

DOI: 10.22363/2313-1683-2024-21-2-657-682

EDN: MCKAAY

УДК 159.9.072

Теоретическая статья

Этические аспекты измерения интеллекта: навстречу компетентности и справедливости

Т.И. Логвиненко¹, Т.Н. Канонир², Е.А. Орел², А.А. Куликова²

¹Научно-технологический университет «Сириус»,
Российская Федерация

²Национальный исследовательский университет «Высшая школа экономики»,
Москва, Российская Федерация
 logvinenkota.spb@gmail.com

Аннотация. Настоящая статья посвящена проблеме измерения интеллекта с акцентом на этический аспект разработки и применения тестов. История психодиагностики интеллекта дает разнообразные примеры, проблемные с точки зрения этики, которые не раз приводили к негативным последствиям как на уровне отдельных индивидов, так и для целых сообществ. Цель данной статьи – описать актуальные этические проблемы в сфере измерения интеллекта, их предпосылки и исторические примеры. Мы обсуждаем этические проблемы, касающиеся: 1) глобальных подходов к операционализации интеллекта; 2) возможных нарушений прав человека в результате применения тестов интеллектов; 3) справедливости тестов интеллекта для различных групп респондентов; 4) оценки качества и выбора тестового инструментария. Данные вопросы мы рассматриваем через призму этических принципов психолога, таких как уважение, честность, компетентность и ответственность. Несмотря на богатую историю измерения интеллекта и исследований в этой области, этические проблемы, поднятые десятилетия назад, не потеряли свою актуальность. Так как этические вопросы часто не имеют однозначно верного ответа, важным является само существование такой дискуссии и обсуждение истоков этических проблем тестирования интеллекта, широкого контекста, в которых они существуют, и предлагаемых сообществом решений. Материалы и выводы данной статьи могут быть полезны как исследователям, так и практикам для принятия информированных решений в контексте измерения интеллекта.

Ключевые слова: интеллект, IQ, тесты интеллекта, психодиагностика, психологическая оценка, этические принципы, этика психолога

Благодарности и финансирование: Исследование выполнено за счет гранта Российского научного фонда (проект № 23-28-01399).

© Логвиненко Т.И., Канонир Т.Н., Орел Е.А., Куликова А.А., 2024



This work is licensed under a Creative Commons Attribution 4.0 International License

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

Введение

Использование тестов для психологической диагностики имеет длительную историю, богатую как на открытия и инновации, так и на ошибки и нарушения прав участников тестирования. В ходе развития психологии как научной и практической области формировались этические принципы профессиональной деятельности. Часто они становились результатом переосмысления методов работы психологов, иногда – после трагических прецедентов. Сегодня профессиональная и исследовательская деятельность психолога регулируется различными этическими кодексами. В России наиболее распространен этический кодекс Российского психологического общества¹, в мире, например, Мета-кодекс этики Европейской федерации ассоциаций психологов² и Этические принципы психологов и кодекс поведения Американской психологической ассоциации³. Кодексы обязательно содержат раздел, посвященный использованию психодиагностических инструментов. Однако кодексы дают общие направления деятельности, но не могут и не должны описывать все возможные этические дилеммы, возникающие в процессе работы психолога. Поэтому важно изучать опыт и ошибки прошлого, чтобы научиться распознавать потенциальные этические дилеммы в ежедневной практике и решать их в соответствии с профессиональными принципами.

Вышесказанное в полной мере относится к тестированию интеллекта, одной из главных и старейших областей психологической диагностики. Для введения в контекст стоит кратко определить интеллект как способность к рассуждению и решению проблем (Gottfredson, 1997; Neisser et al., 1996). Важность измерения интеллекта, как правило, обосновывается тем, что интеллект является предиктором академических, профессиональных и социальных достижений во взрослом возрасте (Deary, 2012; Deary et al., 2010; Sternberg et al., 2001; Strenze, 2007; Zaboski II et al., 2018). Более того, ученые указывают, что интеллект предсказывает эти достижения лучше, чем какие-либо другие черты и способности (Plomin & von Stumm, 2018).

Абсолютное большинство инструментов, используемых для оценки интеллекта (например, тест Стэнфорд-Бине, тест Векслера), базируются на психометрических теориях, которые возникли на основе факторного анализа результатов уже существующих тестов (об истории тестов интеллекта и способностей см.: Анастаси, Урбина, 2007). Большое количество разработанных тестов помогло психометрическим теориям приобрести солидное эмпирическое обоснование, что и обеспечило популярность этого подхода среди исследователей и практиков, которые добавляют все новые и новые данные. Длительная история тестирования интеллекта и развития психометрических теорий дает нам возможность рассмотреть множество этических дилемм

¹ Российское психологическое общество. *Этический кодекс психолога*. <http://psyurus.ru/rpo/documentation/ethics.php> (дата обращения: 30.01.2024).

² European Federation of Psychologists' Associations. *Meta-Code of Ethics*. 2005. <https://www.efpa.eu/sites/default/files/2023-04/meta-code-of-ethics.pdf> (accessed: 30.01.2024).

³ American Psychological Association. *Ethical Principles of Psychologists and Code of Conduct*. 2017. <https://www.apa.org/ethics/code> (accessed: 30.01.2024).

и потенциальных нарушений прав участников тестирования. Анализ этих кейсов актуален и для современной психодиагностики.

Цель данной статьи – рассмотреть этические дилеммы, возникавшие в ходе длительной истории измерения интеллекта и имеющие актуальность на сегодняшний день. Мы обозначим проблемные вопросы, где возможно – предлагаемые пути их решения, и рассмотрим их в контексте сегодняшних представлений об этической профессиональной деятельности психолога. Международное психологическое сообщество выделяет следующие основные этические принципы⁴: 1) уважение прав, свободы и достоинства личности (конфиденциальность, добровольное участие, осведомленность, право на самоопределение, и непосредственно уважение); 2) компетентность (профессиональные знания специалиста, осознание границ своей компетентности, знание профессиональной этики, профессиональное развитие, невозможность действий в определенных условиях); 3) ответственность (непричинение вреда, решение этических дилемм, ответственность по отношению к клиентам, профессиональному сообществу и обществу в целом); 4) честность (справедливость, последовательность, уважение к коллегам, избегание конфликта интересов, ответственность и открытость перед профессиональным сообществом). Мы будем обращаться к этим принципам при описании этических проблем.

В первом разделе статьи опишем предпосылки, предусмотренные в самой концепции интеллекта и истории ее развития и измерения IQ, которые лежат в основе многих частных этических и методологических вопросов. Далее через призму четырех этических принципов рассмотрим несколько конкретных кейсов, как исторических, так и современных, иллюстрирующих нетривиальные этические вопросы, которые стоят перед исследователями и практиками.

Что измеряют тесты интеллекта?

Современные представления об интеллекте и, соответственно, его измерении формировались в контексте западной науки и главенствующих в то или иное время концепций. Так, например, большое влияние на развитие представлений об интеллекте оказала евгеника, отголоски которой мы будем находить в истории измерения интеллекта на протяжении всего двадцатого века.

В начале двадцатого века в Старом Свете наука была более ориентирована на теорию и выявление общих закономерностей функционирования человеческой психики (традиция, заданная В. Вундтом и продолженная, например, в гештальт-психологии). В это же время основным двигателем развития инструментов для психологической диагностики становится США, где традиционно был силен уклон на прикладное использование результатов тестирования. Этот подход затем распространился на весь западный мир: опросы и тестирования ориентировались на немедленное прикладное использование, например, в целях отбора в образовании и в профессиональной деятельности. Определенно тесты, их содержание начали отражать запросы

⁴ European Federation of Psychologists' Associations. Meta-Code of Ethics. 2005. <https://www.efpa.eu/sites/default/files/2023-04/meta-code-of-ethics.pdf> (accessed: 30.01.2024).

общества на измерения. Такая ситуация в какой-то момент привела к монополии на представления о том, кого считать достаточно интеллектуальным, или, другими словами, интеллект какого типа будет востребован в обществе. Востребованным в обществе был интеллект, понимаемый как способность решать когнитивные задания и ориентированный на набор знаний, преподаваемых в учебных заведениях. Монополия такого «аналитического интеллекта» (близкого к академическим достижениям) была нарушена только в начале 1980-х годов, когда в свет сначала вышла книга Г. Гарднера «Структура разума: теория множественного интеллекта» (Gardner, 1983), а следом Р. Стернберг презентовал свою триархическую теорию интеллекта (Sternberg, 1985). Параллельно, начиная с 1920-х годов развивались идеи социального интеллекта – «мудрого поведения в человеческих отношениях», как определил его Э. Торндайк (Thorndike, 1920). Эти идеи к середине – концу 1980-х приобрели большую популярность в психологическом сообществе и продолжили развиваться в начале 1990-х, когда благодаря бестселлеру Д. Гоулмана (Goleman, 1995) стал набирать свою популярность интеллект эмоциональный. Можно резюмировать, что в истории изучения и измерения интеллекта со временем происходит переосмысление того, какие характеристики человека являются значимыми для процессов, происходящих в обществе (Герасимова & Орел, 2022).

Критики утверждают, что современные популярные тесты IQ охватывают не весь спектр человеческого интеллекта, часто упуская из вида такие аспекты, как креативность, эмоциональный интеллект, социальные навыки и практическую мудрость (Холодная, 2004а). Так, в 2004 году на страницах журнала «Психология. Журнал ВШЭ» состоялась знаковая дискуссия, посвященная вопросам тестирования интеллекта и роли результатов тестирования для принятия решений. Часть специалистов (Ушаков, 2004; Шмелев, 2004) показывали вклад интеллекта в различные достижения человека – академические или профессиональные. Другие же (Холодная, 2004b) призывали с осторожностью использовать тесты интеллекта, так как в руках непрофессионалов они могут навредить интересам личности.

К. Станович (Stanovich, 2009) соглашается с тем, что существует много типов поведения, которое можно охарактеризовать как интеллектуальное, но разделяет со многими психологами обеспокоенность тем, что применение слова «интеллект» к широкому спектру различных видов поведения усугубляет существующую путаницу. Он отмечает, что некоторые люди, несмотря на наличие находящегося в рамках нормы интеллекта, не могут мыслить или действовать рационально. Для описания этого феномена он использует термин «дизрационалия» (*dysrationalia*): к этому же феномену он относит примеры, когда выдающиеся и успешные в своей профессии люди принимают иррациональные, а порой даже кажущиеся глупыми личные решения. Один из выводов К. Становича состоит в том, что традиционные тесты IQ не могут в полной мере оценить то, что они предназначены оценивать – способность к рассуждению и решению проблем (Stanovich et al., 2016). Эти сообщения необходимо иметь в виду и когда мы предполагаем, что коэффици-

ент IQ ничего не значит в реальном мире, и когда мы рассчитываем, что высокий IQ гарантирует рациональное поведение и личный и профессиональный успех.

Дополнительно встает вопрос, что понимается под успешностью, которую предсказывают традиционные тесты интеллекта. Исследования показывают, что интеллект, измеренный с помощью традиционных тестов IQ, является довольно точным предиктором академических и профессиональных достижений (Roth et al., 2015). Однако существуют и другие характеристики успеха, такие как индивидуальное чувство самореализации, благополучие. Они в том числе зависят от личных качеств и мотивации, поэтому прогностическую достоверность тестов интеллекта не следует переоценивать и рассматривать как окончательный показатель будущего успеха (Müska, 2014). Более того, в своей недавней статье Р. Стернберг, Д. Прейсс и С. Караме (Sternberg et al., 2023) заявляют, что господствующая сегодня концепция интеллекта в значительной степени отражает не современные культурные потребности, а скорее потребности Европы и Америки начала XX века, когда эта концепция формировалась. Авторы утверждают, что определение того, что мы понимаем под интеллектом, должно рассматриваться в рамках соответствующих исторических контекстов и может и должно меняться вместе с преобразованиями в обществе.

Таким образом, доминирующее понимание интеллекта как аналитических способностей, развиваемых в рамках системы образования (не во всех странах доступного для каждого ребенка), в последние 30 лет все сильнее уступает свои лидирующие позиции более широким взглядам на интеллект и критерии социальной успешности в целом. Образование к тому же отражает общественный консенсус о том, чему нужно учить для успеха в этом конкретном сообществе, а значит индикаторы успеха и лежащие в их основе знания и навыки могут различаться от культуры к культуре.

Однако несмотря на критику психометрических теорий, на сегодняшний день пока еще абсолютное большинство тестов, используемых для психодиагностики интеллекта, разработаны в рамках этого подхода. Популярность этих тестов основывается на большом количестве доступных эмпирических данных, что помогает практикам принимать обоснованные решения на основе полученных результатов (хотя и тут ведется научная дискуссия о том, какие именно данные и в какой форме правомерно использовать при интерпретации – см., например, Canivez et al., 2020). Вместе с тем нам кажется, что накопившиеся противоречия и новые общественные запросы подталкивают к парадигмальному сдвигу в глобальных подходах к пониманию и измерению интеллекта. Однако в любом случае исследователям и практикам важно ясно понимать теоретические основания, на которых построены и доминирующие, и альтернативные и еще только разрабатываемые тесты интеллекта, так как теоретическая основа напрямую связана с интерпретацией результатов.

Как в результате оценки интеллекта могут нарушаться права различных групп респондентов?

А. Бине, основоположник современного подхода к измерению интеллекта, утверждал, что: 1) тест должен применяться только для тех целей и той целевой группы, для которой он создается, и 2) тест не должен быть инструментом стигматизации тех, кому не удалось с ним справиться (Gould, 1981). Однако именно это и произошло с тестированием интеллекта с течением времени. Применение результатов тестирования интеллекта для принятия решений в различных сферах на протяжении всей истории измерений IQ часто становились предметом споров и разбирательств как в научных кругах, так и в средствах массовой информации (The IQ controversy..., 1976; Kamin & Egerton, 1973). Задача данного раздела – попытаться обобщить опыт подобных конфликтов и выявить, какие факторы важно учитывать при принятии тех или иных решений на основе результатов тестирования интеллекта.

Первая группа кейсов касается использования показателей IQ как фактора дискриминации меньшинств. Интерпретация результатов некоторых тестов на интеллект на многие годы повлияла на отношение к интеллектуальным способностям людей, не принадлежащих к среднему классу «белой» Америки. Таким важным тестом бесспорно стал тест Army Alpha, разработанный под руководством Р. Йеркса (R.M. Yerkes) в 1917 году и применяемый для тестирования и распределения рекрутов в американской армии. Этим тестом было протестировано почти два миллиона человек (Gould, 1981). Разрабатываемый и интерпретируемый под влиянием евгеники, тест давал преимущество в результатах образованным респондентам, знакомым с бытом и практиками, присущими среднему классу. Очень быстро выяснилось, что тест не подходит для тестирования людей, не владеющих грамотой, и была разработана невербальная версия теста Army Beta. Как будто появление «адаптированной» версии может говорить о том, что разработчики учли то обстоятельство, что тест должен учитывать знакомство потенциальных респондентов с контекстом заданий. Однако невербальные задания также содержали в себе искажения, так как изображали объекты и действия, не знакомые, например, многим чернокожим рекрутам. Естественно, полученные результаты показывали «интеллектуальное преимущество» белых американцев и очень низкий уровень интеллекта у чернокожих. Учитывая масштабность тестирования и влияние вовлеченных в его организацию людей, именно результаты этого теста в конце концов легли в основу принятия закона «Об ограничении иммиграции» 1924 года, зафиксировавшего квоты на въезд для жителей Европы и запрещавшего иммиграцию жителям Азиатского региона. В законе предусматривалось обязательное тестирование всех въезжающих в США. Поначалу это тестирование оценивало не столько интеллект вновь прибывших иммигрантов, сколько их знание английского языка. Позже были разработаны невербальные тесты, и экзамены стали проводиться на родном языке иммигранта, однако тем, кто не мог сдать тест, могло быть отказано во въезде. Стоит отметить, правда, что процент отказов на основании именно результатов тестирования был очень низким (Warne, 2020 на основе данных Н. Unrau, 1984). Тем не менее этот пример хорошо

иллюстрирует, как легко тестирование интеллекта становится механизмом сегрегации и стигматизации различных групп населения.

Реакцией на истории о сегрегации и разделения общества на основе тестирования в некоторых случаях становился полный запрет на любые тестирования умственных способностей. Для примера достаточно вспомнить постановление ЦК ВКП(б) от 4 июля 1936 года «О педологических извращениях в системе Наркомпросов»⁵ (ЦК ВКП(б), 1985), поставившим крест на развитии отечественной психодиагностики на долгие годы. В качестве обоснования запрета на проведение обследований умственных способностей учащихся авторы постановления указывают на следующие обстоятельства:

«Практика педологов, протекавшая в полном отрыве от педагога и школьных занятий, свелась в основном к ложно-научным экспериментам и проведению среди школьников и их родителей бесчисленного количества обследований в виде бессмысленных и вредных анкет, тестов и т. п., давно осужденных партией. Эти, якобы, научные «обследования», проводимые среди большого количества учащихся и их родителей, направлялись, по преимуществу, против неуспевающих или неукладывающихся в рамки школьного режима школьников и имели своей целью доказать, якобы, с «научной» «биосоциальной» точки зрения современной педологии наследственную и социальную обусловленность неуспеваемости ученика или отдельных дефектов его поведения, найти максимум отрицательных влияний и патологических извращений самого школьника, его семьи, родных, предков, общественной среды и тем самым найти повод для удаления школьников из нормального школьного коллектива.

В этих же целях действовала обширная система обследований умственного развития и одаренности школьников, некритически перенесенная на советскую почву из буржуазной классовой педологии и представляющая из себя форменное издевательство над учащимися, противоречащая задачам советской школы и здравому смыслу. Ребенку 6–7 лет задавались стандартные казуистические вопросы, после чего определялся его так называемый «педологический» возраст и степень его умственной одаренности.

Все это вело к тому, что все большее и большее количество детей зачислялось в категории умственно отсталых, дефективных и «трудных».

На основании отнесения подвергшихся педологическому «изучению» школьников к одной из указанных категорий педологи определяли подлежащих удалению из нормальной школы детей в «специальные» школы и классы для детей «трудных», умственно отсталых, психо-невротиков и т. д.»

Как видно из приведенной цитаты, основной обозначенной причиной запрета деятельности педологов являлась сегрегация детей на основе результатов измерения их умственных способностей. В отсутствии стандартов

⁵ Постановление ЦК ВКП(б) «О педологических извращениях в системе наркомпросов». 4 июля 1936 г. // Коммунистическая партия Советского Союза в резолюциях и решениях съездов, конференций и пленумов ЦК (1898–1988) / под общ. ред. А.Г. Егорова, К.М. Боголюбова. В 16 т. Т. 6. 1933–1937 гг. М.: Ин-т марксизма-ленинизма при ЦК КПСС, 1985. С. 364–367.

разработки и администрирования тестов у сообщества педологов не было аргументов в споре с советской властью, решившей не работать с культурой оценки умственных способностей, чтобы использовать ее результаты во благо обществу, а просто запретить всяческие измерения, чтобы убрать основания для разделения учащихся по уровню способностей.

Пример Советского Союза не является единственным. В Соединенных Штатах подобные кейсы неправомерного использования тестов IQ в конечном итоге вылились в серию судебных исков, повлекших ограничения в использовании тестов. В одном из наиболее известных таких разбирательств – «Ларри П. против Райлза» – федеральный суд США запретил использовать тесты на интеллект для отнесения учащихся к категории «умственно отсталых» и последующего направления их в специальные классы. Аргументация суда заключалась в том, что тесты носили незаконный дискриминационный характер, поскольку при их использовании непропорционально большой процент афроамериканских детей получал результаты IQ ниже 70 баллов, что является одним из критериев умственной отсталости. Беспокойство и общественная критика имели веские основания, однако при принятии итоговых решений валидность результатов самого тестирования не была полноценно оценена и потому, как минимум часть детей впоследствии не получили необходимой образовательной поддержки. Несмотря на эти неоднозначные результаты, запрет действовал еще в течение нескольких десятилетий (Sattler, 2008; Powers et al., 2004).

В приведенных примерах наблюдаются нарушения принципов уважения и честности, которые в дальнейшем вызывали контрреакции, зачастую тоже с неоднозначными последствиями в контексте общественного блага. Оценка интеллекта — это чувствительный вопрос для общества, так как она имеет значительное влияние на судьбу людей. Мы полагаем, что реализация принципов уважения и честности возможна при выполнении следующих условий: во-первых, четкого определения целей тестирования и регулярного мониторинга соответствия его результатов этим целям; во-вторых, максимальной нейтральности по отношению к различным группам людей и обеспечении справедливой оценки. Профессиональному сообществу необходимо постоянно отслеживать, на каких предпосылках основываются измерения и выводы, чтобы в будущем своевременно предотвращать и выявлять возможные этические нарушения.

Вслед за этим встает вопрос о том, при каких условиях оценивание может быть справедливым и какую роль в спорных последствиях тестирования играет сам тест? Чтобы ответить на этот вопрос, обратимся к проблеме необъективности тестов.

Справедливы ли тесты интеллекта для различных групп?

Есть две крайние точки зрения: первая, что тесты интеллекта в своей массе позволяют объективно оценить способности, и противоположная, что многие тесты способностей вообще и тесты интеллекта в частности несправедливы по отношению к некоторым группам респондентов. Началом подобной критики стали наблюдения, что этнические меньшинства, отдельные

расы, малоимущие и другие группы систематически получают при тестировании баллы ниже, чем доминирующие или привилегированные группы. Критики объясняют это свойствами тестов и несовершенными подходами к оцениванию. Опасения связаны, в частности, с тем, что людям из группы меньшинств могут чаще ставить чрезмерный (ложноположительный) психиатрический диагноз, учащиеся непропорционально часто попадают в специальные классы, а кандидатам отказывают в приеме на работу или в образовательное учреждение из-за несправедливых стандартизированных тестов. В этом контексте кейс «Ларри П. против Райлза», описанный выше, был знаковой вехой этих дебатов.

Чтобы описать пространство между этими крайними точками зрения, стоит ввести два важных понятия – справедливость теста (*test fairness*) и предвзятость теста (*test bias*). В литературе можно встретить разные позиции: 1) справедливость и непредвзятость являются тождественными понятиями; 2) непредвзятость является одним из необходимых условий справедливости теста; 3) предвзятость теста не исключает справедливость, и наоборот.

Одними из первых разделять понятия (не)предвзятости и справедливости предложили Дж. Хантер и Ф. Шмидт (Hunter & Schmidt, 1976), делая акцент на том, что (не)предвзятость относится к техническим и статистическим свойствам теста, которые могут быть оценены; а справедливость теста является социальным и этическим концептом, прежде всего связанным с вопросом принятия решений на основе тестов и их социальными последствиями. Вопрос справедливости, таким образом, авторы тесно связывают с философскими концепциями отбора и приводят описания возможных позиций: например, позиции неквалифицированного индивидуализма, общественных квот, квалифицированного индивидуализма. С точки зрения, которую представляют Дж. Хантер и Ф. Шмидт, справедливость и предвзятость хоть и связаны, но не являются взаимоисключающими, и вопрос справедливого использования любого теста выходит из плоскости объективного и становится вопросом морали и ценностей отдельного человека, организаций и сообществ, принимающих решения.

Иную позицию можно встретить в стандартах, совместно разработанных Американской ассоциацией исследований в области образования, Американской психологической ассоциацией и Национальным советом по измерениям в образовании (American Educational Research Association et al., 1999, 2014), являющихся влиятельными игроками в области тестирования в образовании и психологии. В последней версии стандартов тема «Справедливость тестирования» была вынесена в отдельную секцию-книгу (*Fairness in Educational...*, 2022), наряду с традиционными секциями по валидности и надежности тестов. Стандарты определяют четыре основания справедливого тестирования: 1) сопоставимость интерпретации результатов для разных испытуемых должна быть приоритетнее стандартизации процедур тестирования; 2) отсутствие предвзятости измерения (*measurement bias*) при разработке теста; 3) доступность измеряемых конструктов для всех тестируемых при разработке теста и 4) интерпретация и использование тестовых баллов должны отражать взаимодействие индивидуальных особенностей тестируе-

мого с тестом и контекстом, в котором он проводится. То есть согласно этой точке зрения, непредвзятость (прежде всего – отсутствие предвзятости измерения) является одним из необходимых условий справедливости теста.

Так или иначе, несмотря на некоторую путаницу между понятиями справедливости и предвзятости в неакадемической среде (Reynolds & Suzuki, 2012) и различий в определении «справедливости теста» среди ученых, в научном сообществе есть относительно консолидированное мнение, что (не)предвзятость теста относится к статистическим показателям. Предвзятость в этом смысле означает неодинаковую валидность теста для различных подгрупп, когда индивиды с гипотетически одним уровнем латентной способности (например, одинаковым уровнем интеллекта) получают разные результаты по тесту. Предвзятый тест систематически завышает или занижает значение переменной, для оценки которой он предназначен. Если эта погрешность зависит от культурной переменной, такой как этническая принадлежность, то говорят о наличии культурной предвзятости теста, однако смещение может происходить и относительно других переменных – например, возраста и пола.

Очень долгое время среди пользователей тестов и исследователей было распространено мнение, что для использования теста на другом языке и в другой культуре не требуется каких-то изменений в тесте, кроме перевода. Более того, если тест предполагался для использования на том же языке, на котором он разработан, но для другой целевой популяции, или если это невербальный тест, никакие дополнительные процедуры не нужны. Представление о том, что тесты нуждаются в адаптации, которая не является просто переводом, хотя и существовали в научном сообществе долгое время, но официально и полно закрепились не столь давно, ознаменовавшись выходом Руководства по адаптации тестов от Международной комиссии по тестам в 2004 году (Bartram & Hambleton, 2016). Вся процедура адаптации как раз направлена на достижение эквивалентности и сопоставимости конструкторов, методов и отдельных заданий для всех целевых групп двух версий инструментов, результаты которых могут быть сравнимы. В случаях, когда эквивалентности по каким-либо причинам невозможно достичь, сравнения не проводятся. В Руководстве описаны способы оценки и преодоления предвзятости, которые в целом соответствует так называемой концепции предвзятости (Van de Vijver & Leung, 2021; Van de Vijver & Poortinga, 1997; Van de Vijver & Tanzer, 2004).

Предвзятость на уровне конструкторов может быть, если оцениваемый конструктор неодинаково выражается или должен быть определен по-разному у различных групп. В контексте оценки интеллекта вопрос может звучать так: «Является ли интеллект представителя одной культуры/расы/группы тем же самым конструктором, что и интеллект представителя другой культуры/расы/группы?».

Предвзятость на уровне методов может возникнуть из-за несопоставимости выборок (различия в образовании и доступе к образованию, социально-экономическом статусе, жители городов или сельской местности и др.) или, например, из-за несопоставимости условий оценивания (неодинаковая понят-

ность инструкций для разных респондентов, коммуникативные проблемы во взаимодействии с администратором тестирования и его установки и др.).

Предвзятость на уровне отдельных заданий, пожалуй, является самым очевидным и известным пунктом. Так, даже на заре создания тестов на интеллект при переводе учитывались как минимум нюансы географических названий, меры длины, валюты и так далее. Например, когда тест А. Бине был переведен на английский язык для использования в США, американским психологам было очевидно, что арифметическое задание, требующее знания французских денег, необходимо изменить, поскольку в США нет монет в $\frac{1}{2}$ цента или 2 цента (Terman, 1916). Однако сопоставимость отдельных утверждений не всегда настолько очевидна, и ее проверка требует применения специальных психометрических процедур.

В своем справочнике по психологическому тестированию Р. Грегори (Gregory, 2015) указывает, что в большинстве тестов способностей и достижений не обнаружена предвзятость, и ссылается на работы А. Дженсена (Jensen, 1980), С. Рейнольдса (Reynolds, 1994), Н. Кансела и П. Сакетта (Kuncel & Sackett, 2007) и других. Вместе с тем есть работа Й. Вихертс (Wicherts, 2016) с обзором исследований предвзятости батарей тестов интеллекта (в том числе самых популярных, например, WAIS-IV), где автор приходит к выводу, что в более чем половине статей не подтверждается, например, измерительная инвариантность (эквивалентность) методик в отношении этнических групп, пола, образования или возраста. Нет также и консенсуса и о конкретных способах измерения предвзятости (Putnick & Bornstein, 2016), а также о степени влияния разных аспектов предвзятости на валидность выводов о групповых различиях – например, см. полемику Б. Меулемана и соавторов (Meuleman et al., 2023) и К. Вельцеля и соавторов (Welzel et al., 2023).

Так или иначе, можно выделить как минимум три альтернативных направления развития тестов интеллекта, возникших в том числе как ответ на критику о предвзятости «классических» тестов. Одно направление включает развитие идеи невербальных культурно-неспецифических тестов. Примерами подобных комплексных и современных тестов могут быть Leiter-3 (Roid & Koch, 2017, доступен для использования в России) и UNIT II (Friedlander Moore et al., 2017; использовался в России в исследовательских целях). Оба теста разрабатывались в так называемой парадигме «универсального дизайна», являются полностью невербальными и были специально созданы для применения как в общих, так и особых популяциях: различных этнических группах, респондентов с ограниченным слухом, речью или другими нарушениями развития.

Вторым направлением можно обозначить то, что Дж. Наглиери назвал тестами второго поколения (Naglieri, 2015): прежде всего, Cognitive Assessment System (Naglieri et al., 1997) и Kaufman Assessment Battery for Children II (Bain & Jaspers, 2010; Kaufman & Kaufman, 1983). Особенностью этих тестов является добавление нейропсихологических классификаций и теорий в теоретическое основание теста и наличие полностью невербальных вариантов, что обеспечивает более справедливое оценивание самых разных детских популяций. Согласно авторам, эти тесты предсказывают академические до-

стижения не хуже классических, но при их применении различия между расами в сравнении с классическими тестами уменьшаются вдвое.

Третье направление можно обозначить как «культурно-специфическое» или «экологическое» (Berry, 1976, 2022). Эта точка зрения подчеркивает, что интеллект адаптируется к экологическим и культурным условиям, а потому развитие и проявление индивидуального интеллекта будут меняться по мере изменения этих условий – и этот контекст необходимо учитывать при оценке и создании тестов интеллекта. Примером реализации «культурно-специфического» подхода может быть тест «Panga Munthu», разработанный в Замбии для измерения интеллекта африканских детей (Kathura & Serpell, 1998), где исследователи просили вылепить фигурку человека из глины или проволоки вместо выполнения бумажных заданий. Еще одним примером может быть концепция «практического интеллекта» (Sternberg, 1985; Sternberg et al., 2021), в рамках которой исследователи оценивали интеллект у коренных жителей Аляски с помощью разработанного теста таситного (неявного) знания в сферах информации о травах, рыболовстве, выживании, фольклоре и охоте (Grigorenko et al., 2004).

Резюмируя, можно сказать, что вопрос непредвзятого оценивания является составной частью общих вопросов о качестве и валидности инструментария, а также – более глобальных дискуссий об определении интеллекта вообще (см. раздел 1) и природе групповых различий. Нам представляется, что этическая задача исследователей – как минимум быть в курсе этих обсуждений и учитывать их в собственных научных задачах, особенно если они касаются сравнительных исследований. Задача разработчиков тестов – пользоваться лучшими из существующих гайдлайнов для разработки и оценки свойств тестов, проводить полноценные исследования валидности на репрезентативных выборках и проверять, не является ли сам тест и задания источниками различий между группами. Эти задачи напрямую связаны с профессиональной *компетентностью*, которая, в свою очередь, позволяет реализовываться *принципу ответственности* – ответственности исследователей и разработчиками перед отдельными индивидами и обществом в целом. Очевидно, что не для всех инструментов будут проведены необходимые исследования и анализы – в этом случае критически важным кажется как минимум *описывать* тесты согласно международным стандартам, демонстрировать и те аспекты валидности, для которых были получены обнадеживающие результаты, и те аспекты, для которых таких результатов не обнаруживалось или которые не были оценены. Например, для этого возможно использовать Европейскую модель оценки качества тестов. Последняя версия модели была выпущена в 2013 году⁶, и имеет перевод на русский, новая версия⁷ уже доступна для предварительного просмотра и обсуждения.

⁶ Evers A., Hagemester C., Høstmælingen A., Lindley P., Muñiz J., Sjöberg A. EFPA review model for the description and evaluation of psychological and educational tests // Test review form and notes for reviewers. Version 4.2, 6. 2013. https://www.efpa.eu/sites/default/files/2023-06/110c_EFPA_BOA_TEST_REVIEW_MODEL_version426.pdf (дата обращения: 30.01.2024).

⁷ A Task Force of the Board of Assessment of EFPA. EFPA model for the review, description and evaluation of psychological and educational tests // Test review form and notes for re-

В данном случае это станет реализацией *принципа честности*. Соблюдение этого принципа позволит не опускать проблемные вопросы в область «не-сказанного» или даже «замалчиваемого», а потенциальным пользователям тестов – помнить про качество тестов, выбирать тест с учетом целевой аудитории респондентов и думать про валидность интерпретации. Мы подробнее остановимся на этических аспектах, касающихся выбора и использования качественного (или некачественного) инструментария в следующем разделе.

Как оценить качество оценочного инструментария?

Вопрос качества тестов закреплен в различных этических кодексах. Например, в этическом кодексе, принятом Российским психологическим обществом, в пункте 2.3.1 говорится об обязанности психолога пользоваться только валидными и надежными инструментами⁸. Однако именно качество используемых тестов зачастую вызывает самые большие дискуссии, и ситуация в России – не исключение.

В России при формировании рынка оценочных инструментов действовал еще один важный фактор – экономический. Современная российская психодиагностика зарождалась в начале 1980-х годов (так, первый перевод классического учебника А. Анастаси «Психологическое тестирование» увидел свет в 1982 году), и к моменту социальных и экономических потрясений 1990-х годов отечественных инструментов диагностики было еще не очень много. Как пишет А.Г. Шмелев, «...производство отечественных тестов погибло, фактически не поднявшись из младенческой колыбели. Оказалось гораздо выгоднее пользоваться краденными западными методиками, чем оригинальными отечественными, за которые отечественные специалисты законно хотели получить компенсацию вложенного квалифицированного труда.» (Шмелев, 2004, с. 42). В открытой продаже начали появляться сборники тестов (например, *Лучшие психологические тесты*, 1992), которые содержали в себе исключительно тексты заданий и способы подсчета баллов, без указаний на то, каким образом проводилась их валидизация и стандартизация. Культура публикации результатов психометрического анализа инструментов начала формироваться гораздо позже, а понимание того, что тест должен иметь опубликованные результаты проверки его качества у массового пользователя тестов, не до конца сформировано и по сей день. «Лучшие психологические тесты» теперь опубликованы на открытых порталах в интернете, а в продаже все также можно найти тесты (в том числе тест Векслера), не имеющими в комплекте технических отчетов о результатах их психометрического анализа.

Кроме того, почти любой психологический тест для диагностики интеллекта детей и подростков, используемый в России, можно найти в открытом доступе в сети интернет. Кроме вопроса о правомерном использовании закрытых профессиональных тестов, речь о котором порой не идет, мы

viewers. Version 5.0. https://www.efpa.eu/sites/default/files/2024-11/2024-11-18_test-review-model_draft-for-consultation.pdf (дата обращения: 28.12.2024).

⁸ Российское психологическое общество. *Этический кодекс психолога*. <http://psyrus.ru/rpo/documentation/ethics.php> (дата обращения: 30.01.2024).

опять возвращаемся к вопросу качества тестов. Доступность профессиональных инструментов для непрофессионалов самым негативным образом сказывается на валидности результатов и их интерпретации, так как потенциальный диагностируемый может натренироваться на прохождение тестирования.

Еще один аспект, который необходимо обсудить, – стандартизация тестов. Любой психологический тест, который предполагает использование в индивидуальной психодиагностике, требует проведения процедуры стандартизации, так как психолога в первую очередь интересует вопрос соотнесения результата отдельного человека либо с определенным критерием, либо, что более применимо к тестам на интеллект, с представляющей его популяцией. Однако использование тестов на интеллект в России уже дает нам примеры, когда тесты не имеют актуальных норм, но все равно применяются в психодиагностике. Например, широко используемый для диагностики интеллекта у детей и подростков тест Векслера до сих пор использует не просто старые нормы, но еще и полученные не на российских выборках (Филимоненко, Тимофеев, 2016). Употребление таких норм приводит к серьезным ошибкам в интерпретации результатов, особенно в тех случаях, когда интеллект является диагностирующим критерием и результаты диагностики определяют серьезные последствия для жизни ребенка.

Более того, даже нормы, полученные для целевой популяции для тестов на интеллект, требуют пересмотра не реже чем раз в десять лет, что было наглядно продемонстрировано эффектом Флинна (Flynn, 1984, 1987). В рамках этого эффекта было продемонстрировано, что каждое следующее поколение показывает более высокие показатели интеллекта, чем предыдущее. Несмотря на то, что вокруг эффекта Флинна ведется обширная дискуссия (например, Dutton et al., 2016; Kaufman, 2010; McGrew, 2010), его выявление сыграло не последнюю роль в стимулировании разработчиков тестов регулярно пересматривать нормы.

Лейтмотивом темы качества профессиональных тестов звучит применение *принципа компетентности*. Принцип компетентности преломляется во всех ситуациях, когда психолог должен сделать осознанный выбор релевантного и качественного инструментария и отказаться от использования инструментов с неподтвержденным психометрическим качеством. Также компетентность будет проявляться в использовании только тех тестов, работать с которыми психолог обучился, так как большинство тестов на интеллект, предназначенных для психодиагностики, достаточно объемны, сложны в администрировании и требуют специального обучения для применения. Нельзя игнорировать тот факт, что для того, чтобы реализовывать принцип компетентности, необходимо соответствующим образом готовить специалистов. Такая подготовка должна включать не только и не столько знакомство с существующими методиками, но овладение основами разработки тестов, психометрики и принципами проведения диагностики и представления результатов.

Выбор инструмента для диагностики неизбежно затрагивает и еще один этический принцип – *принцип ответственности*. Когда психолог использует некачественный инструмент, он всегда получает искаженный

результат, ведущий к ошибкам, которые могут иметь драматические последствия в жизни респондента. Решения, принятые на основе искаженных результатов, имеют последствия для ребенка и его родителей, так как ребенок не получает необходимой для него поддержки. Это также создает сложности школе, так как педагоги не могут учесть особенные потребности ребенка. И глобально использование некачественных инструментов наносит вред профессиональному сообществу и обществу в целом, так как ведет к недоверию к специалистам и избеганию обращения за помощью в случае необходимости.

Обсуждение

Измерение интеллекта занимает отдельное место в истории развития психодиагностики. К настоящему моменту проделан большой путь, появились процедуры адаптации тестов к различным культурным контекстам, знание о необходимости обновления норм и правила стандартизации – психометрическая наука шагнула вперед, дав возможности более точного оценивания и применения новых форматов, таких как, например, компьютерное тестирование. Однако до конца этот путь не пройден, и перед исследователями возникают запросы, на которые необходимо реагировать.

В рамках данной статьи обсуждались основные проблемы, связанные с измерением интеллекта, включая глобальные подходы к его операционализации, потенциальные нарушения прав человека в результате применения тестов, справедливость оценивания для различных групп респондентов, а также вопросы качества и выбора тестовых инструментов. С нашей точки зрения, преодоление этих трудностей требует от психологического сообщества постоянной рефлексии, направленной на обеспечение соблюдения этических принципов в психодиагностике, таких как уважение, честность, компетентность и ответственность. Такая рефлексия необходима для движения в сторону адекватного тестирования интеллекта в нашей стране. Данная задача требует совместных усилий всех специалистов — как исследователей, так и практиков.

Исходя из проанализированных вопросов, можно сделать несколько ключевых выводов. Во-первых, интеллект представляет собой важный социальный конструкт, результаты его оценки имеют значительное влияние как на уровне отдельного индивида, так и на уровне общества. Например, результаты оценки интеллекта могут стать основой для принятия важных решений, касающихся здоровья, образования или профессиональной деятельности. А на макроуровне они связаны с различными экономическими характеристиками региона или страны, а также такими страновыми характеристиками, как политические свободы, уровень здравоохранения, продолжительность жизни и так далее. Подробнее о концепции макропсихологии интеллекта в исследованиях А.В. Юревича и коллег (Юревич и др., 2007), Д.В. Ушакова и А.А. Григорьева (Ushakov, Grigoriev, 2016)]. Наша позиция заключается в том, что, учитывая большое значение оценки интеллекта для психодиагностики и исследований индивидуальных различий в целом, психологам (как исследователям, так и практикам) важно ориентироваться

в дискуссии, касающейся операционализации интеллекта и не упускать из виду «большую картину», так как теоретическая основа напрямую связана с интерпретацией результатов.

Во-вторых, справедливая оценка интеллекта представляет собой сложную задачу, затрагивающую как исследователей-разработчиков тестов, так и специалистов-практиков. Для первых проблема заключается в создании валидных и надежных тестов, которые учитывают культурные и социальные различия респондентов. Для вторых — в корректной организации тестирования и интерпретации полученных данных. Современные общественные запросы требуют от научного сообщества создания доказательной базы и обеспечения справедливости тестирования. Однако используемые в России инструменты для измерения интеллекта пока не могут в полной мере удовлетворить эти требования, что подчеркивает необходимость разработки новых диагностических инструментов, соответствующих современным реалиям.

В-третьих, необходимо отметить, что даже самый современный и валидный тест не может сам по себе решить всех проблем, связанных с оценкой интеллекта. Мы хотели бы подчеркнуть значимость контекста использования теста, и вклад каждого конкретного специалиста в валидность тестирования, результатов и их интерпретации. Например, неправомерное использование «некачественного» метода и интерпретация результатов, не подкрепленная научной аргументацией, — это решение конкретного специалиста, даже если оно не до конца осознанное и/или поддерживается определенной «традицией» использования. Мы видим необходимость в повышении чувства индивидуальной ответственности специалистов за свою профессиональную деятельность, а также безусловно — в создании благоприятной среды для развития их профессиональных компетенций и формировании общественных и государственных ожиданий как в отношении работы специалистов, так и по поводу качества психометрического инструментария.

Таким образом, общей задачей и практического, и научного профессиональных сообществ является обеспечение обоснованного и корректного использования результатов измерения интеллекта. Принятие решений на основе этих данных должно не только не причинять вреда, но приносить пользу отдельным участникам тестирования и обществу в целом.

Заключение

Измерение интеллекта является одним из ключевых вопросов психодиагностики и требует постоянного научного развития и этической рефлексии. Несмотря на достигнутый прогресс в создании более точных и совершенных инструментов измерения, остаются нерешенные вопросы, связанные с обеспечением валидности, справедливости и соблюдением прав человека при оценке интеллектуального развития. Продолжение междисциплинарного диалога и повышение ответственности специалистов как в области исследований, так и в практике тестирования представляется необходимым условием для дальнейшего совершенствования методов и повышения их полезности как для отдельных индивидов, так и для общества в целом.

Список литературы

- Анастаси А., Урбина С. Психологическое тестирование. 7-е междунар. изд. СПб.: Питер, 2007. 687 с.
- Герасимова Ю.О., Орел Е.А. Измеряя школьника: зачем мы развиваем и оцениваем социально-эмоциональные навыки // Психология. Журнал Высшей школы экономики. 2022. Т. 19. № 1. С. 61–75. <https://doi.org/10.17323/1813-8918-2022-1-61-75>
- Лучшие психологические тесты для профотбора и профориентации: Описание и руководство к использованию / отв. ред. А.Ф. Кудряшов. Петрозаводск: Петроком, 1992. 318 с.
- Ушаков Д.В. Тесты интеллекта, или Горечь самопознания // Психология. Журнал высшей школы экономики. 2004. Т. 1. № 2. С. 76–93.
- Филимоненко Ю.И., Тимофеев В.И. Тест Д. Векслера: Диагностика структуры интеллекта (детский вариант): методическое руководство. СПб.: ИМАТОН, 2016. 106 с.
- Холодная М.А. Профессиональные иллюзии как следствие упрощенных представлений о человеческом интеллекте // Психология. Журнал Высшей школы экономики. 2004а. Т. 1. № 4. С. 38–44.
- Холодная М.А. Психологическое тестирование и право личности на собственный вариант развития // Психология. Журнал Высшей школы экономики. 2004б. Т. 1. № 2. С. 66–75.
- Шмелев А.Г. Тест как оружие // Психология. Журнал Высшей школы экономики. 2004. Т. 1. № 2. С. 40–53.
- Юревич А.В., Ушаков Д.В., Цапенко И.П. Количественная оценка макropsychологического состояния современного российского общества // Психологический журнал. 2007. Т. 28. № 4. С. 23–34.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing (1st ed.). Washington, DC: American Educational Research Association, 1999. 194 p.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing (2014th ed.). Washington, DC: American Educational Research Association, 2014. 230 p.
- Bain S.K., Jaspers K.E. Test review: Review of Kaufman brief intelligence test, Second edition. Bloomington, MN: Pearson, Inc. // Journal of Psychoeducational Assessment. 2010. Vol. 28. No 2. Pp. 167–174. <https://doi.org/10.1177/0734282909348217>
- Bartram D., Hambleton R.K. The ITC guidelines: International standards and guidelines relating to tests and testing // The ITC International Handbook of Testing and Assessment / Ed. by F.T.L. Leong, D. Bartram, F. Cheung, K.F. Geisinger, D. Ilescu. Oxford: Oxford University Press, 2016. Pp. 35–46. <https://doi.org/10.1093/med:psych/9780199356942.003.0004>
- Berry J.W. Cross-cultural research and methodology series: III. Human ecology and cognitive style: Comparative studies in cultural and psychological adaptation. New York: Sage Publications/Halstead Press, 1976. 242 p.
- Berry J.W. The forgotten field: Contexts for cross-cultural psychology // Journal of Cross-Cultural Psychology. 2022. Vol. 53. No 7–8. Pp. 993–1009. <https://doi.org/10.1177/00220221221093810>
- Canivez G.L., McGill R.J., Dombrowski S.C., Watkins M.W., Pritchard A.E., Jacobson L.A. Construct validity of the WISC-V in clinical cases: Exploratory and confirmatory factor analyses of the 10 primary subtests // Assessment. 2020. Vol. 27. No 2. Pp. 274–296. <https://doi.org/10.1177/1073191118811609>

- Deary I.J. Intelligence // *Annual Review of Psychology*. 2012. Vol. 63. No 1. Pp. 453–482. <https://doi.org/10.1146/annurev-psych-120710-100353>
- Deary I.J., Penke L., Johnson W. The neuroscience of human intelligence differences // *Nature Reviews Neuroscience*. 2010. Vol. 11. No 3. Pp. 201–211. <https://doi.org/10.1038/nrn2793>
- Dutton E., van der Linden D., Lynn R. The negative Flynn Effect: A systematic literature review // *Intelligence*. 2016. Vol. 59. Pp. 163–169. <https://doi.org/10.1016/j.intell.2016.10.002>
- Fairness in Educational and Psychological Testing: Examining Theoretical, Research, Practice, and Policy Implications of the 2014 Standards / Ed. by J.L. Jonson, K.F. Geisinger. Washington: American Educational Research Association, 2022. 466 p. <https://doi.org/10.3102/9780935302967>
- Flynn J.R. Massive IQ gains in 14 nations: What IQ tests really measure // *Psychological Bulletin*. 1987. Vol. 101. No 2. Pp. 171–191. <https://doi.org/10.1037/0033-2909.101.2.171>
- Flynn J.R. The mean IQ of Americans: Massive gains 1932 to 1978 // *Psychological Bulletin*. 1984. Vol. 95. No 1. Pp. 29–51. <https://doi.org/10.1037/0033-2909.95.1.29>
- Friedlander Moore A., McCallum R.S., Bracken B.A. The universal nonverbal intelligence test: Second edition // *Handbook of Nonverbal Assessment* / Ed. by R.S. McCallum. Cham: Springer, 2017. Pp. 105–125. https://doi.org/10.1007/978-3-319-50604-3_7
- Gardner H. *Frames of mind: The theory of multiple intelligences*. New York: Basic Books, 1983. 440 p.
- Goleman D. *Emotional intelligence*. New York: Bantam Books, Inc., 1995. 352 p.
- Gottfredson L.S. Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography // *Intelligence*. 1997. Vol. 24. No 1. Pp. 13–23. [https://doi.org/10.1016/s0160-2896\(97\)90011-8](https://doi.org/10.1016/s0160-2896(97)90011-8)
- Gould S.J. *The Mismeasure of Man*. New York: WW Norton, 1981. 352 p.
- Gregory R.J. *Psychological Testing: History, Principles and Applications* (7th ed.). Harlow: Pearson, 2015. 589 p.
- Grigorenko E.L., Meier E., Lipka J., Mohatt G., Yanez E., Sternberg R.J. Academic and practical intelligence: A case study of the Yup'ik in Alaska // *Learning and Individual Differences*. 2004. Vol. 14. No 4. Pp. 183–207. <https://doi.org/10.1016/j.lindif.2004.02.002>
- Hunter J.E., Schmidt F.L. Critical analysis of the statistical and ethical implications of various definitions of test bias // *Psychological Bulletin*. 1976. Vol. 83. No 6. Pp. 1053–1071. <https://doi.org/10.1037/0033-2909.83.6.1053>
- Jensen A.R. *Bias in Mental Testing*. New York: The Free Press, 1980. 786 p.
- Kamin L., Egerton J. The misuse of IQ testing // *Change: The Magazine of Higher Learning*. 1973. Vol. 5. No 8. Pp. 40–43. <https://doi.org/10.1080/00091383.1973.10568572>
- Kaufman A.S. Looking through Flynn's rose-colored scientific spectacles // *Journal of Psychoeducational Assessment*. 2010. Vol. 28. No 5. Pp. 494–505. <https://doi.org/10.1177/0734282910373573>
- Kaufman A.S., Kaufman N.L. Kaufman assessment battery for children (KABC, K-ABC) [Database record] // *APA PsycTests*. 1983. <https://doi.org/10.1037/t27677-000>
- Kuncel N.R., Sackett P.R. Selective citation mars conclusions about test validity and predictive bias // *American Psychologist*. 2007. Vol. 62. No 2. Pp. 145–146. <https://doi.org/10.1037/003-066x62.2.145>
- McGrew K.S. The Flynn effect and its critics: Rusty linchpins and “lookin’ for g and Gf in some of the wrong places” // *Journal of Psychoeducational Assessment*. 2010. Vol. 28. No 5. Pp. 448–468. <https://doi.org/10.1177/0734282910373347>
- Meuleman B., Żółtak T., Pokropek A., Davidov E., Muthén B., Oberski D.L., Billiet J., Schmidt P. Why measurement invariance is important in comparative research. A response to Welzel et al. (2021) // *Sociological Methods & Research*. 2023. Vol. 52. No 3. Pp. 1401–1419. <https://doi.org/10.1177/00491241221091755>

- Naglieri J.A. Hundred years of intelligence testing: Moving from traditional IQ to second-generation intelligence tests // *Handbook of Intelligence* / Ed. by S. Goldstein, D. Princiotta, J. Naglieri. New York: Springer, 2015. Pp. 295–316. https://doi.org/10.1007/978-1-4939-1562-0_20
- Naglieri J.A., Das J.P. Cognitive assessment system (Vol. 96). Itasca, IL: Riverside Publishing, 1997.
- Neisser U., Boodoo G., Bouchard T.J., Boykin A.W., Brody N., Ceci S.J., Halpern D.F., Loehlin J.C., Perloff R., Sternberg R.J., Urbina S. Intelligence: Knowns and unknowns // *American Psychologist*. 1996. Vol. 51. No 2. Pp. 77–101. <https://doi.org/10.1037/0003-066x.51.2.77>
- Powers K.M., Hagans-Murillo K.S., Restori A.F. Twenty-five years after Larry P.: The California response to overrepresentation of African Americans in special education // *The California School Psychologist*. 2004. Vol. 9. No 1. Pp. 145–158. <https://doi.org/10.1007/bf03340915>
- Putnick D.L., Bornstein M.H. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research // *Developmental Review*. 2016. Vol. 41. Pp. 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Reynolds C. Bias in testing // *Encyclopedia of human intelligence* / Ed. by R.J. Sternberg. New York: Macmillan, 1994. Vol. 1. Pp. 175–178.
- Reynolds C.R., Suzuki L.A. Bias in psychological assessment: An empirical review and recommendations // *Handbook of Psychology*. 2nd ed. / Ed. by I. Weiner. Hoboken: John Wiley & Sons, Inc., 2012. Pp. 82–113. <https://doi.org/10.1002/9781118133880.hop210004>
- Roid G.H., Koch C. Leiter-3: Nonverbal cognitive and neuropsychological assessment // *Handbook of Nonverbal Assessment* / Ed. by R. McCallum. Cham: Springer, 2017. Pp. 127–150. https://doi.org/10.1007/978-3-319-50604-3_8
- Roth B., Becker N., Romeyke S., Schäfer S., Domnick F., Spinath F. M. Intelligence and school grades: A meta-analysis // *Intelligence*. 2015. Vol. 53. Pp. 118–137. <https://doi.org/10.1016/j.intell.2015.09.002>
- Sattler J.M. Assessment of children: Cognitive foundations and applications. 5th ed. La Mesa, CA: Jerome M. Sattler, Publisher, Inc., 2008. 790 p.
- Stanovich K.E. What intelligence tests miss: The psychology of rational thought. New Haven, CT: Yale University Press, 2009. 308 p. <https://doi.org/10.12987/9780300142532>
- Stanovich K.E., West R.F., Toplak M.E. The Rationality Quotient: Toward a Test of Rational Thinking. Cambridge, MA: The MIT Press, 2016. 480 p. <https://doi.org/10.7551/mitpress/9780262034845.001.0001>
- Sternberg R.J., Glaveanu V., Karami S., Kaufman J.C., Phillipson S.N., Preiss D.D. Meta-intelligence: Understanding, control, and interactivity between creative, analytical, practical, and wisdom-based approaches in problem solving // *Journal of Intelligence*. 2021. Vol. 9. No 2. P. 19. <https://doi.org/10.3390/jintelligence9020019>
- Sternberg R.J., Grigorenko E., Bundy D.A. The predictive value of IQ // *Merrill-Palmer Quarterly*. 2001. Vol. 47. No 1. Pp. 1–41. <https://doi.org/10.1353/mpq.2001.0005>
- Sternberg R.J., Preiss D.D., Karami S. An historical causal-chain theory of conceptions of intelligence // *Review of General Psychology*. 2023. Vol. 27. No 3. Pp. 320–335. <https://doi.org/10.1177/10892680231158790>
- Sternberg, R.J. Beyond IQ: A triarchic theory of human intelligence. Cambridge: Cambridge University Press., 1985. 432 p.
- Strenze T. Intelligence and socioeconomic success: A meta-analytic review of longitudinal research // *Intelligence*. 2007. Vol. 35. No 5. Pp. 401–426. <https://doi.org/10.1016/j.intell.2006.09.004>

- Terman L.M.* The uses of intelligence tests // The measurement of intelligence. Boston: Houghton, Mifflin and Company, 1916. Pp. 3–21. <https://doi.org/10.1037/10014-001>
- The IQ controversy: Critical readings / Ed. by N.J. Block, G. Dworkin. New York: Pantheon Books, 1976. 557 p.
- Thorndike E.L.* Intelligence and its uses // Harper's Magazine. 1920. Vol. 140. Pp. 227–235.
- Ushakov D.V., Grigoriev A.A.* Macropsychology of intelligence: through emotions to theoretical depth // Психология. Журнал Высшей школы экономики. 2016. Т. 13. № 4. С. 629–635. <https://doi.org/10.17323/1813-8918-2016-4-629-635>
- Van de Vijver F., Tanzer N.K.* Bias and equivalence in cross-cultural assessment: An overview // European Review of Applied Psychology. 2004. Vol. 54. No 2. Pp. 119–135. <https://doi.org/10.1016/j.erap.2003.12.004>
- Van de Vijver F.J.R., Leung K.* Methods and Data Analysis for Cross-Cultural Research. 2nd ed. Cambridge: Cambridge University Press, 2021. <https://doi.org/10.1017/9781107415188>
- Van de Vijver F.J.R., Poortinga Y.H.* Towards an integrated analysis of bias in cross-cultural assessment // European Journal of Psychological Assessment. 1997. Vol. 13. No 1. Pp. 29–37. <https://doi.org/10.1027/1015-5759.13.1.29>
- Warne R.T.* In the Know: Debunking 35 Myths about Human Intelligence. Cambridge: Cambridge University Press, 2020. 434 p. <https://doi.org/10.1017/9781108593298>
- Welzel C., Brunkert L., Kruse S., Inglehart R.F.* Non-invariance? An overstated problem with misconceived causes // Sociological Methods & Research. 2023. Vol. 52. No 3. Pp. 1368–1400. <https://doi.org/10.1177/0049124121995521>
- Wicherts J.M.* The importance of measurement invariance in neurocognitive ability testing // The Clinical Neuropsychologist. 2016. Vol. 30. No 7. Pp. 1006–1016. <https://doi.org/10.1080/13854046.2016.1205136>
- Zaboski B.A.II, Kranzler J.H., Gage N.A.* Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-Horn-Carroll Theory // Journal of School Psychology. 2018. Vol. 71. Pp. 42–56. <https://doi.org/10.1016/j.jsp.2018.10.001>

История статьи:

Поступила в редакцию 31 января 2024 г.

Принята к печати 2 апреля 2024 г.

Для цитирования:

Логвиненко Т.И., Канонир Т.Н., Орел Е.А., Куликова А.А. Этические аспекты измерения интеллекта: навстречу компетентности и справедливости // Вестник Российского университета дружбы народов. Серия: Психология и педагогика. 2024. Т. 21. № 2. С. 657–682. <http://doi.org/10.22363/2313-1683-2024-21-2-657-682>

Вклад авторов:

Т.И. Логвиненко – концепция и дизайн исследования, первоначальный текст рукописи, итоговое редактирование текста. Т.Н. Канонир – концепция и дизайн исследования, первоначальный текст рукописи. Е.А. Орел – концепция исследования, написание отдельных разделов статьи. А.А. Куликова – концепция исследования, написание отдельных разделов статьи.

Заявление о конфликте интересов:

Авторы заявляют об отсутствии конфликта интересов.

Сведения об авторах:

Логвиненко Татьяна Игоревна, старший специалист, Научный центр когнитивных исследований, Научно-технологический университет «Сириус», Российская Федерация, 354340, Федеральная территория «Сириус», Олимпийский пр-т, 1. ORCID: 0000-0001-7430-1963, SPIN-код: 8068-9856. E-mail: logvinenkota.spb@gmail.com.

Канонир Татьяна Николаевна, Dr.Psych. (PhD), доцент, Институт образования, Национальный исследовательский университет «Высшая школа экономики», Российская Федерация, 101000, Москва, ул. Мясницкая, д. 20. ORCID: 0000-0001-5606-8379. E-mail: tkanonir@hse.ru.

Орел Екатерина Алексеевна, кандидат психологических наук, PhD, старший научный сотрудник, Институт образования, Национальный исследовательский университет «Высшая школа экономики», Российская Федерация, 101000, Москва, ул. Мясницкая, д. 20. ORCID: 0000-0002-9100-0713. E-mail: eorel@hse.ru.

Куликова Алена Александровна, кандидат наук об образовании, приглашенный преподаватель, Институт образования, Национальный исследовательский университет «Высшая школа экономики», Российская Федерация, 101000, Москва, ул. Мясницкая, д. 20. ORCID: 0000-0002-4296-3521. E-mail: aponomareva@hse.ru.

DOI: 10.22363/2313-1683-2024-21-2-657-682

EDN: MCKAAY

UDC 159.9.072

Theoretical article

Ethical Aspects of Measuring Intelligence: Towards Competence and Fairness

Tatiana I. Logvinenko¹, Tatiana N. Kanonir²,
Ekaterina A. Orel², Alena A. Kulikova²

¹Sirius University of Science and Technology, The Sirius Federal Territory,
Russian Federation

²HSE University,
Moscow, Russian Federation

 logvinenkota.spb@gmail.com

Abstract. The article is focused on the problem of intelligence measurement, with an emphasis on the ethical aspects of developing and using tests. The history of intelligence measurement provides a variety of examples, problematic from an ethical point of view, which have repeatedly led to negative consequences for both individuals and entire communities. The purpose of this article is to describe current ethical issues in the field of intelligence measurement, their background and historical examples. We discuss the ethical issues in terms of: (1) global approaches to operationalizing intelligence; 2) possible human rights violations resulting from the use of intelligence tests; 3) the fairness of intelligence tests for different groups of respondents; and 4) assessment of test quality in test selection. These issues are examined through the prism of the ethical principles of psychologists, such as respect, honesty, competence, and responsibility. Despite the extensive history of measuring intelli-

gence and research in this area, ethical issues raised decades ago have not lost their relevance. Since ethical questions often do not have clear-cut answers, we believe that engaging in discussions about ethical issues in intelligence testing and exploring potential solutions is itself important and warranted. The content and conclusions of this article may be useful for both researchers and practitioners to make informed decisions in the context of intelligence measurement.

Key words: intelligence, IQ, intelligence test, psychodiagnostics, psychological assessment, ethical principles, ethics code

Acknowledgements and Funding: The study was supported by a grant from the Russian Science Foundation (project No. 23-28-01399).

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing* (1st ed.). Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing* (2014th ed.). Washington, DC: American Educational Research Association.
- Anastasi, A., & Urbina, S. (2007). *Psychological testing* (7th ed.). Saint Petersburg: Piter. (In Russ.)
- Bain, S.K., & Jaspers, K.E. (2010). Test review: Review of Kaufman brief intelligence test, Second edition. Bloomington, MN: Pearson, Inc. *Journal of Psychoeducational Assessment*, 28(2), 167–174. <https://doi.org/10.1177/0734282909348217>
- Bartram, D., & Hambleton, R.K. (2016). The ITC guidelines: International standards and guidelines relating to tests and testing. In F.T.L. Leong, D. Bartram, F. Cheung, K.F. Geisinger, & D. Iliescu (Eds.). *The ITC International Handbook of Testing and Assessment* (pp. 35–46). Oxford: Oxford University Press. <https://doi.org/10.1093/med:psych/9780199356942.003.0004>
- Berry, J.W. (1976). *Cross-cultural research and methodology series: III. Human ecology and cognitive style: Comparative studies in cultural and psychological adaptation*. New York: Sage Publications/Halstead Press.
- Berry, J.W. (2022). The forgotten field: Contexts for cross-cultural psychology. *Journal of Cross-Cultural Psychology*, 53(7–8), 993–1009. <https://doi.org/10.1177/00220221221093810>
- Block, N.J., & Dworkin, G. (Eds.). (1976). *The IQ controversy: Critical readings*. New York: Pantheon Books.
- Canivez, G.L., McGill, R.J., Dombrowski, S.C., Watkins, M.W., Pritchard, A.E., & Jacobson, L.A. (2020). Construct validity of the WISC-V in clinical cases: Exploratory and confirmatory factor analyses of the 10 primary subtests. *Assessment*, 27(2), 274–296. <https://doi.org/10.1177/1073191118811609>
- Deary, I.J. (2012). Intelligence. *Annual Review of Psychology*, 63(1), 453–482. <https://doi.org/10.1146/annurev-psych-120710-100353>
- Deary, I.J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, 11(3), 201–211. <https://doi.org/10.1038/nrn2793>
- Dutton, E., van der Linden, D., & Lynn, R. (2016). The negative Flynn Effect: A systematic literature review. *Intelligence*, 59, 163–169. <https://doi.org/10.1016/j.intell.2016.10.002>
- Filimonenko, Yu. I., & Timofeev, V.I. (2016). *Wechsler Intelligence Scale for Children (WISC)*. Saint Petersburg: IMATON. (In Russ.)
- Flynn, J.R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95(1), 29–51. <https://doi.org/10.1037/0033-2909.95.1.29>

- Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101(2), 171–191. <https://doi.org/10.1037/0033-2909.101.2.171>
- Friedlander Moore, A., McCallum, R.S., & Bracken, B.A. (2017). The universal nonverbal intelligence test: Second edition. In R.S. McCallum (Eds.). *Handbook of Nonverbal Assessment* (pp. 105–125). Cham: Springer. https://doi.org/10.1007/978-3-319-50604-3_7
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Gerasimova, Yu., & Orel, E. (2022). Measuring a schoolchild: Why we develop and assess social-emotional skills. *Psychology. Journal of the Higher School of Economics*, 19(1), 61–75. (In Russ.) <https://doi.org/10.17323/1813-8918-2022-1-61-75>
- Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books, Inc.
- Gottfredson, L.S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13–23. [https://doi.org/10.1016/s0160-2896\(97\)90011-8](https://doi.org/10.1016/s0160-2896(97)90011-8)
- Gould, S.J. (1981). *The Mismeasure of Man*. New York: WW Norton.
- Gregory, R.J. (2015). *Psychological Testing: History, Principles and Applications* (7th ed.). Harlow: Pearson.
- Grigorenko, E.L., Meier, E., Lipka, J., Mohatt, G., Yanez, E., & Sternberg, R.J. (2004). Academic and practical intelligence: A case study of the Yup'ik in Alaska. *Learning and Individual Differences*, 14(4), 183–207. <https://doi.org/10.1016/j.lindif.2004.02.002>
- Hunter, J.E., & Schmidt, F.L. (1976). Critical analysis of the statistical and ethical implications of various definitions of test bias. *Psychological Bulletin*, 83(6), 1053–1071. <https://doi.org/10.1037/0033-2909.83.6.1053>
- Jensen, A.R. (1980). *Bias in Mental Testing*. New York: The Free Press.
- Jonson, J.L., & Geisinger, K.F. (Eds.). (2022). *Fairness in Educational and Psychological Testing: Examining Theoretical, Research, Practice, and Policy Implications of the 2014 Standards*. Washington: American Educational Research Association. <https://doi.org/10.3102/9780935302967>
- Kamin, L., & Egerton, J. (1973). The misuse of IQ testing. *Change: The Magazine of Higher Learning*, 5(8), 40–43. <https://doi.org/10.1080/00091383.1973.10568572>
- Kaufman, A.S. (2010). Looking through Flynn's rose-colored scientific spectacles. *Journal of Psychoeducational Assessment*, 28(5), 494–505. <https://doi.org/10.1177/0734282910373573>
- Kaufman, A.S., & Kaufman, N.L. (1983). *Kaufman assessment battery for children (KABC, K-ABC)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t27677-000>
- Kholodnaia, M.A. (2004a). Professional illusions resulting from simplified ideas about human intelligence. *Psychology. Journal of the Higher School of Economics*, 1(4), 38–44. (In Russ.)
- Kholodnaia, M.A. (2004b). Psychological testing and the right of a person to chose her/his own path of development. *Psychology. Journal of the Higher School of Economics*, 1(2), 66–75. (In Russ.)
- Kudryashov, A.F. (Ed.). (1992). *Best psychological tests for vocational selection and career guidance: Description and user guide*. Petrozavodsk: Petrokom (In Russ.)
- Kuncel, N.R., & Sackett, P.R. (2007). Selective citation mars conclusions about test validity and predictive bias. *American Psychologist*, 62(2), 145–146. <https://doi.org/10.1037/003-066x62.2.145>
- McGrew, K.S. (2010). The Flynn effect and its critics: Rusty linchpins and “lookin’ for g and Gf in some of the wrong places”. *Journal of Psychoeducational Assessment*, 28(5), 448–468. <https://doi.org/10.1177/0734282910373347>

- Meuleman, B., Żółtak, T., Pokropek, A., Davidov, E., Muthén, B., Oberski, D.L., Billiet, J., & Schmidt, P. (2023). Why measurement invariance is important in comparative research. A response to Welzel et al. (2021). *Sociological Methods & Research*, 52(3), 1401–1419. <https://doi.org/10.1177/00491241221091755>
- Naglieri, J.A. (2015). Hundred years of intelligence testing: Moving from traditional IQ to second-generation intelligence tests. In Goldstein, S., Princiotta, D., Naglieri, J. (Eds.). *Handbook of Intelligence* (pp. 295–316). New York: Springer. https://doi.org/10.1007/978-1-4939-1562-0_20
- Naglieri, J.A., Das, J.P. (1997). *Cognitive assessment system* (Vol. 96). Itasca, IL: Riverside Publishing.
- Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77–101. <https://doi.org/10.1037/0003-066x.51.2.77>
- Powers, K.M., Hagans-Murillo, K.S., & Restori, A.F. (2004). Twenty-five years after Larry P.: The California response to overrepresentation of African Americans in special education. *The California School Psychologist*, 9(1), 145–158. <https://doi.org/10.1007/bf03340915>
- Putnick, D.L., & Bornstein, M.H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- Reynolds, C. (1994). Bias in testing. In R.J. Sternberg (Ed.). *Encyclopedia of human intelligence* (Vol. 1, pp. 175–178). New York: Macmillan.
- Reynolds, C.R., & Suzuki, L.A. (2012). Bias in psychological assessment: An empirical review and recommendations. In I. Weiner (Ed.). *Handbook of Psychology*, 2nd ed. (pp. 82–113). Hoboken: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118133880.hop210004>
- Roid, G.H., & Koch, C. (2017). Leiter-3: Nonverbal cognitive and neuropsychological assessment. In R. McCallum (Ed.). *Handbook of Nonverbal Assessment* (pp. 127–150). Cham: Springer. https://doi.org/10.1007/978-3-319-50604-3_8
- Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, 53, 118–137. <https://doi.org/10.1016/j.intell.2015.09.002>
- Sattler, J.M. (2008). *Assessment of children: Cognitive foundations and applications* (5th ed.). La Mesa, CA: Jerome M. Sattler, Publisher, Inc.
- Shmelev, A.G. (2004). A test as a weapon. *Psychology. Journal of the Higher School of Economics*, 1(2), 40–53. (In Russ.)
- Stanovich, K.E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press. <https://doi.org/10.12987/9780300142532>
- Stanovich, K.E., West, R.F., & Toplak, M.E. (2016). *The Rationality Quotient: Toward a Test of Rational Thinking*. Cambridge, MA: The MIT Press. <https://doi.org/10.7551/mitpress/9780262034845.001.0001>
- Sternberg, R.J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: Cambridge University Press.
- Sternberg, R.J., Glaveanu, V., Karami, S., Kaufman, J.C., Phillipson, S.N., & Preiss, D.D. (2021). Meta-intelligence: Understanding, control, and interactivity between creative, analytical, practical, and wisdom-based approaches in problem solving. *Journal of Intelligence*, 9(2), 19. <https://doi.org/10.3390/jintelligence9020019>
- Sternberg, R.J., Grigorenko, E., & Bundy, D.A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, 47(1), 1–41. <https://doi.org/10.1353/mpq.2001.0005>
- Sternberg, R.J., Preiss, D.D., & Karami, S. (2023). An historical causal-chain theory of conceptions of intelligence. *Review of General Psychology*, 27(3), 320–335. <https://doi.org/10.1177/10892680231158790>

- Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, 35(5), 401–426. <https://doi.org/10.1016/j.intell.2006.09.004>
- Terman, L.M. (1916). The uses of intelligence tests. In *The measurement of intelligence* (pp. 3–21). Boston: Houghton, Mifflin and Company. <https://doi.org/10.1037/10014-001>
- Thorndike, E.L. (1920). Intelligence and its uses. *Harper's Magazine*, 140, 227–235.
- Ushakov, D.V. (2004). Intelligence tests, or the bitter taste of self-knowledge. *Psychology. Journal of the Higher School of Economics*, 1(2), 76–93. (In Russ.)
- Ushakov, D.V., & Grigoriev, A.A. (2016). Macropsychology of intelligence: through emotions to theoretical depth. *Psychology. Journal of the Higher School of Economics*, 13(4), 629–635. <https://doi.org/10.17323/1813-8918-2016-4-629-635>
- Van de Vijver, F., & Tanzer, N.K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 54(2), 119–135. <https://doi.org/10.1016/j.erap.2003.12.004>
- Van de Vijver, F.J.R., & Leung, K. (2021). *Methods and Data Analysis for Cross-Cultural Research* (2nd ed.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781107415188>
- Van de Vijver, F.J.R., & Poortinga, Y.H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13(1), 29–37. <https://doi.org/10.1027/1015-5759.13.1.29>
- Warne, R.T. (2020). *In the Know: Debunking 35 Myths about Human Intelligence*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108593298>
- Welzel, C., Brunkert, L., Kruse, S., & Inglehart, R.F. (2023). Non-invariance? An overstated problem with misconceived causes. *Sociological Methods & Research*, 52(3), 1368–1400. <https://doi.org/10.1177/0049124121995521>
- Wicherts, J.M. (2016). The importance of measurement invariance in neurocognitive ability testing. *The Clinical Neuropsychologist*, 30(7), 1006–1016. <https://doi.org/10.1080/13854046.2016.1205136>
- Yurevich, A.V., Ushakov, L.V., & Tsapenko, I.P. (2007). Quantitative evaluation of present-day Russian society macropsychological state. *Psikhologicheskii zhurnal*, 28(4), 23–34. (In Russ.)
- Zaboski, B.A.II, Kranzler, J.H., & Gage, N.A. (2018). Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-Horn-Carroll Theory. *Journal of School Psychology*, 71, 42–56. <https://doi.org/10.1016/j.jsp.2018.10.001>

Article history:

Received 31 January 2024

Revised 31 March 2024

Accepted 2 April 2024

For citation:

Logvinenko, T.I., Kanonir, T.N., Orel, E.A., & Kulikova, A.A. (2024). Ethical aspects of measuring intelligence: Towards competence and fairness. *RUDN Journal of Psychology and Pedagogics*, 21(2), 657–682. (In Russ.) <http://doi.org/10.22363/2313-1683-2024-21-2-657-682>

Author's contribution:

Tatiana I. Logvinenko – study concept and design, manuscript draft, final text editing. Tatiana N. Kanonir – study concept and design, manuscript draft. Ekaterina A. Orel – study concept, writing of individual article sections. Alena A. Kulikova – study concept, writing of individual article sections.

Conflicts of interest:

The authors declare that there is no conflict of interest.

Bio notes:

Tatiana I. Logvinenko, Research Fellow, Center for Cognitive Sciences, Sirius University of Science and Technology, 1 Olimpiyskiy Ave., 354340 The Sirius Federal Territory, Russian Federation. ORCID: 0000-0001-7430-1963, SPIN-code: 8068-9856. E-mail: logvinenkota.spb@gmail.com.

Tatiana N. Kanonir, Dr. Psych. (PhD), Associate Professor, Institute of Education, HSE University, 20 Myasnitskaya St, 101000 Moscow, Russian Federation. ORCID: 0000-0001-5606-8379. E-mail: tkanonir@hse.ru.

Ekaterina A. Orel, PhD in Psychology, Senior Research Fellow, Institute of Education, HSE University, Candidate of Psychological Sciences, 20 Myasnitskaya St, 101000 Moscow, Russian Federation. ORCID: 0000-0002-9100-0713. E-mail: eorel@hse.ru.

Alena A. Kulikova, PhD in Education, Research Fellow, Institute of Education, HSE University, 20 Myasnitskaya St, 101000 Moscow, Russian Federation. ORCID: 0000-0002-4296-3521. E-mail: aponomareva@hse.ru.