

# Общая психология, психология личности, история психологии

УДК 159.923.35  
ББК Ю95

DOI: 10.14529/psy180101

## «ЗАКОН УМЕНЬШЕНИЯ ОТДАЧИ» СПИРМЕНА: ИССЛЕДОВАНИЕ НА МАСШТАБНЫХ РОССИЙСКИХ ВЫБОРКАХ

К.В. Сугоняев<sup>1</sup>, Ю.И. Радченко<sup>2</sup>

<sup>1</sup> Институт психологии Российской академии наук, г. Москва, Россия

<sup>2</sup> Военная академия Генерального штаба ВС РФ, г. Москва, Россия

Снижение уровня корреляций между тестами способностей в группах с более высоким уровнем интеллекта было открыто Ч. Спирменом в 20-е годы прошлого века и получило в современной психологической литературе обозначение «закон уменьшения отдачи» Спирмена (SLODR). Многочисленные попытки подтвердить данный феномен за рубежом дали противоречивые результаты. В России подобных исследований до настоящего времени не проводилось. На двух масштабных выборках численностью 11335 и 11934 человека, выполнявших одну и ту же батарею тестов способностей, и характеризующихся высокой гомогенностью, было подтверждено наличие эффектов, постулированных Спирменом. Используемая в исследовании технология анализа данных позволила исключить артефактное происхождение выявленных зависимостей. Полученные данные свидетельствуют о том, что эффекты SLODR проявляют себя на уровне групповых факторов способностей и являются высоко воспроизводимыми. Предложены гипотезы относительно возможных причин отрицательных результатов, полученных некоторыми авторами, и рекомендации относительно возможных направлений будущих исследований.

*Ключевые слова:* «закон уменьшения отдачи» Спирмена (SLODR), гипотеза дифференциации, общая когнитивная способность, фактор *g*, психометрические характеристики.

### Введение

Тесты способностей находят широкое применение в прогнозировании успешности обучения и профессиональной деятельности в промышленности, образовании, коммерческих и государственных структурах, в том числе и в силовых (Campbell, 2001; Gottfredson, 2002; Шнайдер, 2004). Среди множества индивидуально-психологических характеристик именно общая когнитивная способность (ОКС или фактор *g*) является наиболее валидным предиктором профессиональной пригодности кандидатов (Ree, 1992; Salgado, 2003; Schmidt, 2004; Kuncel, 2010; Drasgow, 2012). Этот факт – в сочетании с относительной экономичностью психометрического тестирования – обеспечивает тестам способностей статус обязательного компонента любой тестовой батареи, предназначенной для отбора персонала (Schmidt, 1998, 2016; Le, 2007; Ones, 2017). В последние годы наблюдается бурный рост интереса к новым сферам при-

менения тестов ОКС: от прогнозирования состояния здоровья и продолжительности жизни [Deary, 2010; Christensen, 2016] до оценки роли когнитивных факторов в социально-экономическом развитии стран и регионов (Lynn, 2002, 2012; Meisenberg, 2011; Ushakov, 2016; Grigoriev, 2016).

Существуют, однако, различные подходы к оценке ОКС: от применения одного-двух когнитивных тестов с высокой нагрузкой на фактор *g* до использования комплексных тестовых батарей способностей (ТБС), включающих обычно от 8 до двух десятков субтестов. Среди наиболее известных зарубежных ТБС такого рода можно назвать тесты Д. Векслера (WISC, WAIS), ASVAB, SAB, WIT, WJ COG и др. В России, наряду с русскоязычными версиями зарубежных ТБС (Д. Векслера, Р. Амтхауэра и др.), нашли применение и отечественные разработки – ШТУР, УИТ СПЧ, ТИПС и др. Вне зависимости от того, опираются эти ТБС на иерархиче-

ские или на многофакторные модели интеллекта, их состав характеризуется высокой степенью сходства, а прогностичность рассчитываемых в них интегральных оценок оказывается в большей степени зависимой от психометрического качества, нежели от конкретного содержания тестов, т. е. многокомпонентные ТБС являются взаимозаменяемыми (Johnson, 2004; Salthouse, 2014).

Учитывая, что ТБС применяются для прогнозирования значимых аспектов поведения, практически важным является вопрос о том, на какую оценку (или оценки) следует опираться при формировании прогностического заключения. Достаточно ли рассчитать единственную интегральную оценку (например, IQ) или же можно добиться более высокой валидности прогноза на основе учета *нескольких* оценок (за отдельные субтесты или группы субтестов)? Мнения на этот счет существенно разнятся. Хотя имеется немало сторонников важности учета частных способностей (Mount, 2008; Lang, 2010; Stanhope, 2014; Reeve, 2015; Kell, 2017), на рубеже веков возобладала точка зрения о предпочтительности использования в отборе персонала именно интегральных оценок ОКС, поскольку они являются более робастными<sup>1</sup> и охватывают большую часть вариативности критерия, тогда как прирост валидности, обеспечиваемый дополнительным включением в уравнение прогноза оценок специальных способностей, обычно не превышает нескольких процентов (Ree, 1994, 2015; Jensen, 1998; Carretta, 2000; Salgado, 2003; Kuncel, 2004; Schmidt, 2004; Brown, 2006).

На фоне дискуссии между сторонниками более интегративного либо более дифференцированного подхода к оценке способностей в последние 25 лет активизировался интерес к открытому Ч. Спирменом феномену снижения уровня интеркорреляций между тестами способностей (т. е. их нагруженности на фактор  $g$ ) в группах с более высоким уровнем интеллекта, названному им «законом уменьшения отдачи» (Spearman, 1927). Впоследствии данный феномен стал обозначаться в литературе как SLODR (Spearman's Law of

<sup>1</sup> Робастность (англ. *robustness*, от *robust* – «крепкий», «сильный», «твёрдый», «устойчивый») – свойство статистического метода, характеризующее независимость влияния на результат исследования различного рода выбросов, устойчивости к помехам (прим. ред.).

Diminishing Returns) или гипотеза дифференциации<sup>2</sup> (Deary, 1996). Предложено несколько объяснений данному феномену, среди которых: повышение когнитивной специализации с ростом уровня способностей (Deary, 1996); влияние формального образования (Abad, 2003); влияние диспозиционных свойств и интересов, направляющих перераспределение когнитивных усилий (Ackerman, 1996; Woodley, 2011) и др.

Выполненные к настоящему времени исследования SLODR дают противоречивые результаты. Большинство публикаций подтверждает факт ослабления взаимосвязей между субтестами в группах с высоким уровнем способностей (Detterman, 1989; Lynn, 1992; Deary, 1996; Legree, 1996; Abad, 2003; Jensen, 2003; Coyle, 2003; Reynolds, 2007, 2010; Tucker-Drob, 2009; McGill, 2015; Blanch, 2017; Molenaar, 2017; Blum, 2017). Более того, получены данные, свидетельствующие о том, что у представителей наций с более высоким средним уровнем IQ нагруженность частных тестов способностей и образовательных достижений на генеральный фактор<sup>3</sup> ниже, чем в нациях с менее высоким уровнем интеллекта (Coyle, 2013). Хотя величина различия оказалась невелика (0,04), авторы интерпретировали полученный результат как согласующийся с феноменом SLODR. Вместе с тем, даже в случаях положительных для SLODR результатов отмечается, во-первых, относительно небольшая величина эффекта (различия в средних уровнях интеркорреляций между субтестами в сравниваемых группах редко превышают 0,1–0,2), и, во-вторых, в ряде исследований эффект SLODR проявлял себя лишь для части тестов. Так, в исследовании А. Дженсена максимальные межгрупповые различия были зафиксированы для тестов с меньшей нагруженностью на фактор  $g$  (Jensen, 2003). Напротив, М. Рейнольдс и Т. Кит для тестов с низкой нагруженностью на фактор  $g$  эффекта SLODR не обнаружили (Reynolds, 2007). В исследовании Ф. Абада с

<sup>2</sup> Вслед за Спирменом Гэрретт распространил эту гипотезу на возрастные изменения когнитивных функций (Garrett, 1946), однако рассмотрение этого вопроса выходит за рамки настоящей публикации (прим. авт.).

<sup>3</sup> На национальном уровне фактор  $G$  трактуется авторами как сочетание интеллекта, знаний и интеллектуального использования полученных знаний (Coyle, 2013, p. 408).

соавт. этот эффект сильнее проявился в результатах обследования по батарее когнитивных тестов с более выраженным кристаллизованным компонентом интеллекта, чем по батарее тестов с выраженным флюидным его компонентом (Abad, 2003). Наконец, имеется немало работ, в которых эффекты SLODR подтвердить не удалось (Fogarty, 1995; Facon, 2004; Hartmann, 2006; Arden, 2007; Kroczek, 2016). Более того, в одной из недавних публикаций группа авторов поставила под сомнение эмпирические свидетельства в пользу SLODR, выдвинув предположение о том, что своим происхождением эффект обязан различиям в скошенности распределений оценок в сравниваемых группах, т. е. является статистическим артефактом (Murphy, 2013).

Следует признать, что ряд опубликованных исследований SLODR действительно уязвим для критики с точки зрения методологии. Некоторые из них характеризуются недостаточной статистической мощностью: при тех величинах эффекта, которые обычно регистрируются для индикаторов SLODR, необходимо, чтобы каждая из групп была представлена не менее чем тысячей испытуемых (Cohen, 1992). Деление выборки более чем на две группы также представляется неоптимальным, поскольку может приводить к существенному ограничению вариативности оценок и снижению любых величин эффекта, а также провоцировать появление искусственных отрицательных корреляций между субтестами (если деление осуществляется на основе интегральной оценки ОКС) (Jensen, 2003). В некоторых исследованиях надежность тестовых оценок в сравниваемых группах существенно ниже минимально приемлемого уровня 0,7, что ставит вопрос о качестве данных и делает проблематичным выявление искомым эффектов. Многие исследования выполнялись на данных, полученных в условиях добровольного (low stake) тестирования. В качестве исследовательских выборок нередко используются выборки стандартизации известных тестовых батарей (характеризующихся, обычно, очень высокой гетерогенностью), например, детские и подростковые выборки, отличающиеся нестабильностью структуры способностей (Martins, 2016). Полученные в подобных случаях результаты не только irrelevantны контексту отбора персонала, но и могут быть в существенной мере загрязнены влиянием неконтролируемых переменных, в частности, возраста.

Мюррей с соавт. называют в качестве возможных причин получения противоречивых результатов также использование субтестов с малым числом пунктов, с чрезмерной представленностью слишком простых или слишком сложных заданий, а также различия в форме распределения тестовых оценок в сравниваемых группах и др. (Murphy, 2013).

Значительную роль в уточнении факта наличия или отсутствия эффекта SLODR могли бы сыграть данные, полученные в ходе обследования кандидатов на военную службу. Подобные обследования проводятся практически во всех странах мира, охватывают значительные по численности контингенты и нередко предусматривают применение достаточно диверсифицированных батарей когнитивных тестов. Однако нам удалось найти гораздо меньше публикаций такого рода, чем можно было бы ожидать; к тому же они не вносят ясности в вопрос о реальности SLODR. В частности, известны работы с использованием данных, полученных с помощью тестовой батареи ASVAB (Батарея профессиональных способностей для вооруженных сил США), в одной из которых (свыше 9000 респондентов) эффект SLODR получил подтверждение (Legree, 1996), а в другой ( $N = 6980$ ) – нет (Hartmann, 2006).

В России, насколько нам известно, подобные исследования до сих пор не проводились, поэтому представляло бы интерес выяснить, проявляется ли эффект SLODR на отечественной популяции, и если да, то в какой степени? Получение ответов на эти вопросы потребовало обращения к данным, полученным в ходе обследования кандидатов на обучение в образовательных учреждениях Минобороны России. При проведении отборочных мероприятий с этим контингентом регламентировано применение ТБС «КР-3» в составе 10 субтестов, ядро которой было разработано в 80-х годах прошлого века Б.В. Кулагиным и М.М. Решетниковым (Методики..., 2005) и позже подверглось незначительному редактированию в ходе компьютеризации. Данная батарея достаточно диверсифицирована с точки зрения содержания тестовых заданий и необходимых для их выполнения когнитивных операций, поэтому использование полученных с ее помощью данных в контексте проблемы SLODR представляется вполне уместным. При проведении исследования мы стремились учесть предостережения, содер-

жащиеся в известных источниках, чтобы минимизировать влияние возможных методологических погрешностей.

Исследование проводилось в два этапа. На первом предполагалось установить факт наличия или отсутствия эффекта SLODR на данных, полученных в ходе обследования кандидатов одного года набора. На втором этапе планировалось кросс-валидизировать полученные результаты на данных, полученных в ходе обследования кандидатов другого набора.

### Методы

#### Выборка

Для того чтобы исключить возможные влияния на результаты исследования временных трендов в тестовых оценках («эффекта Флинна»), было решено ограничить анализ данными, полученными в ходе обследования абитуриентов одного набора. По итогам набора 2013 года было получено 12827 протоколов, однако после исключения данных трех вузов, в которых были выявлены признаки нарушения регламента обследования, протоколов обследования девушек и повторных обследований, осталось 11335 протоколов, представляющих абитуриентов 21 вуза, расположенных в 18 регионах Российской Федерации.

Таким образом, исследовательская выборка состояла только из мужчин сопоставимого образовательного уровня (среднее общее или среднее профессиональное образование) в возрасте от 16 до 28 лет ( $M = 18,42 \pm 1,29$ ), при этом около 80 % выборки

было представлено юношами в возрасте 17 и 18 лет. Подавляющее большинство респондентов представляли выпускники средних общеобразовательных школ (88,9 %), остальные являлись выпускниками суворовских или кадетских училищ (2,57 %) либо военнослужащими (8,53 %).

#### Инструментарий

Тестовая батарея «КР-3» состоит из 10 тестов, краткая характеристика которых представлена в табл. 1. Все тесты состоят из одинакового числа заданий ( $n=30$ ), каждое из которых предполагает выбор единственного правильного ответа среди 5 вариантов.

Все тесты выполняются в течение одной сессии в контролируемом групповом бланковом формате в той последовательности, в которой они представлены в табл. 1. Поскольку во всех вузах применяются одни и те же средства автоматизации тестирования, все полученные тестовые материалы полностью идентичны. Свои ответы респонденты фиксируют на стандартных регистрационных бланках, которые в дальнейшем вводятся в программу обработки посредством сканирования.

Все включенные в анализ протоколы были полными, т. е. включали результаты выполнения всех 10 методик. Каждый электронный протокол содержал не только итоговые оценки выполнения тестов, но и ответы на отдельные задания, что позволяло контролировать психометрические характеристики как тестовых оценок, так и отдельных заданий.

Таблица 1

Характеристика методик тестовой батареи «КР-3»

Наименование методики	Краткая характеристика	Временной лимит (мин)
Аналогии	Оценка способности к пониманию характера логических связей и отношений между понятиями	5
Числовые ряды	Оценка способности к индуктивному числовому мышлению	8
Память на фигуры	Оценка кратковременной зрительной памяти на абстрактные контурные изображения	4
Узоры	Оценка способности к мысленному конструированию целостного образа объекта по его части	8
Арифметический счет	Оценка практических навыков устного счета	5
Вербальная память	Оценка кратковременной зрительной памяти на слова	4
Установление закономерности	Нахождение лексического эквивалента для знаковой последовательности	8
Силлогизмы	Оценка способности к логическому рассуждению (индуктивному выводу)	10
Исключение слова	Оценка способности к обобщению при оперировании вербальной информацией	8
Кубы	Оценка способности к мысленному вращению объекта в трех измерениях	10

Таблица 2

Количественные индексы, характеризующие качество выполнения тестов в интегральной выборке

Тест	Статистики				
	$N_{пр}$	СКО	Асимметрия	Экссесс	ОМН
Аналогии (Ан)	18,786	5,451	-0,309	-0,281	0,862
Числовые ряды (ЧР)	16,371	3,958	-0,092	0,162	0,708
Память на фигуры (ПФ)	18,770	6,121	-0,302	-0,594	0,888
Узоры (Уз)	16,253	6,850	-0,079	-0,934	0,911
Арифметический счет (АС)	13,559	6,064	0,420	-0,221	0,885
Вербальная память (ВП)	20,482	5,476	-0,361	-0,291	0,862
Установление закономерности (УЗк)	22,898	4,464	-1,037	1,801	0,844
Силлогизмы (Сил)	15,600	4,903	-0,223	-0,233	0,806
Исключение слова (ИС)	20,190	4,418	-0,710	0,617	0,756
Кубы (Кб)	14,502	5,572	0,125	-0,439	0,816

Условные обозначения. Здесь и далее:  $N_{пр}$  – среднее число правильно решенных заданий; СКО – среднеквадратическое отклонение; ОМН – одномоментная надежность тестовых оценок (коэффициент Кьюдера – Ричардсона № 20).

### Способ деления выборки

Из множества возможных способов деления выборки на группы, различающиеся по уровню интеллекта, предпочтение было отдано рекомендованному А. Дженсеном медианному делению пополам на основе интегральной оценки ОКС (Jensen, 2003). Последняя рассчитывалась как средняя десяти стандартизированных оценок «эффективности» выполнения каждого теста<sup>4</sup>. По результатам выполнения каждого теста автоматически рассчитывались стандартизированные оценки (стэны) для всех первичных показателей в соответствии с технологией форсированной процентильной нормализации (Шмелев, 2013, с. 271–274). По стэновым оценкам эффективности десяти тестов рассчитывалась средняя, медианное значение которой использовалось для деления выборки на две части. Поскольку одинаковую медианную оценку 5,6 получили 309 человек, для деления выборки пополам была принята во внимание также средняя сы-

рая оценка эффективности, являющаяся континуальной величиной; граничной оказалась оценка 42,17. Интегральные оценки ОКС, на основе значений которых осуществлялось деление выборки на «высокую» и «низкую» группы, были распределены нормально: асимметрия и эксцесс распределения сырых оценок ОКС составили, соответственно, -0,059 и -0,144; аналогичные данные по стандартизированным оценкам: -0,211 и 0,019.

Для получения равных по численности групп ( $N = 5667$ ) медианный протокол был исключен. Далее по тексту сравниваемые группы обозначаются как «высокая» и «низкая».

В качестве возможных индикаторов эффектов SLODR рассчитывались: средний уровень межтестовых корреляций; доля дисперсии, объясняемой первым фактором; средний уровень межфакторных корреляций. Все вторичные статистические расчеты выполнялись с нормализованными стандартными оценками продуктивности<sup>4</sup> ( $N_{пр}$ ). Deskриптивные статистики рассчитывались с помощью табличного процессора Microsoft Excel; корреляционный и факторный анализ данных выполнялся с использованием пакета прикладных программ StatSoft Statistica v. 8.0.

## Результаты исследования

### 1. Психометрические характеристики интегральной выборки

Основные дескриптивные статистики и индексы надежности тестовых оценок, рассчитанные на интегральной выборке, представлены в табл. 2.

<sup>4</sup> Оценка «эффективность» представляет собой нормированную версию «классической» оценки успешности выполнения когнитивного теста (сумма правильно решенных задач за фиксированное время выполнения; далее по тексту – *продуктивность*), скорректированную с учетом вероятности случайного угадывания. В результате данного преобразования «сырая» оценка эффективности принимает значения от 0 до 100, где 0 соответствует качеству решения, не отличающемуся от случайного угадывания (20 % правильных ответов), а оценка 100 соответствует максимально возможному результату (применительно к тестам «КР-3» – 30 правильных ответов из 30 возможных).

## Общая психология, психология личности, история психологии

По представленным в табл. 2 данным, большинство тестов батареи имеют низкую или умеренную асимметрию распределения оценок и вполне respectable надежность. Тем не менее, можно отметить более выраженную левостороннюю асимметрию оценок методики «Установление закономерности» (причиной которой является преобладание легких для данного контингента заданий) и недостаточную с точки зрения ведомственного стандарта качества тестов (Методические указания..., 2014) надежность оценок в методиках «Числовые ряды» (причина – дефицит времени для выполнения) и «Исключение слова» (причина – высокая гетерогенность заданий) (Сугоняев, 2013). Межтестовые корреляции на интегральной выборке (без коррекции на ошибку измерения) представлены в табл. 3.

При факторизации тестовых оценок (без вращения, метод главных осей) первый фактор описывал 45,18 % общей вариативности оценок; он был принят в качестве меры ОКС (g) (Ree, 1994; Jensen, 2003). Нагрузки тестовых оценок на фактор g представлены в табл. 4 (графа 1ФР). По представленным в ней данным, максималь-

ные нагрузки на фактор g имеют вербальные и математические тесты, а наименьшие – тесты памяти; тесты пространственных способностей занимают промежуточное положение.

График собственных нагрузок факторов («каменистая осыпь») указывал на возможность выделения четырех факторов. Действительно, результаты 4-факторного косоугольного вращения оказались наиболее легко интерпретируемыми: в этом случае выделенные размерности легко отождествлялись с известными групповыми факторами интеллекта. В табл. 4 в графе «4ФР» представлено распределение факторных нагрузок тестов на выделенные факторы, идентифицированные как математические способности (№ 1), память (№ 2), пространственные (№ 3) и вербальные способности (№ 4).

### 2. Психометрические характеристики сравниваемых групп

В табл. 5 и 6 представлены основные психометрические индексы, характеризующие качество выполнения тестов в «высокой» и «низкой» группах.

Межтестовые корреляции на интегральной выборке

Таблица 3

Тест	Ан	ЧР	ПФ	Уз	АС	ВП	УЗк	Сил	ИС
ЧР	0,506								
ПФ	0,356	0,334							
Уз	0,441	0,443	0,360						
АС	0,447	0,590	0,306	0,338					
ВП	0,322	0,309	0,470	0,272	0,315				
УЗк	0,465	0,483	0,332	0,502	0,412	0,347			
Сил	0,534	0,476	0,293	0,402	0,407	0,317	0,461		
ИС	0,486	0,401	0,283	0,340	0,341	0,324	0,420	0,517	
Кб	0,354	0,336	0,294	0,480	0,267	0,242	0,425	0,365	0,303

Факторные нагрузки тестовых оценок при разных вариантах факторизации в интегральной выборке

Таблица 4

Тест	Фактор				
	1ФР	4ФР			
	g	1	2	3	4
Аналогии	-0,745	0,392	0,176	0,278	<b>0,590</b>
Числовые ряды	-0,740	<b>0,765</b>	0,134	0,267	0,279
Память на фигуры	-0,581	0,169	<b>0,796</b>	0,265	0,075
Узоры	-0,686	0,256	0,160	<b>0,754</b>	0,179
Арифметический счет	-0,664	<b>0,858</b>	0,177	0,104	0,179
Вербальная память	-0,562	0,131	<b>0,827</b>	0,061	0,244
Установление закономерности	-0,732	0,347	0,190	<b>0,536</b>	0,359
Силлогизмы	-0,723	0,272	0,110	0,257	<b>0,728</b>
Исключение слова	-0,661	0,101	0,173	0,132	<b>0,846</b>
Кубы	-0,597	0,045	0,117	<b>0,831</b>	0,175

Условные обозначения: 1ФР – однофакторное решение; 4ФР – четырехфакторное решение.

Таблица 5  
Количественные индексы, характеризующие качество выполнения тестов в «высокой» группе

Методика	Статистики				
	N <sub>пр</sub>	СКО	Асимметрия	Экссесс	ОМН
Аналогии	21,808	4,256	-0,376	-0,387	0,816
Числовые ряды	18,499	3,215	0,134	0,289	0,610
Память на фигуры	21,579	5,116	-0,551	-0,184	0,868
Узоры	19,839	5,821	-0,505	-0,386	0,912
Арифметический счет	16,495	5,561	0,319	-0,385	0,905
Вербальная память	22,950	4,586	-0,515	-0,203	0,840
Установление закономерности	25,284	2,961	-0,763	0,955	0,730
Силлогизмы	18,243	3,850	-0,152	0,195	0,725
Исключение слова	22,461	3,232	-0,584	0,590	0,594
Кубы	17,020	5,159	-0,072	-0,330	0,809

Таблица 6  
Количественные индексы, характеризующие качество выполнения тестов в «низкой» группе

Методика	Статистики				
	N <sub>пр</sub>	СКО	Асимметрия	Экссесс	ОМН
Аналогии	15,764	4,802	-0,226	-0,075	0,775
Числовые ряды	14,242	3,456	-0,139	0,286	0,582
Память на фигуры	15,960	5,743	-0,086	-0,567	0,842
Узоры	12,669	5,851	0,200	-0,702	0,838
Арифметический счет	10,624	5,038	0,624	0,555	0,780
Вербальная память	18,014	5,173	-0,214	-0,149	0,810
Установление закономерности	20,513	4,440	-1,021	1,729	0,796
Силлогизмы	12,957	4,393	-0,160	-0,384	0,727
Исключение слова	17,918	4,276	-0,649	0,453	0,706
Кубы	11,984	4,775	0,209	-0,252	0,732

Несмотря на заметное ограничение (по сравнению с интегральной выборкой) вариативности практически всех тестовых оценок (в среднем на 17,9 % – в высокой группе и на 10,0 % в – низкой), большинство (17 из 20) оценок ОМН сохранили значения выше 0,7. Разделение выборки пополам не привело к существенным изменениям в средних уровнях асимметрии распределения оценок, а показатели эксцесса в среднем даже несколько снизились.

Для того, чтобы полностью исключить возможное влияние на результаты сопоставления групп различий в вариативности и форме распределения тестовых оценок, перед проведением корреляционного и факторного анализа оценок в группах была выполнена форсированная процентильная нормализация данных *раздельно для каждой из групп*. В результате этой процедуры среднеквадратические отклонения нормализованных тестовых оценок в двух группах стали практически

одинаковыми (1,95 и 1,96)<sup>5</sup>. Межтестовые корреляции в сравниваемых группах (без коррекции на ошибку измерения) представлены в табл. 7 и 8; результаты факторизации – в табл. 9 и 10.

### 3. Анализ собственно эффектов SLODR

В табл. 11 сведены основные индексы, которые могут отражать эффекты, постулируемые SLODR, а именно: средний уровень межтестовых корреляций (рассчитанных как средняя арифметическая 45 коэффициентов) и по методу Г. Кайзера (Kaiser, 1968; Jensen, 2003); собственные значения генерального фактора и разброс факторных нагрузок в однофакторном решении; средний уровень межфакторных корреляций при косоугольном вращении четырех факторов.

Все представленные в табл. 11 индексы свидетельствуют в пользу реальности большей

<sup>5</sup> Средние составили, соответственно, 5,28 и 5,34. Несущественное отклонение от теоретического значения 5,5 связано с дискретностью исходных оценок.

Таблица 7

Межгестовые корреляции в «высокой» группе\*

Тест	Ан	ЧР	ПФ	Уз	АС	ВП	УЗк	Сил	ИС
ЧР	0,260								
ПФ	0,119	0,078							
Уз	0,224	0,216	0,159						
АС	0,226	0,455	0,091	0,090					
ВП	0,058	0,040	0,318	0,033	0,091				
УЗк	0,211	0,247	0,072	0,319	0,178	0,069			
Сил	0,319	0,216	0,039	0,170	0,156	0,050	0,209		
ИС	0,241	0,104	0,016	0,072	0,073	0,050	0,132	0,272	
Кб	0,160	0,125	0,115	0,364	0,060	0,018	0,234	0,153	0,057

\* Все коэффициенты более 0,034 статистически значимы ( $p < 0,01$ ).

Таблица 8

Межгестовые корреляции в «низкой» группе

Тест	Ан	ЧР	ПФ	Уз	АС	ВП	УЗк	Сил	ИС
ЧР	0,347								
ПФ	0,172	0,160							
Уз	0,225	0,241	0,179						
АС	0,282	0,445	0,127	0,141					
ВП	0,139	0,134	0,354	0,067	0,151				
УЗк	0,260	0,296	0,163	0,309	0,230	0,205			
Сил	0,380	0,314	0,106	0,190	0,247	0,156	0,291		

Таблица 9

Факторные нагрузки тестовых оценок при разных вариантах факторизации в высокой группе

Методика	Фактор				
	1ФР	4ФР			
	<i>g</i>	1	2	3	4
Аналогии	-0,616	0,279	0,086	0,220	<b>0,577</b>
Числовые ряды	-0,612	<b>0,800</b>	-0,006	0,185	0,131
Память на фигуры	-0,303	0,041	<b>0,792</b>	0,179	-0,019
Узоры	-0,571	0,082	0,072	<b>0,783</b>	0,074
Арифметический счет	-0,512	<b>0,850</b>	0,087	-0,009	0,072
Вербальная память	-0,214	0,040	<b>0,821</b>	-0,068	0,079
Установление закономерности	-0,583	0,254	0,017	<b>0,554</b>	0,201
Силлогизмы	-0,554	0,140	-0,011	0,172	<b>0,696</b>
Исключение слова	-0,377	-0,070	0,021	-0,044	<b>0,802</b>
Кубы	-0,467	-0,048	0,031	<b>0,769</b>	0,045

Таблица 10

Факторные нагрузки тестовых оценок при разных вариантах факторизации в низкой группе

Методика	Фактор				
	1ФР	4ФР			
	<i>g</i>	1	2	3	4
Аналогии	-0,637	0,393	0,076	0,152	<b>0,545</b>
Числовые ряды	-0,646	<b>0,763</b>	0,058	0,182	0,213
Память на фигуры	-0,407	0,108	<b>0,795</b>	0,182	-0,016
Узоры	-0,498	0,204	0,039	<b>0,756</b>	0,012
Арифметический счет	-0,543	<b>0,836</b>	0,094	0,013	0,093
Вербальная память	-0,403	0,045	<b>0,822</b>	-0,015	0,178
Установление закономерности	-0,621	0,235	0,169	<b>0,528</b>	0,301
Силлогизмы	-0,640	0,202	0,021	0,163	<b>0,743</b>
Исключение слова	-0,557	0,003	0,112	0,049	<b>0,824</b>
Кубы	-0,418	-0,094	0,047	<b>0,772</b>	0,129



Таблица 11

Сравнение статистических индексов в высокой и низкой группах

Группы	Средние межтестовые корреляции		Собственные значения фактора $g$	Вариативность факторных нагрузок в 1ФР	Средние межфакторные корреляции в 4ФР
	ср. арифм.	по Кайзеру			
Низкая	0,213	0,220	2,976	0,095	0,313
Высокая	0,155	0,165	2,487	0,132	0,204
Соотношение Н/В	1,374	1,333	1,197	0,720	1,534

дифференциации способностей в группах обследованных с более высоким уровнем ОКС. К этому можно добавить, что 41 из 45 межтестовых корреляций в «высокой» группе оказалась меньше, чем в «низкой», причем три исключения из четырех были связаны с тестами оценки пространственных способностей (см. табл. 7, 8). Изменение статуса пространственных способностей в группе с высоким уровнем ОКС проявилось и в том, что в однофакторном решении нагрузки тестов «Кубы» и «Узоры» на генеральный фактор возросли, тогда как нагрузки всех прочих тестов (особенно тестов оценки памяти) снизились; соответственно возросла вариативность нагрузок отдельных тестов на фактор  $g$ .

Данный результат был получен после локальной нормализации оценок для каждой группы – достаточно времязатратной процедуры, проведение которой было мотивировано стремлением нивелировать возможные искажения вследствие различий в вариативности тестовых оценок. Тем не менее представляло интерес уточнить, в какой мере эти различия могли повлиять на количественные индикаторы SLODR.

Для ответа на этот вопрос были повторены все виды анализа, результаты которых представлены в табл. 7–11, с использованием стэнов, рассчитанных только для интегральной выборки. Результаты оказались в существенной мере аналогичными: хотя все коэффициенты, указывающие на более высокий уровень связности частных способностей в «низкой» группе и их большую дифференцированность в «высокой», слегка подросли (рост составил от 1 до 2,5 %). Этот результат подтвердил известную робастность коэффициента корреляции Пирсона к нарушениям требования нормальности распределения (особенно на больших выборках) (Navlicek, 1977; Bishara, 2017). Полученный результат позволил провести кросс-валидизационное исследование на другой выборке аналогичного происхождения и сопоставимого объема,

не прибегая к трудоемкой процедуре локальной нормализации для каждой из групп.

#### 4. Кросс-валидизационное исследование

Для проверки воспроизводимости результатов, полученных на первом этапе исследования, анализ был повторен на массиве данных, полученных в ходе обследования кандидатов на обучение в 25 военных вузах страны в 2014 году. Поскольку объем доступных данных в этом исследовании оказался больше (свыше 15,5 тыс. протоколов), а также, учитывая содержащиеся в ряде публикаций указания на возможность изменения степени дифференциации способностей в зависимости не только от уровня интеллекта (SLODR-ability), но и от возраста (SLODR-age), было решено включить в анализ протоколы только 17- и 18-летних кандидатов ( $N=11934$ , средний возраст  $17,91 \pm 0,47$  лет). Предполагалось, что в случае подтверждения эффекта на еще более однородной по возрасту выборке, можно будет с большей степенью уверенности интерпретировать его в контексте различий в уровне ОКС, а не в возрасте (Deary, 1996; Tucker-Drob, 2009; Blum, 2017). На этой интегральной выборке были рассчитаны основные дескриптивные статистики для показателей «продуктивность» и индексы их одномоментной надежности (табл. 12); все они оказались весьма близки к индексам, рассчитанным для выборки 2013 года (см. табл. 2).

В интегральной выборке средний уровень корреляций между стандартизированными оценками продуктивности выполнения тестов составил 0,385, при этом генеральный фактор описывал 44,9 % вариативности оценок. В 4-факторном решении (с косоугольным вращением осей) средний уровень межфакторных корреляций составил 0,523.

Деление выборки пополам осуществлялось описанным выше способом. В целях уменьшения загруженности публикации табличными данными результаты корреляционного и факторного анализа интегральной и

Таблица 12

Количественные индексы, характеризующие качество выполнения тестов в интегральной выборке 17–18-летних кандидатов 2014 года (N = 11934)

Тест	Статистики				
	N <sub>пр</sub>	СКО	Асимметрия	Экссесс	ОМН
Аналогии	18,730	5,402	-0,344	-0,141	0,858
Числовые ряды	16,687	4,017	-0,297	0,635	0,717
Память на фигуры	19,029	6,037	-0,344	-0,502	0,884
Узоры	16,613	6,793	-0,150	-0,862	0,910
Арифметический счет	13,729	5,875	0,402	-0,137	0,875
Вербальная память	20,541	5,443	-0,470	-0,073	0,858
Установление закономерности	22,915	4,587	-1,276	2,617	0,851
Силлогизмы	16,189	4,766	-0,358	-0,008	0,798
Исключение слова	20,553	4,468	-0,811	0,964	0,766
Кубы	14,456	5,571	0,125	-0,376	0,815

Таблица 13

Сравнение статистических индексов в высокой и низкой группах, сформированных из числа 17–18-летних кандидатов 2014 года

Группы	Средние межтестовые корреляции		Собственные значения фактора g	Вариативность факторных нагрузок в 1ФР	Средние межфакторные корреляции в 4ФР
	ср. арифм.	по Кайзеру			
Низкая	0,239	0,245	3,205	0,087	0,344
Высокая	0,136	0,144	2,300	0,121	0,132
Соотношение низкая/высокая	1,757	1,701	1,393	0,719	2,606

парциальных выборок набора 2014 года здесь не приводятся<sup>6</sup>; а в табл. 13 представлены итоговые индексы, характеризующие эффекты SLODR в кросс-валидизационной выборке.

**Обсуждение**

Настоящее исследование выполнено с учетом актуализировавшегося в последние годы запроса на повышение статистической мощности публикуемых результатов и представление доказательств их воспроизводимости (Open Science..., 2015; LeBel, 2017; Lilienfeld, 2017).

По сравнению с абсолютным большинством публикаций по теме SLODR, исследованные нами выборки являются не только более многочисленными, но и значительно более гомогенными по возрасту, полу и социальному статусу. Они являются также и более гомогенными по уровню ОКС<sup>7</sup>, о чем свидетельствует тот факт, что дистанция между

средним уровнем оценок использованной меры ОКС в «высокой» и «низкой» группах составила 1,61 стандартного отклонения в выборке 2013 года и 1,59 в 2014 году, тогда как во многих публикациях по теме сообщается о более значительных различиях в уровнях IQ между сравниваемыми группами. Тем не менее все индексы, использованные для квантификации эффектов SLODR, недвусмысленно свидетельствуют в пользу его реальности.

Для проведения исследования авторы публикации воспользовались одним из наиболее популярных способов деления выборки – пополам на основе интегральной оценки ОКС, рассчитанной с учетом эффективности выполнения всех тестов батареи. Хотя большинство психометрических показателей в выделенных группах оказалось на сопоставимом уровне, вариативность тестовых оценок в высокой группе оказалась меньше, чем в низкой. Игнорирование этого факта могло бы поставить под сомнение любой результат в пользу SLODR, поскольку снижение тесноты связей между элементами тестовой батареи в высокой группе могло быть приписано психометрическому артефакту. Примененная в исследовании процедура раздельной форсирован-

<sup>6</sup> Соответствующие материалы могут быть предоставлены первым в перечне авторов исследователем по отдельному запросу.

<sup>7</sup> Предположительно – вследствие самоселекции кандидатов.

ной нормализации оценок в каждой из сравниваемых групп выборки 2013 года позволила исключить вероятность того, что выявленный эффект является следствием снижения вариативности тестовых оценок либо различий в форме их распределения (Murray, 2013). Достоинством данного способа нормализации исходных данных является также то, что он минимизирует влияние возможных аутлайеров<sup>8</sup> (т. е. действует аналогично винзоризации<sup>9</sup>) и тем самым устраняет еще один возможный источник артефактного происхождения эффектов SLODR (Tuğran, 2015).

Представленные в нижней строке табл. 11 и 13 значения индексов, отражающих изменения тесноты связей между частными оценками в «высокой» группе, имеют тот же порядок, что и приведенные в табл. 1 в статье А. Дженсена (Jensen, 2003). Следует учитывать, что трансформация сырых тестовых оценок (варьирующих в батарее «КР-3» в диапазоне 0–30 баллов) в шкалу нормализованных стэнов (1–10 баллов) неизбежно приводит к некоторому снижению величин эффекта вследствие ограничения вариативности переменных; в этом же направлении (ослабления значений эффекта) действует и неидеальная надежность тестовых оценок. Из этого следует, что истинная величина эффектов SLODR (на конструктивном уровне) может быть на 10–20 % больше той, которая представлена в табл. 11 и 13.

Сравнение этих таблиц свидетельствует о том, что величина эффектов SLODR в 2014 году оказалась несколько больше, чем в 2013. Можно предположить, что такой результат достигнут благодаря элиминации возрастных (и связанных с ними образовательных) эффектов, вектор которых не обязательно согласуется с действием факторов, детерминирующих SLODR-ability. Для получения свидетельств за или против этого предположения нужны дополнительные исследования.

<sup>8</sup> Выброс (англ. *outlier*), промах – в статистике результат измерения, выделяющийся из общей выборки (прим. ред.).

<sup>9</sup> Винзоризация (англ. *WinzORIZATION*) – вычисления среднего значения  $x$ , стандартного отклонения  $\sigma$ , операции срезки значений временного ряда, выпадающих за уровни  $x \pm 4\sigma$ , и повторения этой последовательности 3 операций до тех пор, пока значения  $x$  и  $\sigma$  не перестанут меняться (прим. ред.).

Факторизация тестовых оценок в общей и парциальных выборках приводила к получению весьма схожих матриц факторных нагрузок, причем конгруэнтность матриц между «общей» и «низкой» выборками неизменно оказывалась выше, чем между «общей» и «высокой»: 0,963 против 0,899 в 2013 г. и 0,948 против 0,846 – в 2014 г. Минимальной конгруэнтностью отличались матрицы, рассчитанные для «высокой» и «низкой» групп: 0,789 в 2013 г. и 0,768 в 2014 г. В то же время конгруэнтность матриц, рассчитанных для одноуровневых выборок *разных* лет, оказалась впечатляюще высокой: 0,970 (интегральные выборки 2013 и 2014 гг.), 0,984 (группы с низким уровнем ОКС) и 0,954 (группы с высоким уровнем ОКС). Эти цифры свидетельствуют о том, что организация способностей в «высокой» и «низкой» группах различна, и это различие отличается высокой степенью воспроизводимости.

Анализ результатов факторизации позволяет предположить, что основным проявлением SLODR является, возможно, не столько общее снижение влияния фактора  $g$  на частные оценки (в нашем случае оно оказалось не столь выраженным), сколько изменение соотношения между  $g$  и групповыми факторами интеллекта. В частности, в рамках использованной в исследовании батареи «КР-3» в «высокой» группе зафиксировано снижение тесноты связей между  $g$  и тестами памяти, но ее увеличение с тестами пространственных способностей. Предложенный нами в качестве дополнительного индикатора SLODR показатель «вариативность факторных нагрузок в 1ФР» в обеих выборках продемонстрировал в высокой группе рост на фоне снижения общего уровня нагрузок на  $g$ . Если наша гипотеза верна, то она позволяет объяснить наличие значительного числа исследований, в которых подтвердить SLODR с помощью недифференцированных уровневых индексов не удалось: по-видимому, это произошло из-за наложения разнонаправленных эффектов.

Еще одним фактором, затрудняющим выявление эффекта SLODR, на наш взгляд, может быть повышенная вероятность контаминации групп с низким уровнем способностей лицами, избирающими тактику хаотического реагирования – если не с самого начала тестирования, то вскоре после решения нескольких наиболее легких заданий, которые традиционно помещаются в начало теста. В пользу

такого предположения свидетельствует тот факт, что после разделения выборки пополам в «низкой» группе зафиксировано относительно большее снижение надежности оценок, чем в «высокой» (соответственно, 0,759 и 0,781; см. табл. 5 и 6), несмотря на большую их вариативность. Понятно, что вероятность присутствия таких респондентов в условиях мотивированного участия в административном тестировании (например, с целью отбора на престижную специальность – как это имеет место в нашем случае) невелика, но она может заметно возрастать в low stake ситуациях (Crede, 2010). С другой стороны, если включенные в ТБС тесты слишком просты для некоторых респондентов, может возрастать доля лиц, выполнивших все или большинство тестов на уровне, близком к максимальному. В таких случаях интеркорреляции между тестовыми оценками в этой части выборки могут оказаться завышенными, что может повлечь за собой подъем среднего уровня корреляций в высокой группе и тем самым замаскировать эффект SLODR<sup>10</sup>. Авторы настоящей публикации рекомендуют будущим исследователям SLODR контролировать качество выполнения тестов (что возможно только на основе сохранения в электронных протоколах всех ответов на задания) и предпринимать меры для чистки выборки от аутлайеров (Desimone, 2015), а также учитывать особенности конструирования включенных в батарею тестов и соответствие их трудности уровню когнитивного развития популяции.

В ряде публикаций, посвященных SLODR, обсуждаются возможные связи проявления эффекта с надежностью тестовых оценок или нагруженностью тестов на фактор  $g$  в интегральной выборке (Abad, 2003; Jensen, 2003; Reynolds, 2007 и др.). В нашем исследовании каких-либо направленных связей подобного рода выявлено не было. Важно отметить, что даже относительно меньшая измерительная точность тестов в низкой группе не смогла нивелировать большую тесноту связей между ними.

Тот факт, что свидетельства большей дифференциации способностей в группе с более высоким уровнем ОКС получены на двух

независимых выборках значительного объема с использованием нормализованных данных, опровергает тезис А. Мюррей с соавт. о том, что эффекты SLODR являются психометрическим артефактом, в частности, функцией от скошенности данных (Murray, 2013, p. 448). Более того, представленные в публикации данные, по нашему мнению, свидетельствуют о том, что эффекты SLODR связаны скорее с содержанием когнитивных операций (принадлежностью к тому или иному групповому фактору интеллекта), нежели с формальными психометрическими характеристиками тестовых оценок (надежностью, асимметрией распределения, нагруженностью на фактор  $g$  и т. п.).

Практическое значение феномена SLODR в контексте отбора персонала состоит в том, что если насыщенность частных тестов способностей фактором  $g$  снижается с ростом интеллектуального уровня кандидата, то единственная оценка результативности выполнения ТБС может уступать в точности прогноза частным оценкам по отдельным широким способностям в *верхней части* диапазона ОКС (McGill, 2015). Например, из представленных в настоящей публикации результатов следует, что интегральная оценка ОКС, рассчитанная по результатам батареи «КР-3», способна лишь в минимальной степени предсказать эффективность оперативной памяти в группе лиц с высоким уровнем интеллекта. В связи с этим в отборе персонала для наиболее сложных видов деятельности валидность комплекса оценок, учитывающих как  $g$ , так и групповые факторы способностей (например, в рамках мультипликативной модели профпригодности, Evans, 2000), может превосходить валидность единственной оценки ОКС. Косвенно в пользу такого вывода могут свидетельствовать результаты, полученные К. Канали (Canali, 2016).

### Ограничения

В описываемом исследовании не проверялось влияние на эффект SLODR альтернативных способов разделения выборки, в частности, на основе оценок по отдельным тестам батареи. Включенная в анализ выборка состояла исключительно из молодых мужчин, что оставляет под вопросом возможность генерализации полученных результатов на иные популяции. Рассчитанные величины эффектов SLODR не корректировались с учетом неидеальной надежно-

<sup>10</sup> В исследованных выборках не было лиц, безошибочно выполнивших все тесты: в выборке 2013 г. максимальная средняя оценка продуктивности составила 28,5; в выборке 2014 г. – 27,7.

сти тестовых оценок и ограничения диапазона их варьирования вследствие селективного характера выборок. Продолжительность выполнения всех тестов ограничивалась, причем жесткость ограничения в ряде случаев превышала оптимальную для тестов такого уровня (van der Linden, 2011; Kyllonen, 2016); данный факт мог приводить к искусственному завышению коэффициентов ОМН для некоторых тестов. Разделение выборки по полам не позволяет установить характер связи между выраженностью

эффектов SLODR и уровнем ОКС; она может быть как линейной, так и нелинейной (Lynn, 1992; McGill, 2015; Canali, 2016). Для получения ответов на эти вопросы необходимы дополнительные исследования.

**Статья подготовлена при поддержке Российского научного фонда, проект № 17-78-30035 «Психологические факторы экономической и социальной конкурентоспособности России».**

### Литература/References

1. Методики военного профессионального психологического отбора: методическое пособие. М.: Военное изд-во, 2005. 524 с. [*Metodiki voennogo professional'nogo psikhologicheskogo otbora: metodicheskoe posobie* [Methods of Military Personnel Selection: Textbook]. Moscow, Military Publishing, 2005. 524 p.]
2. Методические указания по порядку разработки, экспертизы (оценки соответствия психометрическим требованиям), внедрения и использования методик профессионального психологического отбора в Вооруженных Силах Российской Федерации. М.: ВАГШ ВС РФ, 2014. 45 с. [*Metodicheskie ukazaniya po poryadku razrabotki, ekspertizy (otsenki sootvetstviya psikhometricheskim trebovaniyam), vnedreniya i ispol'zovaniya metodik professional'nogo psikhologicheskogo otbora v Vooruzhennykh Silakh Rossiyskoy Federatsii* [The Procedure of Selection Tests Construction, Checking, Introduction and Application in Armed Forces of Russian Federation: Instructions]. Moscow, MAGS Publ., 2014. 45 p.]
3. Сугоняев, К.В. Анализ психометрических характеристик тестов, применяемых в отборе кандидатов на факультеты высшего профессионального образования вузов Минобороны России. I Всероссийский съезд психологов силовых структур Российской Федерации. Сборник материалов; под общ. ред. Ю.П. Зинченко, А.Г. Караяни, Ю.С. Шойгу. М.: НЦУКС МЧС России, 2013. С. 80–83. [Sugonyaev K.V. [Psychometric properties of tests provided for applicant selection for higher education faculties at military institutes of Russian Department of Defense]. *I Vserossiyskiy s'ezd psikhologov silovykh struktur Rossiyskoy Federatsii: sbornik materialov* [Proceedings of the 1st Conference of Military Psychologists of Russian Federation]. Moscow, 2013, pp. 80–83. (in Russ.)]
4. Шмелев, А.Г. Практическая тестология. Тестирование в образовании, прикладной психологии и управлении персоналом / А.Г. Шмелев. М.: Маска, 2013. 688 с. [Shmelyov A.G. *Prakticheskaya testologiya. Testirovanie v obrazovanii, prikladnoy psikhologii i upravlenii personalom* [Practical Testology. Testing in Education, Applied Psychology, and Personnel Management]. Moscow, Mask, 2013. 688 p.]
5. Шнайдер Б., Шмитт Н. Персонал для организации: научный подход к поиску, отбору, оценке и удержанию сотрудников. СПб.: Экономическая школа, 2004. 560 с. [Shneider B., Schmitt N. *Personal dlya organizatsii: nauchnyy podkhod k poisku, otboru, otsenke i uderzhaniyu sotrudnikov* [Staffing Organizations]. St. Petersburg, Economic School, 2004. 560 p.]
6. Abad F.J., Colom R., Juan-Espinosa M., Garcia L.F. Intelligence differentiation in adult samples. *Intelligence*, 2003, vol. 31, no. 2, pp. 157–166. DOI: 10.1016/S0160-2896(02)00141-1.
7. Ackerman P. A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, 1996, vol. 22, no. 2, pp. 227–257. DOI: 10.1016/S0160-2896(96)90016-1.
8. Arden R., Plomin R. Scant evidence for Spearman's law of diminishing returns in middle childhood. *Personality and Individual Differences*, 2007, vol. 42, no. 4, pp. 743–753. DOI: 10.1016/j.paid.2006.08.010.
9. Bishara, A.J., Hittner J.B. Confidence intervals for correlations when data are not normal. *Behavior Research Methods*, 2017, vol. 49, no. 1, pp. 294–309. DOI: 10.3758/s13428-016-0702-8.
10. Blanch A., Garcia H., Llaveria A., Aluja A. The Spearman's law of diminishing returns in chess. *Personality and Individual Differences*, 2017, vol. 104, pp. 434–441. DOI: 10.1016/j.paid.2016.09.003.

11. Blum D., Holling H. Spearman's law of diminishing returns: A meta-analysis. *Intelligence*, 2017, vol. 65, pp. 60–66. DOI: 10.1016/j.intell.2017.07.004.
12. Brown K.G., Le H., Schmidt F.L. Specific aptitude theory revisited: Is there incremental validity for training performance? *International Journal of Selection and Assessment*, 2006, vol. 14, no. 2, pp. 87–100. DOI: 10.1111/j.1468-2389.2006.00336.x.
13. Campbell J.P., Knapp D.J. (Eds.), *Exploring the Limits in Personnel Selection and Classification*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc. 2001. 637 p.
14. Canali K., Legree P., Meaden J. Specific abilities may increment psychometric g for high ability populations. *Conference paper and poster delivered at the 2016 SIOP Conference in Anaheim, California*, 2016. Available at: <http://www.dtic.mil/dtic/tr/fulltext/u2/1010433.pdf> (accessed: 29.11.2017).
15. Carretta T.R., Ree M.J. General and specific cognitive and psychomotor abilities in personnel selection: The prediction of training and job performance. *International Journal of Selection and Assessment*, 2000, vol. 8, no. 4, pp. 227–236. DOI: 10.1111/1468-2389.00152.
16. Christensen G.T., Mortensen E.L., Christensen K., Osler M. Intelligence in young adulthood and cause-specific mortality in the Danish Conscription Database – A cohort study of 728,160 men. *Intelligence*, 2016, vol. 59, pp. 64–71. DOI: 10.1016/j.intell.2016.08.001.
17. Cohen J. A power primer. *Psychological Bulletin*, 1992, vol. 112, no. 1, pp. 155–159. DOI: 10.1037/0033-2909.112.1.155.
18. Coyle T.R. IQ, the worst performance rule, and Spearman's law: A reanalysis and extension. *Intelligence*, 2003, vol. 31, no. 5, pp. 473–489. DOI: 10.1016/S0160-2896(02)00175-7.
19. Coyle T.R., Rindermann H. Spearman's law of diminishing returns and national ability // *Personality and Individual Differences*, 2013, vol. 55, no. 4, pp. 406–410. DOI: 10.1016/j.paid.2013.03.023.
20. Crede M. Random responding as a threat to the validity of effect size estimates in correlational research. *Educational and Psychological Measurement*, 2010, vol. 70, no. 4, pp. 596–612. DOI: 10.1177/0013164410366686.
21. Deary I.J. Cognitive epidemiology: Its rise, its current issues, and its challenges. *Personality and Individual Differences*, 2010, vol. 49, no. 4, pp. 337–343. DOI: 10.1016/j.paid.2009.11.012.
22. Deary I.J., Egan V., Gibson G.J., Austin E.J., Brand C.R., Kellaghan T. Intelligence and the differentiation hypothesis. *Intelligence*, 1996, vol. 23, no. 2, pp. 105–132. DOI: 10.1016/S0160-2896(96)90008-2.
23. Desimone J.A., Harms P.D., Desimone A.J. Best practice recommendations for data screening. *Journal of Organizational Behavior*, 2015, vol. 36, no. 2, pp. 171–181. DOI: 10.1002/job.1962.
24. Detterman D.K., Daniel M.H. Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence*, 1989, vol. 13, no. 4, pp. 349–359. DOI: 10.1016/S0160-2896(89)80007-8.
25. Drasgow F. Intelligence and workplace. In I.B. Weiner, N.W. Schmitt & S. Highhouse (Eds.), *Industrial and Organizational Psychology. Handbook of Psychology*, vol. 12 (2nd ed.). New York: Wiley, 2012, pp. 184–210.
26. Evans M.G. Implications of the asymmetry of g for predictive validity. *Poster presented at the APA/Yale Conference on Intelligence, June 2000*. Available at: [https://www.researchgate.net/profile/Martin\\_Evans/publication/267260638](https://www.researchgate.net/profile/Martin_Evans/publication/267260638) (accessed 26.11.2017).
27. Facon B. Are correlations between cognitive abilities highest in low-IQ groups during childhood? *Intelligence*, 2004, vol. 32, no. 4, pp. 391–401. DOI: 10.1016/j.intell.2004.06.002.
28. Fogarty G.J., Stankov L. Challenging the “Law of diminishing returns”. *Intelligence*, 1995, vol. 21, no. 2, pp. 157–174. DOI: 10.1016/0160-2896(95)90024-1.
29. Garrett H.E. A developmental theory of intelligence. *American Psychologist*, 1946, vol. 1, no. 9, pp. 372–378. DOI: 10.1037/h0056380.
30. Gottfredson L.S. Where and why g matters: Not a mystery. *Human Performance*, 2002, vol. 15, no. 1/2, pp. 25–46. DOI: 10.1080/08959285.2002.9668082.
31. Grigoriev A., Ushakov D., Valueva E., Zirenko M., Lynn R. Differences in educational attainment, socio-economic variables and geographical location across 79 provinces of the Russian Federation. *Intelligence*, 2016, vol. 58, pp. 14–17. DOI: 10.1016/j.intell.2016.05.008.

32. Hartmann P., Reuter M. Spearman's "Law of diminishing returns" tested with two methods // *Intelligence*, 2006, vol. 34, no. 1, pp. 47–62. DOI:10.1016/j.intell.2005.06.002.
33. Havlicek L., Peterson N. Effect of the violation of assumptions upon significance levels of the Pearson r. *Psychological Bulletin*, 1977, vol. 84, no. 2, pp. 373–377. DOI: 10.1037/0033-2909.84.2.373.
34. Jensen A.R. The g Factor: The Science of Mental Ability. Westport, CT: Praeger, 1998. 648 p.
35. Jensen A.R. Regularities in Spearman's law of diminishing returns. *Intelligence*, 2003, vol. 31, no. 2, pp. 95–105. DOI: 10.1016/S0160-2896(01)00094-0.
36. Johnson W., Bouchard T.J., Krueger R.F., McGue M., Gottesman I.I. Just one g: consistent results from three test batteries. *Intelligence*, 2004, vol. 32, no. 1, pp. 95–107. DOI: 10.1016/S0160-2896(03)00062-X.
37. Kaiser H.F. A measure of the average intercorrelation. *Educational and Psychological Measurement*, 1968, vol. 28, no. 2, pp. 245–247.
38. Kell H.J., Lang J.W.B. Specific abilities in the workplace: More important than g? *Journal of Intelligence*, 2017, vol. 5, no. 2, 13. DOI: 10.3390/jintelligence5020013.
39. Kroczek B., Ociepa M., Chuderski A. No Spearman's law of diminishing returns for the working memory and intelligence relationship. *Polish Psychological Bulletin*, 2016, vol. 47, no. 1, pp. 73–80. DOI: 10.1515/ppb-2016-0008.
40. Kuncel N.R., Hezlett S.A. Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science*, 2010, vol. 19, no. 6, pp. 339–345. DOI: 10.1177/0963721410389459.
41. Kyllonen P.C., Zu J. Use of response time for measuring cognitive ability. *Journal of Intelligence*, 2016, vol. 4, no. 4, 14. DOI: 10.3390/jintelligence4040014.
42. Lang J.W.B., Kersting M., Hülshager U.R., Lang J. General mental ability, narrower cognitive abilities, and job performance: The perspective of the nested-factors model of cognitive abilities. *Personnel Psychology*, 2010, vol. 63, no. 3, pp. 595–640. DOI: 10.1111/j.1744-6570.2010.01182.x.
43. Le H., Oh I.-S., Shaffer J., Schmidt F. Implications of methodological advances for the practice of personnel selection: How practitioners benefit from meta-analysis. *Academy of Management: Perspectives*, 2007, vol. 21, no. 3, pp. 6–15. DOI: 10.5465/AMP.2007.26421233.
44. LeBel E.P., Campbell L., Loving T.J. Benefits of open and high-powered research outweigh costs. *Journal of Personality and Social Psychology*, 2017, vol. 113, no. 2, pp. 230–243. DOI: 10.1037/pspi0000049.
45. Legree P.J., Pifer M.E., Grafton F.C. Correlations among cognitive abilities are lower for higher ability groups. *Intelligence*, 1996, vol. 23, no. 1, pp. 45–57. DOI: 10.1016/S0160-2896(96)80005-5.
46. Lilienfeld S.O. Psychology's replication crisis and the grant culture: Righting the ship. *Perspectives on Psychological Science*, 2017, vol. 12, no. 4, pp. 660–664. DOI: 10.1177/174569161668877.
47. Lynn R. Does Spearman's g decline at high IQ levels? Some evidence from Scotland. *The Journal of Genetic Psychology*, 1992, vol. 153, no. 2, pp. 229–230. DOI: 10.1080/00221325.1992.10753717.
48. Lynn R., Vanhanen T. *IQ and the Wealth of Nations*. Westport, CT: Praeger Publishers, 2002. 320 p.
49. Lynn R., Vanhanen T. National IQs: A review of their educational, cognitive, economic, political, demographic, sociological, epidemiological, geographic and climatic correlates. *Intelligence*, 2012, vol. 40, no. 2, pp. 226–234. DOI: 10.1016/j.intell.2011.11.004.
50. Martins A.A., Alves A.F., Almeida L.S. The factorial structure of cognitive abilities in childhood. *European Journal of Education and Psychology*, 2016, vol. 9, no. 1, pp. 38–45. DOI: 10.1016/j.ejeps.2015.11.003.
51. McGill R.J. Spearman's law of diminishing returns (SLODR): Examining effects at the level of prediction. *Journal of Psychology and Behavioral Science*, 2015, vol. 3, no. 1, pp. 24–36. DOI: 10.15640/jpbs.v3n1a3.
52. Meisenberg G., Lynn R. Intelligence: A measure of human capital in nations. *The Journal of Social, Political and Economic Studies*, 2011, vol. 36, no. 4, pp. 421–454.
53. Molenaar D., Kő N., Rózsa S., Mészáros A. Differentiation of cognitive abilities in the WAIS-IV at the item level. *Intelligence*, 2017, vol. 65, pp. 48–59. DOI: 10.1016/j.intell.2017.10.004.

54. Mount M.K., Oh I.S., Burns M. Incremental validity of perceptual speed and accuracy over general mental ability. *Personnel Psychology*, 2008, vol. 61, no. 1, pp. 113–139. DOI: 10.1111/j.1744-6570.2008.00107.x.
55. Murray A.L., Dixon H., Johnson W. Spearman's law of diminishing returns: A statistical artifact? *Intelligence*, 2013, vol. 41, no. 5, pp. 439–451. DOI: 10.1016/j.intell.2013.06.007
56. Ones D.S., Viswesvaran C., Schmidt F.L. Realizing the full potential of psychometric meta-analysis for a cumulative science and practice of human resource management. *Human Resource Management Review*, 2017, vol. 27, no. 2, pp. 201–215. DOI: 10.1016/j.hrmr.2016.09.011.
57. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 2015, vol. 349, no. 6251. DOI: 10.1126/science.aac4716.
58. Ree M.J., Earles J.A. Intelligence is the best predictor of job performance. *Current Directions in Psychological Sciences*, 1992, vol. 1, no. 3, pp. 86–89.
59. Ree J.M., Earles J.A., Teachout M.S. Predicting job performance: Not much more than g. *Journal of Applied Psychology*, 1994, vol. 79, no. 4, pp. 518–524. DOI: 10.1037/0021-9010.79.4.518.
60. Ree M.J., Carretta T.R., Teachout M.S. Pervasiveness of dominant general factors in organizational measurement. *Industrial and Organizational Psychology*, 2015, vol. 8, no. 3, pp. 409–427. DOI: 10.1017/iop.2015.16.
61. Reeve C.L., Scherbaum C., Goldstein H. Manifestations of intelligence: Expanding the measurement space to reconsider specific cognitive abilities. *Human Resources Management Review*, 2015, vol. 25, no. 1, pp. 28–37. DOI: 10.1016/j.hrmr.2014.09.005.
62. Reynolds M.R., Keith T.Z. Spearman's law of diminishing returns in hierarchical models of intelligence for children and adolescents. *Intelligence*, 2007, vol. 35, no. 3, pp. 267–281. DOI: 10.1016/j.intell.2006.08.002.
63. Reynolds M.R., Keith T.Z., Beretvas S.N. Use of factor mixture modeling to capture Spearman's law of diminishing returns. *Intelligence*, 2010, vol. 38, no. 2, pp. 231–241. DOI: 10.1016/j.intell.2010.01.002.
64. Salgado J.F., Anderson N., Moscoso S., Bertua C., de Fruyt F. International validity generalization of GMA and cognitive abilities: A European Community meta-analysis. *Personnel Psychology*, 2003, vol. 56, no. 3, pp. 573–605. DOI: 10.1111/j.1744-6570.2003.tb00751.x.
65. Salthouse T.A. Evaluating the correspondence of different cognitive batteries. *Assessment*, 2014, vol. 21, no. 2, pp. 131–142. DOI: 10.1177/1073191113486690.
66. Schmidt F.L., Hunter J. General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 2004, vol. 86, no. 1, pp. 162–173. DOI: 10.1037/0022-3514.86.1.162.
67. Schmidt F.L., Oh I.-S., Shaffer J.A. The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 100 years of research findings. *Working paper*. October 2016. Available at: <http://www.researchgate.net/publication/309203898> (accessed 24.02.2017). DOI: 10.13140/RG.2.2.18843.26400.
68. Spearman C.E. *The Abilities of Man: Their Nature and Measurement*. New York: Macmillan, 1927. 415 p.
69. Stanhope D.S., Surface E.A. Examining the incremental validity and relative importance of specific cognitive abilities in a training context. *Journal of Personnel Psychology*, 2014, vol. 13, no. 3, pp. 146–156. DOI: 10.1027/1866-5888/a000116.
70. Tucker-Drob E. Differentiation of cognitive abilities across the lifespan. *Developmental Psychology*, 2009, vol. 45, no. 4, pp. 1097–1118. DOI: 10.1037/a0015864.
71. Tuğran E., Kocak M., Mirtağoğlu H., Yiğit, S., Mendes M. A simulation based comparison of correlation coefficients with regard to type I error rate and power. *Journal of Data Analysis and Information Processing*, 2015, vol. 3, no. 3, pp. 87–101. DOI: 10.4236/jdaip.2015.33010.
72. Ushakov D.V., Grigoriev A.A. Macropsychology of intelligence: Through emotions to theoretical depth. *Psychology: Journal of the Higher School of Economics*, 2016, vol. 13, no. 4, pp. 629–635.
73. Van der Linden W.J. Setting time limits on tests. *Applied Psychological Measurement*, 2011, vol. 35, no. 3, pp. 183–199. DOI: 10.1177/0146621610391648.
74. Woodley M.A. The cognitive differentiation-integration effort hypothesis: A synthesis between the fitness indicator and life history models of human intelligence. *Review of General Psychology*, 2011, vol. 15, no. 3, pp. 228–245. DOI: 10.1037/a0024348.



Сугоняев Константин Владимирович, ассоциированный сотрудник, Институт психологии Российской академии наук (Москва), skv-354@yandex.ru

Радченко Юрий Иванович, начальник Научно-практического центра Военной академии Генерального штаба Вооруженных Сил Российской Федерации (Москва), savaur@mail.ru

Поступила в редакцию 2 февраля 2018 г.

DOI: 10.14529/psy180101

## SPEARMAN'S LAW OF DIMINISHING RETURNS: INVESTIGATION ON LARGE-SCALES RUSSIAN SAMPLES

K.V. Sugonyaev<sup>1</sup>, skv-354@yandex.ru

Yu.I. Radchenko<sup>2</sup>, savaur@mail.ru

<sup>1</sup> Institute of Psychology of Russian Academy of Sciences, Moscow, Russian Federation

<sup>2</sup> Military Academy of General Staff, Moscow, Russian Federation

The reduction of average correlations among cognitive tests in groups of individuals of above-average IQ was invented by C. Spearman (1927) and now is known as Spearman's "Law of Diminishing Returns" (SLODR). Although many studies also founded that inter-correlations between subtests decreased from low to high ability groups, several empirical investigations have challenged the SLODR. In Russia similar studies didn't carry out by now. We investigated the presence of SLODR in large-scaled dataset consisted of sample of 11335 applicants for joining to military institutes (mean age  $18.42 \pm 1.29$  years). The subjects were tested with the same battery of 10 ability tests. SLODR was assessed (tested) by dividing the sample into low/high ability groups based on the total score. In order to account for criticism of Murray et. al. (2013), test scores in each half-sample were transformed to normalized stens separately to equate variability and minimize skewness. Our findings were fully consistent with SLODR-based predictions, thus refuting suggestions about artifactual genesis of this phenomenon. To cross-validate this result analysis was repeated on second analogous sample after exclusion of all applicants older than 19 years ( $N = 11934$ , mean age =  $17.91 \pm 0.47$ ). And again, all indexes accounted were consistent with SLODR. The results indicate that SLODR effects are emerged on the level of group ability factors, are quite reproducible and relatively independent from test's g-loading and psychometric properties of their scores. Hypotheses are suggested to account for possible reasons of negative outcomes in some earlier studies.

*Keywords:* Spearman's "Law of Diminishing Returns" (SLODR), the differentiation hypothesis, general cognitive ability, g factor, psychometric characteristics.

The article was prepared with the support of the Russian Science Foundation, project No. 17-78-30035 "Psychological factors of economic and social competitiveness of Russia".

Received 2 February 2018

### ОБРАЗЕЦ ЦИТИРОВАНИЯ

Сугоняев, К.В. «Закон уменьшения отдачи» Спирмена: исследование на масштабных российских выборках / К.В. Сугоняев, Ю.И. Радченко // Вестник ЮУрГУ. Серия «Психология». – 2018. – Т. 11, № 1. – С. 5–21. DOI: 10.14529/psy180101

### FOR CITATION

Sugonyaev K.V., Radchenko Yu.I. Spearman's Law of Diminishing Returns: Investigation on Large-Scales Russian Samples. *Bulletin of the South Ural State University. Ser. Psychology*. 2018, vol. 11, no. 1, pp. 5–21. (in Russ.). DOI: 10.14529/psy180101