЕГЭ в середине 2000-х.

В 2011/12 учебном году ПИ РАО совместно с МГППУ было проведено научное исследование на тему «Комплексный подход в диагностике метапредметных результатов начального школьного образования в соответствии с новыми образовательными стандартами» в ЦАО Москвы (руководитель — И.М. Улановская). Целью исследования стала экспериментальная оценка различных аспектов компетентности учащихся 5-х классов, а одной из его важнейших задач — апробация ряда диагностических методик, разработанных психологами ПИ РАО на материале школьных предметов. Согласно междуна-

родным стандартам, хороший тест должен удовлетворять ряду требований, таких как внутренняя согласованность, рассматриваемая одними авторами как составляющая конструктивной валидности [1; 3], а другими — как составляющая надежности теста [5; 9; 10; 2], а также дифференциальная валидность, понимаемая здесь как способность теста выявить различия по социо-демографическим факторам [9]. Настоящая статья имеет целью всестороннюю проверку **гипотезы**: опросник Г.А. Цукерман «Всегда — иногда — никогда» обладает указанными характеристиками и, следовательно, пригоден для оценки метапредметных навыков учащихся 5-х классов.

### Метод

Исследование проводилось в центральном административном округе (ЦАО) г. Москвы. В нем в качестве стандартизированной выборки приняли участие  $n_1 = 1398$  учащихся пятых классов из 40 школ ЦАО, из них 702 мальчика и 696 девочек. В качестве контрольной группы выступили  $n_2 = 74$  пятиклассника гимназии № 1567 Западного административного округа (ЗАО) г. Москвы, из них 46 мальчиков, 27 девочек, пол одного испытуемого не указан. Эта школа является многопрофильным образовательным учреждением и осуществляет набор в пятые классы. Школьники V–VII классов обучаются по гимназической программе с культурологической доминантой, в старших классах осуществляется профильное обучение. Таким образом, предполагалось, что учащиеся этой школы покажут более высокие результаты по сравнению с учащимися ЦАО Москвы.

Все учащиеся были обследованы, в частности, с помощью опросника Г.А. Цукерман «Всегда — иногда — никогда», проверяющего их умение классифицировать и приводить доказательство на материале русского языка и математики. Методика включает по 10 утверждений, предполагающих знание русского языка и математики в объеме начальной школы. Каждое утверждение может быть верно «всегда», «иногда» или «никогда», и это предлагается определить испытуемому. Если он считает, что утверждение верно только «иногда», он должен привести доказательство — примеры, когда оно является верным, а когда — нет. Например, одно из утверждений по математике звучит так: «Если из трехзначного числа вычесть двузначное, то результат будет однозначным числом». Правильный ответ — «иногда», так как  $100 - 99 = 1$  (однозначное число), а  $100 - 50 = 50$  (двузначное число). Опросник содержит 12 итоговых шкал.

Математическая обработка полученных эмпирических данных проводилась в статистическом пакете SPSS 13-й и 19-й версий и включала исследование отдельных пунктов опросника (описательную статистику и проверку различий в распределениях школьников стандартизированной и контрольной групп); проверку надежности итоговых шкал опросника; исследование итоговых шкал (описательная статистика, проверка на нормальность и на предмет дифференциальной валидности, определение норм, а также сопоставление результатов по итоговому шкалам одинакового диапазона). При этом использовались следующие методы: критерий Колмогорова — Смирнова, критерий однородности Хи-квадрат, критерий Манна — Уитни, критерий знаковых ранговых сумм Уилкоксона и двухфакторный дисперсионный анализ ANOVA  $2 \times 2$ .

### Результаты

#### 1. Выявление различий между распределениями учащихся школ ЦАО и гимназии № 1567 по пунктам опросника «Всегда — иногда — никогда»

Для всех вопросов (пунктов) на классификацию вычислялись следующие показатели описательной статистики: число и процент детей, ответивших на вопрос задачи верно, неверно, а также число и процент отказов от ответа. Общее число детей, участвовавших в диагностике (100%), равно сумме частот этих трех групп детей. Проверка статистической значимости различий между распределениями ответов в обеих группах по каждому пункту методики производилась при помощи критерия однородности Хи-квадрат.

Выявлены достоверные различия в распределениях ответов на вопросы, проверяющие способность к классификации по русскому языку, по пяти пунктам из десяти. При этом в ЦАО больший процент ошибок и меньший процент правильных ответов, чем в школе № 1567. Процент отказов в школе № 1567, как правило, немного меньше, чем в ЦАО, но разница незначительна и составляет не более 4 %. Наиболее яркие различия, показывающие характерную тенденцию, отражены в табл. 1.

Здесь в ЦАО на 20,5 % больше ошибок и на 24,6 % меньше правильных ответов, чем в школе № 1567. Процент отказов в ЦАО несколько выше.

Еще более контрастная картина получается при проверке различий по вопросам, оценивающим способность к классификации по математике. Выявлены достоверные различия в распределениях ответов по девяти из десяти пунктов, кроме «Математика, 2а». При этом тенденция та же, что и для русского языка,

Таблица 1

Распределения ответов по пункту «Русский язык, 6а» в группах учащихся ЦАО и школы № 1567

Группа		Русский язык, 6а			Всего
		Неверно	Верно	Отказ	
ЦАО Москва	Количество	627	715	56	1398
	%	44,8	51,1	4,0	100,0
Школа № 1567	Количество	18	56	0	74
	%	24,3	75,7	0,0	100,0

но по некоторым пунктам увеличивается разница процента отказов с 4 % до 7,5–10 %. Так, например, по пункту «Математика, 9а» в ЦАО на 21,7 % больше ошибок и на 29,3 % меньше правильных ответов, чем в школе № 1567, а процент отказов больше на 7,5 %.

Если умение классифицировать проверяют по 10 вопросов методики Г.А. Цукерман в обеих дисциплинах, то способность к доказательству — только по 5. По русскому языку получены достоверные различия по четырем из пяти пунктов, а по математике — по всем пяти. При этом, как и в случае с классификацией, в ЦАО больший процент ошибок и меньший процент правильных ответов, чем в школе № 1567, а процент отказов порой больше на 8,6–19,2 %. В качестве примера приведем наиболее характерный результат (табл. 2). Здесь вместо порядкового номера данного вопроса мы использовали его код «1б». Обратим внимание, что выборки существенно уменьшились по сравнению с первоначальными, так как в рассмотрение были взяты только те школьники, которые правильно ответили «иногда» на соответствующий вопрос «а».

Таким образом, дифференцирующей способностью по фактору «категория учащихся» обладают 23 из 30 пунктов опросника.

## 2. Проверка надежности шкал опросника «Всегда-иногда-никогда»

В качестве меры надежности использовался коэффициент Альфа Кронбаха ( $\alpha$ ) [5; 9; 10; 2]. Чем ближе этот показатель к единице, тем лучше согласованы пункты соответствующей шкалы. Надежность шкал проверялась отдельно для каждой из рассматриваемых выборок. Для большинства шкал показатель  $\alpha$  Кронбаха для обеих выборок лежит в пределах от 0,5 до 0,7, что в целом следует признать удовлетворительным. Как правило, на выборке ЦАО, значительно большей по объему, он выше. Максимальных значений он достигает для шкал «Доказательство — Иногда» (0,723), «Классификация — Итог» (0,736), «Математика — Компетентность» (0,795) и «Общий итог: Русский язык и Математика» (0,829). Только для шкал «Русский язык — Доказательство» и «Верная классификация — Всегда» его значение падает примерно до 0,4, однако существенно повысить его за счет выбрасывания пунктов не получится: на выборке ЦАО таких пунктов нет. Здесь вообще практически нет пунктов, повышающих  $\alpha$  Кронбаха при их удалении. Эта выборка существенно больше, чем школы № 1567, поэтому результаты выглядят более убедительными.

Одним из важных вопросов разработчиков диагностических методик является вопрос о том, какие пункты плохо согласованы с соответствующей шкалой и, возможно, нуждаются в удалении из опросника. В данном случае прирост  $\alpha$  Кронбаха, который дают отдельные пункты методики Г.А. Цукерман при их удалении, очень мал (это сотые и тысячные доли). Мы считаем, что эти пункты методики следует оставить. Например, такие пункты, как Р2а и Р3а, очень незначительно — на 0,006 и 0,02 соответственно — повышают при удалении показатель  $\alpha$  Кронбаха для шкал «Верная классификация — Иногда» и «Верная классификация — Никогда» соответственно, но они хорошо работают на другие шкалы методики, как следует из данных ЦАО. Если же мы их удалим, то чуть-чуть улучшим согласованность указанных шкал, зато ухудшим другие шкалы. По школе № 1567 оказался плохо работающим пункт 8 по русскому языку. Надо сказать, что это один из двух наиболее сложных для ответов пунктов в обеих выборках.

Мы полагаем, что удаление отдельных пунктов данной методики нецелесообразно еще и по следующим причинам. Во-первых, удаление вопросов «а» пунктов с правильным ответом «Иногда» автоматически влечет за собой удаление соответствующих вопросов «б». От этого диапазон шкал уменьшится, что затруднит в дальнейшем их применение для выявления различий между группами. Во-вторых, в исходном варианте данной методики диапазон шкал типа «Классификация», «Доказательство» и итоговых шкал «Компетентность» попарно одинаков для русского языка и математики. Это позволяет ответить на вопрос, какой же из двух предметов лучше развивает метапредметные навыки (см. ниже). Если неравномерно выбросить пункты, это будет невозможно.

## 3. Исследование итоговых шкал опросника «Всегда-иногда-никогда»

### 3.1. Описательная статистика и проверка нормальности для трех основных итоговых шкал. Нормы для пятых классов

Основными в данной методике являются шкалы, измеряющие компетентность по русскому языку, математике и по обоим предметам в целом. Для каждой из них по объединенной выборке ЦАО и школы № 1567 ЗАО города Москвы были построены гистограммы и проведена проверка на нормальность по критерию Колмогорова — Смирнова. Все три распределения отличаются от нормального ( $p < 0,001$ ), что можно предположить и визуально по гистограммам.

Таблица 2

Распределения ответов по пункту «Русский язык, 1б» в группах учащихся ЦАО и школы № 1567, правильно ответивших на вопрос «Русский язык, 1а»

Группа		Русский язык, 1б			Всего
		Неверно	Верно	Отказ	
ЦАО Москва	Количество	177	172	46	395
	%	44,8	43,5	11,6	100,0
Школа № 1567	Количество	4	23	0	27
	%	14,8	85,2	0,0	100,0

Так, например, на гистограмме (см. вкладку) можно увидеть левостороннюю асимметрию распределения показателей по итоговой шкале, что говорит о некотором преобладании более низких результатов в выборке. Тем не менее стандартизацию выборку следует признать репрезентативной, поскольку она отражает характерные для генеральной совокупности равные пропорции учащихся обоего пола, преобладание пятиклассников 11-летнего возраста с включением меньшей доли 10-летних и незначительного процента 12–13-летних, широкий охват детей разных национальностей. О репрезентативности свидетельствует также ее весьма значительный объем и явно выраженный случайный характер отбора респондентов.

Кроме того, для трех итоговых шкал были вычислены следующие статистики: среднее, стандартное отклонение, асимметрия, эксцесс, минимальное и максимальное значения в выборке, а также квантили и 16-я и 84-я процентиля. Последние вычислялись, чтобы определить нормы для учащихся пятых классов. Известно, что для нормального распределения нормой обычно считаются показатели, лежащие в пределах  $M \pm \sigma$  для стандартизированной выборки, между которыми находится 68 % выборки — по 34 % слева и справа от среднего, совпадающего также с медианой, т. е. 50-й процентилью. В таком случае для распределения, отличающегося от нормального, разумно ориентироваться на границы, определяемые  $50 \pm 34$  % выборки, т. е. считать нормой показатели, лежащие между 16-й и 84-й процентилями. Таким образом, нормой для учащихся пятых классов по шкале «Русский язык — Компетентность» следует считать диапазон от 3 до 8 баллов, по шкале «Математика — Компетентность» — диапазон от 3 до 11 баллов, а по шкале «Общий итог: Русский язык и Математика» — диапазон от 7 до 18 баллов (см. табл. 3).

Для всех трех итоговых шкал распределение объединенной выборки имеет положительную, т. е. левостороннюю асимметрию, как мы уже наблюдали на гистограмме (см. рис. 1). Эксцесс показывает, что самое «плосковершинное» распределение получает-

ся по шкале «Математика — Компетентность». По всем трем шкалам в выборке принимаются оба крайних значения диапазона.

Приведем здесь также аналогично рассчитанные нормы для остальных итоговых шкал: «Русский язык — Классификация» от 2 до 6 баллов, «Русский язык — Доказательство» от 1 до 2 баллов, «Математика — Классификация» от 3 до 8 баллов, «Математика — Доказательство» от 1 до 3 баллов, «Верная классификация — Всегда» от 1 до 3 баллов, «Верная классификация — Никогда» от 2 до 5 баллов, «Верная классификация — Иногда» от 2 до 6 баллов, «Доказательство — Иногда» от 1 до 6 баллов, «Классификация — Итог» от 6 до 14 баллов.

Поскольку далее производилась проверка статистической однородности выборок ЦАО и школы № 1567 г. Москвы, предварительно была проведена проверка нормальности обеих выборок по этим шкалам по критерию Колмогорова — Смирнова. Распределение показателей по каждой из трех итоговых шкал в ЦАО отличается от нормального ( $p < 0,001$ ), а в школе № 1567 является нормальным, что, по-видимому, объясняется значительно меньшим объемом последней выборки. Исходя из этого, для сравнения показателей обеих групп 5-классников был выбран непараметрический критерий Манна—Уитни [6; 7; 8].

### 3.2 Исследование влияния социодемографических факторов «категория учащихся» и «пол» на развитие метапредметных навыков

Результаты оценки влияния фактора «категория учащихся» на показатели, измеряемые итоговыми шкалами, отражены в табл. 4. По всем двенадцати шкалам показатели школы № 1567 достоверно превышают соответствующие показатели ЦАО ( $p < 0,001$ ), что вполне согласуется с ожиданиями.

Шкалы «Русский язык — Доказательство», «Математика — Доказательство» и «Верная классификация — Всегда» имеют небольшой диапазон (5–6 градаций), поэтому помимо сравнения двух групп испытуемых по уровню измеряемого признака по критерию Манна—Уитни мы сопоставили также распределения показате-

Таблица 3  
Описательные статистики для трех итоговых шкал по объединенной выборке ЦАО и школы № 1567 Москвы

Статистика	Процентиль	Русский язык — Компетентность	Математика — Компетентность	Общий итог: Русский язык и Математика
N		1472	1472	1472
Среднее		5,42	7,07	12,50
Стандартное отклонение		2,80	3,56	5,56
Асимметрия		0,62	0,27	0,53
Эксцесс		0,13	-0,60	-0,21
Минимум		0,00	0,00	0,00
Максимум		15,00	15,00	30,00
Процентили	16	3,00	3,00	7,00
	25	3,00	4,00	8,00
	50	5,00	7,00	12,00
	75	7,00	10,00	16,00
	84	8,00	11,00	18,00

Таблица 4  
**Сравнительный анализ показателей ЦАО и школы № 1567 по итоговым шкалам опросника «Всегда—иногда—никогда» (по критерию Манна—Уитни)\***

Шкала	Средние	
	ЦАО	Школа № 1567
Русский язык – Классификация	4,5	5,7
Русский язык – Доказательство	0,8	1,8
Русский язык – Компетентность	5,3	7,5
Математика – Классификация	5,6	7,8
Математика – Доказательство	1,2	2,9
Математика – Компетентность	6,9	10,7
Верная классификация – Всегда	2,3	3,2
Верная классификация – Иногда	4,0	4,7
Верная классификация – Никогда	3,8	5,6
Доказательство – Иногда	2,0	4,7
Классификация – Итог	10,2	13,5
Общий итог: Русский язык и Математика	12,2	18,2

\*По всем шкалам различия значимы на уровне  $p < 0,001$ .

лей обеих групп по этим шкалам по критерию однородности Хи-квадрат и получили различия во всех трех случаях ( $p < 0,001$ ). Тенденция везде одна и та же: в ЦАО больший, чем в школе № 1567, процент учащихся набирает более низкие баллы и меньший процент – более высокие. Так, например, для шкалы «Верная классификация – Всегда» доли учащихся, получивших 0, 1 и 2 балла, в группе ЦАО больше на 3,6 %, 16,3 % и 15,1 % соответственно, чем в школе № 1567. Напротив, доли учащихся, получивших 3 и 4 балла, в группе ЦАО меньше на 6,4 % и 28,6 %, чем в школе № 1567.

С целью оценки влияния фактора «пол» на измеряемые показатели исследовалась статистическая однородность подвыборок мальчиков и девочек как

в объединенной выборке пятиклассников ЦАО и школы № 1567, так и в каждой из этих выборок отдельно. Поскольку распределение генеральных совокупностей мальчиков и девочек в ЦАО по всем итоговым шкалам отличается от нормального (Колмогоров – Смирнов,  $p < 0,001$ ), как и в школе № 1567 для многих шкал, то в качестве адекватного метода был снова выбран критерий Манна–Уитни. Результаты представлены в табл. 5.

Таким образом, средние показатели девочек в ЦАО Москвы и в объединенной совокупности достоверно превышают показатели мальчиков. Результаты пятиклассников школы № 1567 несколько отличаются от общей картины: здесь по шкалам, оценивающим компетентность на материале математики, различий нет. Кроме того, мальчики и девочки демонстрируют одинаковый уровень знаний при решении задач на классификацию с правильным ответом «всегда» или «никогда».

Для уточнения возможного влияния совокупности факторов на показатели по всем итоговым шкалам был также проведен двухфакторный дисперсионный анализ ANOVA  $2 \times 2$ , где независимыми факторами стали «пол», имеющий, очевидно, два уровня, и «категория учащихся» также с двумя уровнями – «ЦАО Москвы» и «школа № 1567». В качестве отклика рассматривались поочередно все характеристики, измеренные в итоговых шкалах методики «Всегда – иногда – никогда».

Напомним, что двухфакторный ANOVA проверяет три гипотезы: об отсутствии влияния 1-го фактора (о равенстве генеральных средних в группах мальчиков и девочек по всей объединенной совокупности в целом); об отсутствии влияния 2-го фактора (о равенстве генеральных средних в группах пятиклассников ЦАО и школы № 1567) и об отсутствии взаимодействия 1-го и 2-го факторов (влияние 1-го фак-

Таблица 5  
**Сравнительный анализ показателей девочек и мальчиков в объединенной выборке ЦАО и школы № 1567 и в каждой выборке отдельно по итоговым шкалам методики «Всегда – иногда – никогда» (по критерию Манна–Уитни)**

Шкала	Объединенная выборка			ЦАО г. Москвы			Школа № 1567 г. Москвы		
	Средние		p	Средние		p	Средние		p
	Д n <sub>1</sub> = 723	М n <sub>2</sub> = 748		Д n <sub>3</sub> = 696	М n <sub>4</sub> = 702		Д n <sub>5</sub> = 27	М n <sub>6</sub> = 46	
Русский язык – Классификация	4,97	4,23	***	4,91	4,16	***	6,41	5,28	**
Русский язык – Доказательство	1,04	0,63	***	0,99	0,58	***	2,41	1,41	***
Русский язык – Компетентность	6,00	4,86	***	5,90	4,75	***	8,81	6,70	***
Математика – Классификация	5,92	5,57	**	5,84	5,44	**	8,26	7,54	нз
Математика – Доказательство	1,40	1,24	*	1,33	1,14	**	3,22	2,80	нз
Математика – Компетентность	7,34	6,82	**	7,18	6,58	**	11,48	10,34	нз
Верная классификация – Всегда	2,47	2,28	**	2,43	2,23	***	3,37	3,11	нз
Верная классификация – Никогда	4,23	3,92	***	4,20	3,88	***	4,81	4,54	нз
Верная классификация – Иногда	4,20	3,60	***	4,11	3,50	***	6,48	5,17	**
Доказательство – Иногда	2,45	1,88	***	2,33	1,73	***	5,63	4,22	**
Классификация – Итог	10,90	9,80	***	10,75	9,61	***	14,67	12,83	*
Общий итог: Русский язык и Математика	13,35	11,68	***	13,08	11,33	***	20,30	17,04	*

Сокращения: Д – девочки, М – мальчики, p – уровень значимости

\* – различия значимы на уровне  $p < 0,05$ ; \*\* – различия значимы на уровне  $p \leq 0,01$ ; \*\*\* – различия значимы на уровне  $p \leq 0,001$ ; нз – различия не значимые.

тора не различается для разных уровней 2-го фактора, и наоборот). Если выявлено влияние какого-либо отдельного фактора, это значит, что есть значимые различия хотя бы между одной парой генеральных средних, соответствующих градациям этого фактора. Однако гораздо интереснее, если выявлено достоверное взаимодействие обоих факторов, тогда характер этого взаимодействия устанавливается при помощи графиков средних значений. Метод применяется при следующих предположениях: распределение зависимой переменной в сравниваемых генеральных совокупностях для каждого из имеющихся сочетаний градаций факторов должно быть нормальным с одинаковыми дисперсиями. Однако в специальной литературе [6; 4] отмечалось, что нарушение предположения о нормальности распределения не оказывает существенного влияния на результаты ANOVA, поэтому проверка корректности его применения сводится к проверке гипотезы однородности дисперсий по критерию Ливена.

Заметим, что применение этого метода не является в данной ситуации вполне корректным, поскольку практически для всех шкал условие равенства дисперсий не выполняется. При этом влияние фактора «категория учащихся» выявлено для всех шкал, что полностью повторяет результаты, полученные непараметрическими методами. Влияние фактора «пол» не выявлено только для четырех шкал из двенадцати: «Математика – Классификация» ( $p = 0,056$ , Ливен:  $p = 0,002$ ), «Математика – Доказательство» ( $p = 0,084$ , Ливен:  $p = 0,306$ , здесь применение ANOVA корректно); «Верная классификация – Всегда» ( $p = 0,105$ , Ливен:  $p = 0,003$ ) и «Верная классификация – Никогда» ( $p = 0,098$ , Ливен:  $p = 0,045$ ). Таким образом, здесь более адекватные характеру данных непараметрические методы оказались и более чувствительными. Что касается совокупного влияния обоих факторов, то оно было выявлено только на способность приводить доказательство на материале русского языка (шкала «Русский язык – Доказательство»,  $p = 0,014$ , критерий Ливена:  $p = 0,019$ ). Здесь разница в средних между мальчиками и девоч-

ками в школе № 1567 достоверно выше, чем в ЦАО города Москвы (см. вкладку).

### 3.3. Сравнительный анализ развития метапредметных навыков на материале русского языка и математики

Наконец, остался весьма интересный вопрос: различается ли умение пятиклассников классифицировать и доказывать на материале русского языка и математики? Какой из двух предметов лучше развивает эти навыки? Есть ли различия в уровне компетентности по обоим предметам? Такие вопросы правомерны, поскольку пары соответствующих шкал имеют одинаковый диапазон. Мы проверили гипотезу о статистической однородности двух связанных выборок отдельно для пятиклассников ЦАО и школы № 1567 города Москвы с помощью непараметрического критерия знаковых ранговых сумм Уилкоксона [6; 7]. Результаты представлены в табл. 6. В ней рядом с количеством положительных, отрицательных или нулевых сдвигов в скобках указан процент таких сдвигов в соответствующей выборке.

В обеих группах показатели развития умения классифицировать, доказывать и итоговый показатель достоверно выше по математике, чем по русскому языку ( $p < 0,001$ ). Таким образом, математика лучше развивает эти метапредметные навыки. Если сравнить эту тенденцию в двух группах, то в школе № 1567 она выглядит даже более ярко выраженной. Действительно, если в группе ЦАО доли положительных сдвигов показателей вида «Математика ... – Русский язык ...», отражающих превышение показателя по математике над показателем по русскому языку, равны 61 %, 39 % и 64 % по каждому навыку соответственно, то в школе № 1567 доли таких сдвигов составляют уже 62 %, 69 % и 88 %.

### Обсуждение и выводы

Статистический анализ данных апробации опросника Г.А. Цукерман «Всегда – иногда – никогда» поз-

Таблица 6

**Сравнение показателей развития метапредметных навыков на материале русского языка и математики в группах учащихся ЦАО и школы № 1567 (по критерию знаковых ранговых сумм Уилкоксона)**

Метапредметный навык	Знак сдвига	Сдвиги «Математика ... – Русский язык ...»		Общая тенденция	p
		ЦАО, n <sub>1</sub> = 1398 (100 %)	Школа № 1567, n <sub>2</sub> = 74 (100%)		
		Количество, %	Количество, %		
Классификация	-	295 (21)	7 (9)	Математика – Классификация > Русский язык – Классификация	***
	+	854 (61)	62 (84)		
	0	249 (18)	5 (7)		
Доказательство	-	251 (18)	9 (12)	Математика – Доказательство > Русский язык – Доказательство	***
	+	544 (39)	51 (69)		
	0	603 (43)	14 (19)		
Компетентность	-	331 (24)	6 (8)	Математика – Компетентность > Русский язык – Компетентность	***
	+	890 (64)	65 (88)		
	0	177 (12)	3 (4)		

\*\*\* – различия значимы на уровне  $p < 0,001$ .

воляет сделать следующие выводы. Проверка надежности шкал с помощью коэффициента  $\alpha$  Кронбаха показывает удовлетворительную или хорошую согласованность пунктов опросника с соответствующими шкалами: значение  $\alpha$  Кронбаха для большинства шкал лежит в пределах от 0,5 до 0,7, но может достигать и 0,8. Только для шкал «Русский язык — Доказательство» и «Верная классификация — Всегда» его значение падает до 0,4. Удаление пунктов опросника с целью повышения этого показателя нецелесообразно.

Исследование итоговых шкал показало, что распределение по трем основным шкалам «Русский язык — Компетентность», «Математика — Компетентность» и «Общий итог: Русский язык и Математика» в объединенной генеральной совокупности отличается от нормального и имеет левостороннюю асимметрию, поэтому было предложено определять нормы по 16-й и 84-й перцентильям. Нормой для учащихся пятых классов по шкале «Русский язык — Компетентность» следует считать диапазон от 3 до 8 баллов, по шкале «Математика — Компетентность» — диапазон от 3 до 11 баллов, а по шкале «Общий итог: Русский язык и Математика» — диапазон от 7 до 18 баллов. Предложены также нормы для остальных девяти итоговых шкал.

Проверка влияния фактора «категория учащихся» на показатели по всем 12 итоговым шкалам опросника устанавливает достоверно более высокие показатели учащихся гимназии № 1567 по сравнению с учащимися школ ЦАО г. Москвы. Статистический анализ отдельных пунктов опросника выявляет лучшие результаты в умении классифицировать учащихся школы № 1567 по пяти пунктам из десяти по русскому языку

и по девяти из десяти пунктов по математике, а в умении доказывать — по четырем пунктам из пяти по русскому языку и по всем пяти пунктам по математике.

Сравнительный анализ результатов девочек и мальчиков показывает достоверно более высокие результаты девочек в объединенной совокупности и отдельно по ЦАО по всем шкалам опросника, а для школы № 1567 — по всем шкалам, кроме трех шкал, оценивающих компетентность по математике, и еще по двум шкалам на классификацию с верным ответом «всегда» и «никогда» соответственно, где различий нет. Совокупное влияние факторов «пол» и «категория учащихся» подтверждается статистически только для шкалы «Русский язык — Доказательство», где разница в средних между мальчиками и девочками в школе № 1567 достоверно выше, чем в ЦАО.

Кроме того, получены статистически значимые различия в умении классифицировать, доказывать и в общей компетентности учащихся как в ЦАО, так и в школе № 1567: математика лучше развивает эти метапредметные навыки, чем русский язык.

Таким образом, статистически доказана надежность шкал опросника Г.А. Цукерман «Всегда — иногда — никогда» и его несомненная способность определять влияние факторов «пол» и «категория учащихся» на показатели по итоговым шкалам и отдельным пунктам. Определены репрезентативные нормы для итоговых шкал. Тем самым экспериментально подтверждена возможность его использования для оценки заявленных метапредметных результатов начального школьного образования в генеральной совокупности учащихся. В этом состоит практическая значимость данного исследования.

### Литература

1. Анастаси А. Психологическое тестирование. Кн. 1. М., 1982.
2. Астанина Н.Б. Создание и апробация нового личностного опросника «Методика изучения доверия к себе» // Психологическая наука и образование. 2010. № 3. [Электронная версия] [http://www.psyedu.ru/files/articles/psyedu\\_ru\\_2010\\_3\\_Astanina.pdf](http://www.psyedu.ru/files/articles/psyedu_ru_2010_3_Astanina.pdf)
3. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психологической диагностике. Киев, 1989.
4. Гласс Дж., Стенли Дж. Статистические методы в педагогике и психологии. М., 1976.
5. Митина О.В. Математические методы в психологии. М., 2009.
6. Наследов А.Д. Математические методы психологического исследования. СПб., 2006.
7. Сорокова М.Г. Математические методы в психологии: непараметрическая статистика. М., 2011.
8. Хлевная Е.А., Штроо В.А., Киселева Т.С. Экспериментальное исследование возможности развития эмоционального интеллекта // Психологическая наука и образование. 2012. № 3. [Электронная версия] [http://www.psyedu.ru/files/articles/psyedu\\_ru\\_2012\\_3\\_3018.pdf](http://www.psyedu.ru/files/articles/psyedu_ru_2012_3_3018.pdf)
9. Franke G.H. Die Symptom-Checkliste von Derogatis — Deutsche Version. Gottingen, 1995.
10. Franke G.H. Möglichkeiten und Grenzen im Einsatz der Symptom-Checkliste SCL-90-R. // Verhaltenstherapie & Psychosoziale Praxis. 2001. № 33.

# Employing Tsukerman's Inventory "Always – Sometimes – Never" for Assessing Metasubject Skills in Students of 5<sup>th</sup> Class

M.G. Sorokova

PhD in Psychology, PhD in Physics and Mathematics, professor at the Chair of Mathematics and Natural Sciences, Department of Information Technologies, Moscow State University of Psychology and Education

---

The transition to the new federal state standards in primary and elementary education makes the task of developing adequate tools for psychological and educational assessment a very important one. In the 2011/12 academic year the Psychological Institute of the Russian Academy of Education (PIRAE) together with the Moscow State University of Psychology and Education (MSUPE) carried out an extensive research entitled "Complex approach to the assessment of metasubject outcomes of primary school education according to the new educational standards" (leading researcher I.M. Ulanovskaya). One of the research tasks was to test a series of techniques for evaluating various aspects of competence in students that had been developed by psychologists at PIRAE on the material of school subjects. This paper focuses on the statistical analysis of the test outcomes for the "Always – Sometimes – Never" inventory developed by Galina A. Tsukerman. The inventory was used on a sample of 5-class students of the Central Administrative Okrug of Moscow (1398 children from 40 schools, 702 males and 696 females) and the Gymnasium №1567 of the Western Administrative Okrug of Moscow (74 children, 46 males, 27 females, the sex of 1 subject unspecified). The research proved the scales of inventory to be statistically reliable and revealed the effect of 'Sex' and 'Student Category' factors on scales in general and on particular points of the inventory. Representative norms for overall scales were defined as well. Also, it was found that the impact of Mathematics on the development of classification skills, the ability to prove and on general competence of schoolchildren is much greater than that of Russian Language lessons. Thus the possibility of using this inventory for assessing the specified metasubject skills in students of 5<sup>th</sup> classes has been experimentally proved.

**Keywords:** reliability, validity, standardized sample, chi-squared test, Kolmogorov-Smirnov test, Mann-Whitney test, Wilcoxon test, two-factor scheme of dispersion analysis.

## References

1. Anastazi A. Psihologicheskoe testirovanie. Kn. 1. M., 1982.
2. Astanina N.B. Sozdanie i aprobaciya novogo lichnostnogo oprosnika "Metodika izucheniya doveriya k sebe" // Psihologicheskaya nauka i obrazovanie. 2010. № 3. Elektronnaya versiya: [http://www.psyedu.ru/files/articles/psyedu\\_ru\\_2010\\_3\\_Astanina.pdf](http://www.psyedu.ru/files/articles/psyedu_ru_2010_3_Astanina.pdf)
3. Burlachuk L.F., Morozov S.M. Slovar'-spravochnik po psihologicheskoi diagnostike. Kiev, 1989.
4. Glass Dzh., Stenli Dzh. Statisticheskie metody v pedagogike i psihologii. M., 1976.
5. Mitina O.V. Matematicheskie metody v psihologii. M., 2009.
6. Nasledov A.D. Matematicheskie metody psihologicheskogo issledovaniya. SPb., 2006.
7. Sorokova M.G. Matematicheskie metody v psihologii: neparametricheskaya statistika. M., 2011.
8. Hlevnaya E.A., Shtroo V.A., Kiseleva T.S. Eksperimental'noe issledovanie vozmozhnosti razvitiya emocional'nogo intellekta // Psihologicheskaya nauka i obrazovanie. 2012. № 3. Elektronnaya versiya: [http://www.psyedu.ru/files/articles/psyedu\\_ru\\_2012\\_3\\_3018.pdf](http://www.psyedu.ru/files/articles/psyedu_ru_2012_3_3018.pdf).
9. Franke G.H. Die Symptom-Checkliste von Derogatis – Deutsche Version. Gottingen, 1995.
10. Franke G.H. Moglichkeiten und Grenzen im Einsatz der Symptom-Checkliste SCL-90-R. // Verhaltenstherapie & Psychosoziale Praxis. 2001. № 33.