

Оценка информационно-коммуникационной компетентности учащихся: подходы, инструмент, валидность и надежность результатов

**С. М. Авдеева, М. Г. Руднев, Г. М. Васин,
К. В. Тарасова, Д. М. Панова**

Статья поступила
в редакцию
в июле 2017 г.

Авдеева Светлана Михайловна
кандидат технических наук, заместитель исполнительного директора Национального фонда подготовки кадров, заместитель руководителя Центра мониторинга качества образования Института образования Национального исследовательского университета «Высшая школа экономики». Адрес: 123022, Москва, ул. 1905 года, 7, стр. 1. E-mail: avdeeva@ntf.ru

Руднев Максим Геннадьевич
кандидат социологических наук, ведущий научный сотрудник Лаборатории сравнительных исследований массового сознания Национального исследовательского университета «Высшая школа экономики». Адрес: 101000, Москва, ул. Мясницкая, 20. E-mail: mrudnev@hse.ru

Васин Георгий Михайлович
психометрик, департамент развития ИКТ в образовании Национального фонда подготовки кадров. Адрес: 123022, Москва, ул. 1905 года, 7, стр. 1. E-mail: vasin@ntf.ru

Тарасова Ксения Вадимовна
кандидат педагогических наук, ведущий специалист Национального фонда подготовки кадров. Адрес: 123022 Москва, ул. 1905 года, 7, стр. 1. E-mail: kvtarasova@mail.ru

Панова Дарья Михайловна
разработчик тестов, Национальный фонд подготовки кадров. Адрес: 123022, Москва, ул. 1905 года, 7, стр. 1. E-mail: panova@ntf.ru

Аннотация. Представлен инструмент для измерения информационно-коммуникационной компетентности (ИКК) в цифровой среде, предназначенный для учащихся 7–9-х классов основной школы — Information and Communication Literacy Test (ICL-тест). Обзор существующих инструментов измерения показал, что других подходящих для российских учащихся инструментов по оценке ИКК не существует. Тест создавался с использованием инновационного систематического подхода к разработке тестов (evidence centered design). Для обсчета результатов ICL-теста были применены сети Байеса, работающие в данном случае лучше современной теории тестов. ICL-тест показал высокую конструктивную, содержательную, конвергентную, дивергентную и очевидную валидность, а также высокую надежность. Инструмент соответствует как российским, так и международным образовательным стандартам по критериям Вебба. На основе ICL-теста были проведены исследования по оценке ИКК и выявлены факторы, влияющие на ее формирование в регионах РФ, Республике Армения и Республике Белоруссия. Авторы делают вывод, что ICL-тест может быть применен как в российской, так и в других системах образования.

Авторы выражают благодарность за помощь в осуществлении проекта Центру тестирования Республики Армения, коллегам из Белоруссии, Барнаула, Красноярска, Москвы, Республики Татарстан и лично Марку Зелману.

Ключевые слова: информационно-коммуникационная компетентность, тесты, Information and Communication Literacy Test, валидность, компетентность, оценка компетентности, метапредметные образовательные результаты. DOI: 10.17323/1814-9545-2017-4-104-132

Переход к информационному обществу радикально меняет требования к качеству образования и тем компетенциям, которые школьники и студенты приобретают в ходе обучения. В специальном докладе недавнего Всемирного экономического форума [World Economic Forum, 2015] была подчеркнута значимость «навыков XXI века», среди которых способность критически мыслить, решать проблемы, аргументировать, анализировать, синтезировать информацию. Значимость этих навыков обусловлена прежде всего трансформирующейся природой человеческой деятельности: функциональные обязанности работников и требования рынка труда изменяются, в современном обществе основным предметом труда становится информация. Важное место среди качеств, необходимых современному работнику, занимают «информационная и коммуникационная компетентность, медиа- и интернет-грамотность, квалифицированная оценка и анализ информации и данных, компьютерное программирование» [Abbot, 2014]. Высокий уровень развития этих навыков является залогом того, что человек будет способен успешно жить и работать в условиях информационного общества.

Изменившееся общество диктует новые задачи системе образования. К их числу относятся переподготовка взрослых кадров, но в первую очередь — адаптация школьного образования к новым требованиям. В процессе обучения каждый школьник должен не только освоить основные предметы программы, но и приобрести навыки самообучения, критического мышления, быстрого анализа информации и коммуникации. Собственно, для получения современного образования важно уметь самостоятельно работать с информацией.

В 2009–2012 гг. для всех уровней образования были приняты новые федеральные государственные образовательные стандарты, определяющие требования к образовательным результатам выпускников. Помимо знания основных фактов и сведений, входящих в состав предметных областей, теперь требуются метапредметные результаты обучения: умение самостоятельно определять цели своего обучения, устанавливать причинно-следственные связи, строить логические рассуждения, анализировать информацию, выделять проблему, синтезировать решения, обобщать и делать выводы, используя навыки коммуникации, взаимодействия, коллективной работы, разрешения споров. Еще одним новым требованием к образовательным результатам ста-

ло формирование у учащихся способности применять полученные знания и умения для решения как учебных, так и практических задач. Эти отдельные нетехнические компетентности можно объединить под названием информационно-коммуникационной компетентности (ИКК). С изменением представлений об образовательных результатах меняются и методы и инструменты их оценивания. Актуальной задачей, в частности, становится разработка инструментов, позволяющих оценивать ИКК.

Существующие инструменты оценки ИКК чаще всего тестируют техническую грамотность, нередко полученные с их помощью показатели отражают одновременно и технические, и информационные компетенции. Средств оценить навыки, которые позволяют эффективно работать с информацией в высокотехнологичной цифровой среде, сегодня в России практически нет. Существующие измерители компетентности школьников в использовании ИКТ — это итоговая оценка по курсу информатики, результаты Основного государственного экзамена и Единого государственного экзамена или сертификаты ИТ-компаний, например ECDL-сертификат. Использовать эти показатели для определения уровня информационно-коммуникационной компетентности школьников не представляется возможным.

По окончании основной школы часть школьников уже переходит на рынок труда, а большинство остается в системе образования, и, значит, есть возможность повлиять на развитие их компетенций. Поэтому измерение ИКК именно в этот момент важно как для самих школьников, так и для политиков системы образования.

В Национальном фонде подготовки кадров разработан инструмент, позволяющий оценить ИКК выпускников основной школы, — Information and Communication Literacy Test (ICL-тест). В его основу положена компетентностная модель решения информационных задач, включающая основные когнитивные действия при работе с информацией: оценка, интеграция, передача и др. [Educational Testing Service, 2002a]. Особенностью данного инструмента является то, что он предназначен не для определения уровня владения определенным программным продуктом или техническими возможностями компьютера, а для оценки способности выпускников основной школы использовать компьютер и другие современные ИКТ для получения новых знаний, осуществления коммуникации, проведения исследовательской деятельности — способности, которая должна помочь им приобрести навыки непрерывного обучения в течение всей жизни и стать успешными в выбранной профессии.

В статье описан ICL-тест и представлены результаты исследования его валидности и надежности на основании нескольких тестирований в регионах России, Белоруссии и Армении. В первом разделе рассматриваются различные трактовки ИКК

и близких понятий, а также существующие инструменты ее измерения, далее описан сам инструмент и приводятся результаты исследования его валидности, включая соответствие российским и международным образовательным стандартам, связь результатов теста с внешними переменными, такими как техническая грамотность и самооценка ИКК, средний балл в школе и оценки по государственным экзаменам, а также демонстрируется надежность, включая внутреннюю согласованность и устойчивость результатов во времени. В заключении обсуждаются возможности использования ICL-теста как в российской системе образования, так и в международных сравнениях.

Наряду с понятием информационно-коммуникационной компетентности в литературе используются такие термины, как «ИКТ-компетентность», «компьютерная компетентность», «компьютерная грамотность», «технологическая грамотность», «информационная грамотность», «медиаграмотность», «медиапедагогическая компетентность», «информационно-технологическая компетентность», «цифровая компетентность». Эти названия, к сожалению, не отражают различий в содержании соответствующих понятий, которые заключаются прежде всего в том, каким образом в них сочетаются технологическая (навыки владения определенными программными продуктами или компьютерными устройствами) и информационная (обобщенные когнитивные навыки) компетентность/грамотность¹ [Australian Library and Information Association, 2001; Hepworth, 2000; Johnston, Webber, 2003; Blömeke, 2000; Malta Qualifications Council, 2007].

Согласно Федеральному государственному образовательному стандарту основного общего образования РФ, в числе метапредметных образовательных результатов рассматривается и ИКТ-компетенция.

Метапредметные результаты освоения основной образовательной программы основного общего образования должны отражать:

<...>

б) умение определять понятия, создавать обобщения, устанавливать аналогии, классифицировать, самостоятельно выбирать основания и критерии для классификации, устанавливать причинно-следственные связи, строить логическое рассуждение, умозаключение (индуктивное, дедуктивное и по аналогии) и делать выводы;

¹ Большинство исследователей объединяют понятия «компетентность» и «грамотность», далее мы используем термин «компетентность», чтобы не множить сущности.

1. Понятие информационно-коммуникационной компетентности и существующие инструменты для ее измерения

- 7) умение создавать, применять и преобразовывать знаки и символы, модели и схемы для решения учебных и познавательных задач <...>;
- 11) формирование и развитие компетентности в области использования информационно-коммуникационных технологий (далее ИКТ-компетенции)»².

Согласно ФГОС, ИКТ-компетенция — это способность использовать ИКТ для работы с информацией при решении учебных задач, в процессе самостоятельной познавательной деятельности, жизни и труда в современном высокотехнологичном обществе. ИКТ-компетенция включает не только умение учащегося работать с информацией в цифровой среде, но и навыки обращения с определенным оборудованием, программным обеспечением и проч.³

Число инструментов, измеряющих ИКТ и близкие конструкты, приближается к сотне, и полный их обзор представлен в [Virkus, 2003; Rader, 2002; Johnston, Webber, 2003], а наиболее свежий обзор приведен в [Kim, Kil, Shin, 2014]. Мы разбили все найденные инструменты на четыре группы в соответствии с теми конструктами, которые они измеряют⁴.

В первую группу вошли тесты, оценивающие навыки работы с конкретным программным обеспечением (Microsoft Word, Excel и др.) и оборудованием, например International Computer Drivers License (ICDL)⁵.

Вторая группа инструментов измеряет информационную грамотность без прямой связи с технологиями. Например, Information Literacy Test (ILT) [Cameron, Wise, Lottridge, 2007] опирается на стандарты информационной грамотности [American Library Association, 2000], с его помощью можно оценить запас знаний о том, как работать с информацией. Аналогичным образом сконструированы инструменты, измеряющие близкий по смыслу навык — медиаграмотность [Brown, 1998; Livingstone, Thumim, 2003], например разработанные для Евросоюза и ЮНЕСКО инструменты, основанные на обобщенных критери-

² Федеральный государственный образовательный стандарт основного общего образования РФ. <http://минобрнауки.рф/документы/938>

³ Примерная основная образовательная программа основного общего образования, одобренная решением федерального учебно-методического объединения по общему образованию (протокол от 8 апреля 2015 г. № 1/15). <http://fgosreestr.ru/registry/primernaya-osnovnaya-obrazovatelnaaya-programma-osnovnogo-obshhego-obrazovaniya-3/>

⁴ Отдельную категорию составляют инструменты, оценивающие ИКТ и близкие конструкты через самооценку. Поскольку самооценка может быть смещена, инструменты такого рода обычно обладают низкой валидностью в отношении измеряемого навыка.

⁵ <http://www.ecdl.com/>, <http://www.ecdlrussia.org>

ях [Celot, Tornero, 2009] либо позволяющие получить результаты только на уровне целой страны [UNESCO Communication and Information Sector, UNESCO Institute for Statistics, 2013].

Третья группа тестов смешивает в оценке технические и информационные компетенции. Пионерская Программа по измерению ИКТ-компетентности для международного исследования PISA [Educational Testing Service, 2002a; van Joelingen, 2004], в рамках которой был создан инструмент, прошедший пилотные испытания в США, Австралии и Японии, не была реализована из-за недостатков измерительного инструментария: показатели валидности теста оказались неудовлетворительными и он не обеспечивал возможность межстрановых сравнений [Lennon et al., 2003]. Такая же участь постигла сконструированный аналогичным образом национальный тест в Великобритании, Qualification Curriculum Authority Key Stage 3 ICT Test [Neil, 2007].

В 2005 г. представила национальное исследование ИКТ-компетентности для средней школы австралийская группа исследователей [Ainley, Fraillon, Freeman, 2007]. На его основе был разработан инструмент для международных сравнений компьютерной и информационной грамотности International Computer and Information Literacy Study [Fraillon, Ainley, 2010; Fraillon, Schulz, Ainley, 2013]. В отдельных странах появились также тесты для начальной школы: в Корее [Kim, Kil, Shin, 2014] и в Бельгии [Aesaert et al., 2014]. Во всех этих инструментах авторы сознательно совмещают когнитивные и технические навыки в одной шкале измерения.

Четвертая группа — это инструменты для измерения ИКК без включения технических навыков. Одним из немногих примеров в этой категории является тест iCritical Thinking (изначально назывался iSkills [Katz, 2013; Katz, Macklin, 2007]). Он предназначен для измерения ИКК в цифровой среде у студентов американских вузов.

Итак, за последние 20 лет наработан определенный опыт измерения ИКК и достигнут некоторый прогресс в понимании сущности ИКК и приемлемых способов ее измерения. Однако, сталкиваясь с необходимостью измерения ИКК у выпускников средней школы, мы обнаруживаем, что в нашем распоряжении есть либо тесты на техническую грамотность, либо тесты на других языках, предназначенные для других целевых групп (студенты вузов, начальная школа, взрослые и т. п.). При этом примерная основная образовательная программа РФ рекомендует оценивать ИКК в рамках тестирования метапредметных результатов⁶.

⁶ Примерная основная образовательная программа основного общего образования, одобренная решением Федерального учебно-методического объединения по общему образованию от 8 апреля 2015 г., прото-

При создании представляемого в данной статье инструмента по оценке ИКК был учтен опыт разработки вышеупомянутых тестов, в первую очередь мы ориентировались на опыт iCritical Thinking. Однако наш инструмент изначально разработан для другой целевой аудитории и решает другие задачи.

2. Инструмент для оценки ИКК в цифровой среде

2.1. Разработка теста

Поскольку технологии быстро развиваются и становятся все более интуитивными, мы не фокусируемся на этой составляющей ИКТ-компетенции и уделяем основное внимание информационной и коммуникационной компетентности. Согласно определению, предложенному Международным советом по ИКТ-компетентности (International ICT Literacy Panel)⁷, ИКК — это «способность использовать цифровые технологии, инструменты коммуникации и/или сети для получения доступа к информации, управления ею, интеграции, оценивания, создания и передачи информации с соблюдением этических и правовых норм для того, чтобы успешно жить и трудиться в условиях общества знаний» [Educational Testing Service, 2002b]. Таким образом, в основу модели ИКК, на которой строится Information and Communication Literacy Test, положены семь составляющих:

- определение информации — способность корректно сформулировать проблему, чтобы целенаправленно искать и обрабатывать информацию;
- доступ к информации — способность искать и находить релевантную информацию из различных источников в разных средах;
- управление информацией — способность классифицировать или организовывать информацию по существующим критериям для структурирования, размещения, сохранения информации, а также быстрого ее поиска в дальнейшем;
- оценка информации — способность составить мнение о качестве, релевантности, полезности и эффективности информации и источников ее получения;
- интеграция информации — способность интерпретировать и реструктурировать информацию, вычленять главное, сравнивать или противопоставлять сведения, полученные из нескольких источников;

кол № 1/15. <http://fgosreestr.ru/registry/primernaya-osnovnayaobrazovatel'naya-programma-osnovnogo-obshhego-obrazovaniya-3/>

⁷ В соответствии с этим определением наш инструмент, который на начальных этапах разработки назывался тест ИКТ-компетентности [Авдеева и др., 2009; Avdeeva, Zelman, 2011], позже получил более точное название.

- создание информации — способность создавать или адаптировать информацию с учетом конкретной задачи и потребностей определенной целевой группы, приводить аргументы и свидетельства, подтверждающие правильность созданной или адаптированной информации;
- коммуникация (передача) информации — способность адаптировать информацию к потребностям определенной целевой группы и направлять ее определенной целевой группе.

Мы применили систематический подход к разработке тестов (evidence centered design) — «набор взаимосвязанных процедур, призванных прояснить два вопроса: какие данные мы можем извлечь из поведения тестируемого, которые бы стали свидетельством того, что тестируемый обладает или не обладает интересующими нас компетентностями? И каким образом мы можем создать ситуацию, которая помогла бы в получении этих данных?» [Mislevy, Levy 2007]. Разработка теста в рамках этого подхода происходит «послойно»: она начинается с набора очень общих идей (например, определений и составляющих) и заканчивается спецификацией теста (конкретными правилами его разработки).

Вначале уточняется определение измеряемого конструкта (ИКК), определяется количество уровней ИКК и расписывается значение каждого из этих уровней. Затем вычленяются составляющие ИКК, т. е. семь компетенций, прописываются параметры трех возможных уровней каждой из составляющих и их связь с уровнями общего конструкта. На следующих этапах задается теоретическая рамка теста через поиск в деятельности учащихся таких наборов действий (наблюдаемых переменных), которые могли бы свидетельствовать о том или ином уровне каждой из семи отдельных компетенций — в соответствии с их определениями. И только на последнем этапе формируются сами тестовые задания: на готовую структуру наблюдаемых переменных накладываются контекст, содержание и конкретный сценарий. Главное достоинство такого подхода — его систематичность, ясность и обратимость логики. Мы всегда понимаем, как конкретное свидетельство (наблюдаемая переменная) связано с измеряемым конструктом. Наличие такой связи само по себе является источником высокой конструктивной валидности инструмента. Кроме того, такая логика разработки теста соответствует общей логике психометрического измерения — связи латентных переменных (ИКК и ее составляющих) с наблюдаемыми переменными (свидетельствами). Единицей измерения, таким образом, является не тестовое задание, а наблюдаемая переменная, их может быть разное количество в разных заданиях. Каждое задание проходит проверку валидности на выборках учащихся.

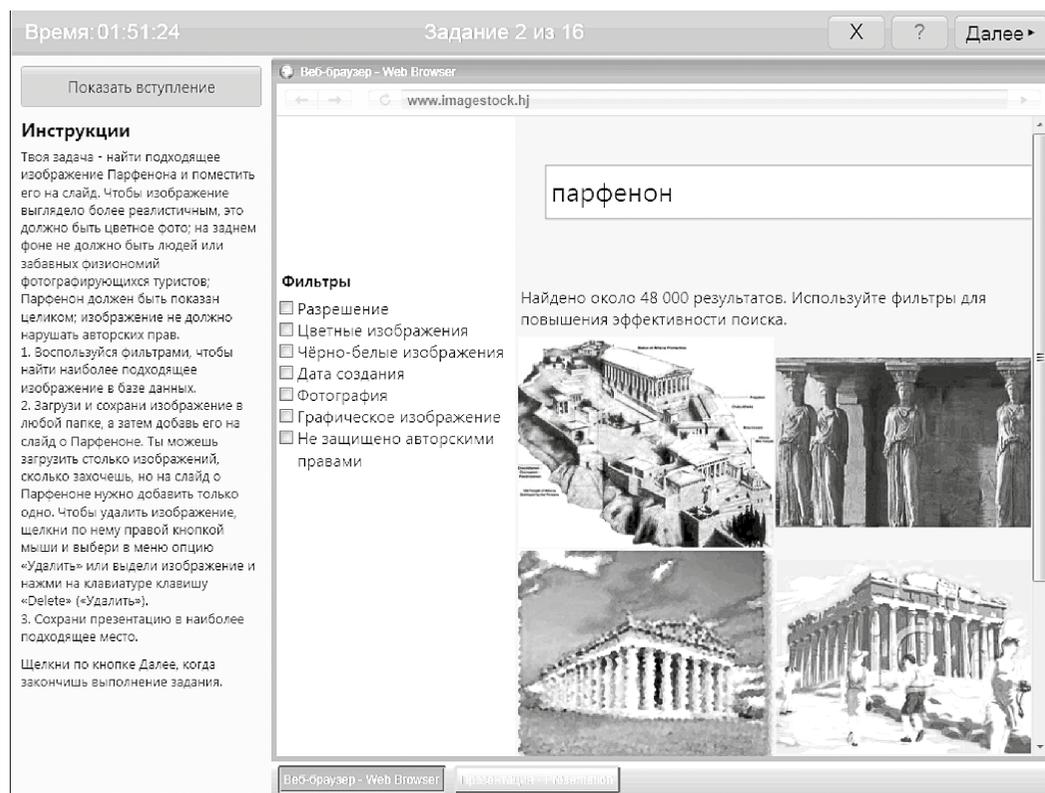
2.2. Дизайн теста ICL-тест предназначен для оценки ИКК у выпускников основной средней школы. Он включает 16 тестовых заданий сценарного типа, в том числе 13 коротких, на выполнение которых планируется не более 4 минут, два средних, которые могут занять до 15 минут, и одно продолжительное, до 30 минут. Задания предъявляются на экране компьютера последовательно, от коротких к более продолжительным, и воспроизводят реальные ситуации, с которыми учащиеся сталкиваются как в учебе, так и в повседневной жизни. Сценарии тестовых заданий разработаны таким образом, что их решение способно заинтересовать целевую аудиторию, тем самым мы решаем задачу создать у учащихся внутреннюю мотивацию к решению задачи и повышаем степень достоверности полученных результатов. В тесте соблюдается 50%-ный баланс учебных и внеучебных по контексту заданий.

Каждое задание начинается с короткого текста, описывающего некоторую ситуацию и призванного приблизить тестовую задачу к реальной жизни. На последующих экранах представлено само задание (рис. 1): в левой части экрана сформулирована конечная цель задания и даны все необходимые инструкции, справа представлен материал задания и инструменты в виде имитации широко распространенных программ или веб-сайтов, таких как текстовый редактор, социальная сеть и т. д. Специальных навыков и знаний о конкретных компьютерных программах для выполнения задания не требуется, за исключением базовых, но и их применение описано в разделе «Помощь». В ходе выполнения задания в зависимости от его длительности могут использоваться одна или несколько имитаций программ с возможностью переключения между ними.

Действия учащихся и комбинации действий записываются в базу данных. Обработка действий учащихся происходит автоматически, и по завершении выполнения всех заданий или по истечении времени тестирования учащемуся предъявляется экран с его результатом. Если показатель выполнения теста не наивысший, он сопровождается рекомендациями о том, какие компоненты ИКК у учащегося наименее развиты и требуют совершенствования.

Перед тестированием администратор дает устную инструкцию, демонстрирует экран с ключевыми правилами тестирования и предлагает выполнить три простейших задания для ознакомления с интерфейсом теста. Как правило, по завершении тестирования, прежде чем предъявить предварительные результаты, он предлагает учащимся заполнить анкету с вопросами об учебе и использовании компьютеров дома и в школе. Полученные с ее помощью сведения необходимы для исследования факторов формирования ИКК. Также анкета может включать различные методики для оценки валидности теста (см. ниже).

Рис. 1. Презентация тестового задания



ICL-тест может администрироваться несколькими способами: онлайн, через локальный сервер и с флеш-накопителей. Вид администрирования не влияет на результаты тестирования, поэтому данные, полученные разными способами, можно считать полностью эквивалентными.

ICL-тест состоит из 16 заданий, однако оценка уровня ИКК респондента происходит при помощи существенно большего числа индикаторов, скрытых от него. Эти индикаторы называются наблюдаемыми переменными, их общее число составляет 60–67 в зависимости от того, какие задания используются в конкретном варианте теста. Каждое задание включает от 3 до 7 наблюдаемых переменных, которые оценивают те или иные аспекты выполнения задания. Каждая наблюдаемая переменная имеет три уровня, характеризующие низкое, среднее или высокое соответствие деятельности респондента заданным критериям. Действия учащихся оцениваются по семи составляющим ИКК, каждая из которых имеет также три уровня: неприемле-

2.3. Система обработки ответов

мый, приемлемый и компетентный. Некоторые задания позволяют получить данные о нескольких составляющих ИКК.

Наиболее распространенная модель оценки результатов тестов — современная теория тестов (Item Response Theory, IRT [Hambleton, Swaminathan, Jane Rogers, 1991]) — по нескольким причинам непригодна для наших задач. Во-первых, латентный конструкт является многомерным: у ИКК семь взаимосвязанных составляющих. Во-вторых, поскольку измеряемые переменные сгруппированы в задания, проблема локальной зависимости заложена в самом дизайне теста. В рамках IRT есть модели для работы с многомерными конструктами (многомерные модели), есть также и модели, позволяющие контролировать локальную зависимость (фасеточные модели). Однако в случае, когда данные и многомерны, и имеют группы локально зависимых заданий, IRT не позволяет производить полностью корректный расчет результатов [Wang, Wilson, 2005]. Наконец, все латентные переменные в нашем случае являются дискретными, в то время как IRT направлена на получение непрерывной латентной переменной.

Чтобы решить перечисленные проблемы, для оценки результатов теста было решено использовать байесовские сети [Mislevy, Levy, 2007]. Простая сеть, в которой одна латентная переменная, а отношения между ней и каждой из дихотомических наблюдаемых переменных можно описать логистической функцией, будет эквивалентна типичной IRT-модели. В байесовских сетях можно дополнять эти взаимосвязи любыми другими, например между наблюдаемыми переменными, чтобы отразить локальную зависимость. Важным преимуществом байесовских сетей является то, что они хорошо сочетаются с систематическим подходом к созданию тестов, поскольку в них отражены причинно-следственные отношения, а значит, легко проследить цепочку от свидетельства (наблюдаемой переменной) до конкретного уровня измеряемого конструкта (ИКК и ее составляющих) [Ibid.].

Для описания связей между переменными мы воспользовались логистической функцией Дибелло — Самеджима [Mislevy et al., 2002]. Поскольку качество составляющих ИКК характеризуется тремя уровнями, требуются две ступени правильности: переход от низкого уровня к среднему и от среднего к высокому. После апробаций и пилотных тестирований меняются не присвоенные параметры модели, а сами задания. Так как все наблюдаемые переменные одинаково важны, параметры дискриминативности фиксируются равными 1. Параметры трудности уменьшаются или увеличиваются в зависимости от уровня подготовленности данной группы учащихся в целом: параметры трудности корректируются так, чтобы увеличить дисперсию итоговых баллов по ИКК внутри выборки и таким образом усилить предсказательную валидность результатов тестирования. Например, для выборки девятиклассников в Москве мы исполь-

зовали трудность –1 и 0, а для менее подготовленной выборки девятиклассников в Республике Армения трудность была снижена до –0,5 и 1. Таким образом, калибровка параметров проходит в экспертном, ручном режиме. Однако в планах развития теста заложена автоматизированная калибровка на основе цепей Маркова и алгоритма Монте-Карло [Mislevy et al., 2002].

При анализе результатов тестирования мы опирались на определение валидности, сформулированное в Стандартах образовательного и психологического тестирования: «Валидность — характеристика того, насколько эмпирические данные и теоретические построения поддерживают адекватность интерпретации результатов измерения» [American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999. P. 20].

Для исследования валидности инструмента было проведено несколько процедур. Одним из источников валидности является систематический подход к разработке теста. Поскольку тест разрабатывался на основе обобщенных стандартов, необходимо показать, насколько он соответствует российским и международным образовательным стандартам, а также возрастным нормам чтения. Затем мы обратимся к традиционным методам проверки валидности через сравнение результатов ICL-теста с внешними переменными.

Согласованность теста с образовательными стандартами определялась с помощью специальной процедуры, разработанной Н. Веббом [Webb, 1997]. Под согласованностью подразумевается степень совпадения прогнозируемых результатов освоения образовательной программы и их оценок, полученных с помощью тестового инструментария, а также степень их соответствия друг другу. Согласованность по сути свидетельствует о содержательной валидности теста.

Процедура определения согласованности по Веббу заключается в сопоставлении каждого из образовательных стандартов с соответствующими частями теста по нескольким критериям: содержательная согласованность; понимание прогресса; равенство и справедливость; педагогические выводы; применимость к системе образования. Здесь мы сосредоточимся на содержательной согласованности⁸. Результаты ее оценки представляют собой долю стандартов, которым было найдено соответствие в тесте. В тех нередких случаях, когда образовательные стан-

3. Исследование валидности ICL-теста

3.1. Согласованность с образовательными стандартами

⁸ Один из критериев содержательной согласованности — «соответствие по глубине знаний» здесь опущен, так как работа над ним еще не окончена.

дарты носят нарративный характер, конкретную долю вычислить не представляется возможным и приходится пользоваться качественными критериями и основываться на суждениях экспертов и разработчиков. В этом случае результатом оценки выступает одно из возможных суждений: «слабая», «приемлемая» или «полная» согласованность.

Тест ИКК проверялся на согласованность с тремя видами образовательных стандартов: Стандарты ИКК для высшего образования Американской библиотечной ассоциации — на них, в частности, основывалась рамка теста [American Library Association, 2000], общие стандарты тестирования «Качество и справедливость» [Educational Testing Service, 2002b] и часть нового российского Федерального государственного образовательного стандарта основного общего образования 2010 г., касающаяся ИКТ, а также его уточнение в виде Примерной основной образовательной программы образовательного учреждения [Кондаков, Кезина, 2013]. Со стороны теста в сравнении участвуют теоретическая рамка, спецификация, банк свидетельств (наблюдаемых переменных) и тексты заданий.

В табл. 2 представлены итоговые результаты исследования согласованности⁹. По большинству критериев согласованность полная или приемлемая. Слабое совпадение знаниевого диапазона объясняется тем, что данный критерий предполагает сравнение конкретного варианта теста со всеми стандартами. Разумеется, один вариант теста не может соответствовать всем перечисленным стандартам в силу ограниченности по времени и количеству заданий. Полный же банк заданий покрывает до 85% индикаторов в стандартах. Большую часть расхождений по другим критериям можно объяснить несоответствием целевой группы: в стандартах Американской библиотечной ассоциации это студенты, в тесте для оценки ИКК — учащиеся средней школы.

В соответствии с Примерной основной образовательной программой образовательного учреждения, разработанной в рамках требований ФГОС, ИКТ-компетенция включает не только умение учащегося работать с информацией, но и навыки работы с определенным оборудованием (принтером, сканером и т. п.), программным обеспечением, в то время как в рамке ICL-теста техническая компетентность отсутствует. Именно этим объясняется неполная согласованность стандартов и теста. При этом найденные уровни согласованности достаточно высоки по всем критериям.

⁹ Полная версия исследования согласованности, в которой подробно изложены все критерии и представлены таблицы сравнения, на основе которых высчитывались проценты и принималось решение об уровне, доступна по запросу у авторов статьи.

Таблица 2. **Результаты исследования согласованности ICL-теста с образовательными стандартами** (доля согласованных единиц и уровень согласованности)

Критерий Вебба	Единица согласования	Американские/ международные стандарты	Российские стандарты
Соответствие общего измеряемого понятия (structure of knowledge comparability)	Определение конструкта (ИКК, ИКТ-компетентность)	66% Приемлемая	Приемлемая
Соответствие измеряемых категорий (categorical congruence)	Названия разделов/рубрик, компетенций	100% Полная	75–90% Приемлемая
Совпадение знаниевого диапазона (range of knowledge correspondence)	Индикаторы, конкретные образовательные результаты и наблюдаемые переменные в конкретных вариантах теста	39–45% Слабая	47–51% Приемлемая
Баланс в подаче информации (balance of representation)	Части, иерархизированные образовательные результаты	100% Полная	Приемлемая
Диспозиционная согласованность (dispositional consonance)	Индикаторы/образовательные результаты, связанные с этическими аспектами	85% Полная	100% Полная
Равенство и справедливость (equity and fairness)	Доступность знаний и теста для уязвимых категорий учащихся (с ограниченными возможностями и т. д.)	Приемлемая*	Полная**
Использование технических средств, материалов и инструментов	Перечисление технических средств	100% Полная	Приемлемая

* Согласованность с общими стандартами тестирования «Качество и справедливость» [Educational Testing Service, 2002b].

** Согласованность с Федеральным законом РФ «Об образовании».

Согласованность по критерию равенства и справедливости основывалась на других документах: «Качество и справедливость» [Educational Testing Service, 2002b] и Федеральный закон от 29 декабря 2012 г. № 273-ФЗ «Об образовании в Российской Федерации»¹⁰. Неполная согласованность со стандартом ETS объясняется тем, что тест не подходит для учащихся со слабым зрением. В остальном же соответствие полное. Одним из подтверждений этого соответствия является анализ дифференцированного функционирования заданий, описанный ниже.

Итак, ICL-тест в приемлемой степени согласован как с российскими, так и с международными образовательными стандартами, а следовательно, пригоден к использованию в российской системе образования и после некоторой адаптации — на международном образовательном рынке.

¹⁰ http://www.consultant.ru/document/cons_doc_LAW_140174/

Таблица 3. **Результаты анализа читабельности русского и английского вариантов ICL-теста**, % от числа тестовых заданий ($N = 29$)

	Категория	Русский	Английский
Формула легкости чтения (Flesch Reading Ease)	Текст понятен для учащихся 5-го класса	4	—
	— » — 6-го класса	28	7
	— » — 7-го класса	90	45
	— » — 8–9-го класса	100	100
	Более высокие уровни	—	—
Формула Флеша — Кинкейда (Flesch, Kincaid)	5–8-й класс	60	66
	9-й класс	94	97
	10-й класс	100	100

3.2. Читабельность тестовых заданий

Инструмент ориентирован на целевую аудиторию определенного возраста (13–15 лет), поэтому перед разработчиками стояла задача сформировать тестовые задания таким образом, чтобы у учащихся не возникало проблем с их восприятием и пониманием содержания. Читабельность зависит от количества слогов в слове, слов в предложении, предложений в абзаце и т. д. Количественная оценка этих параметров позволяет определить степень понятности текста, его соответствия возрасту и образовательному уровню (классу) аудитории, на которую рассчитан текст. Для проверки заданий ICL-теста использовалась формула Р. Флеша — Flesch Reading Ease [Flesh, 1974; 1981] которая прогнозирует легкость чтения, а также формула Флеша — Кинкейда [Kincaid, Gamble, 1977], позволяющая преобразовать показатель легкости чтения в уровень образования, необходимый для понимания оцениваемого текста. Коэффициенты для русского языка были скорректированы [Оборнева, 2006] с учетом большей, чем в английском языке, средней длины слов и предложений.

В ходе исследования были проверены два варианта теста — на русском и английском языках. В табл. 3 показана доля тестовых заданий, которую легко понимают учащиеся на том или ином уровне образования. Установлено, что тексты не менее 90% тестовых заданий соответствуют целевой возрастной группе или являются более простыми. Результаты проведенного исследования свидетельствуют о высоком уровне доступности ICL-теста для учащихся с разным уровнем читательской грамотности.

3.3. Исследование валидности при помощи внешних переменных

Для проверки конвергентной, дивергентной и текущей валидности ICL-теста использовались четыре опросника, встроенные в социальную анкету, которую учащиеся заполняли после

выполнения теста: опросник самооценки ИК-компетентности, опросник школьной успеваемости, опросник частоты обращения к навыкам ИКК, опросник технической грамотности.

Анализ проводился на основе данных, собранных в ходе пилотных исследований в нескольких странах и нескольких регионах России. Пилотное тестирование в населенных пунктах Армении было организовано на стихийной выборке. Операциональное тестирование выборки девятиклассников Армении, операциональное тестирование в двух регионах Белоруссии (Минской и Гомельской областях) основаны на случайных стратифицированных по типу местности выборках, которые также были кластерными. Кластером выступала школа, при этом сами школы внутри страт и классы внутри школ отбирались случайным образом. В Красноярске и Барнауле выборки были простыми кластерными со случайным отбором школ и классов внутри школ. В Москве использовалась неслучайная выборка доступных случаев. Размеры выборок указаны в табл. 4.

Посредством корреляционного анализа данные, полученные с помощью этих опросников, сопоставлялись с результатами тестирования. Общая логика этого анализа такова: если тест действительно измеряет нужный конструкт — ИКК, то его результаты должны положительно коррелировать с самооценкой ИКК, поскольку она в некоторой степени отражает реальный уровень ИКК, и с оценками в школе, поскольку ИКК как метапредметная компетентность призвана помогать в учебе, и не должны быть связаны с технической грамотностью, поскольку это один из ключевых критериев определения ИКК в отличие от схожих конструктов.

Если ICL-тест действительно измеряет указанный конструкт, его результаты должны быть умеренно и положительно связаны с оценкой учащимися своей ИК-компетентности. Согласно результатам других исследований, коэффициент корреляции между ними должен составить от 0,20 до 0,40 [Love, Hughes, 1994].

Средняя оценка по предметам в школе традиционно используется при валидации тестов, предназначенных для системы образования. Учащиеся, получившие высокие баллы по ICL-тесту, обладают когнитивными способностями, которые позволяют им хорошо учиться, — эта гипотеза вполне согласуется с нейропсихологическими и генетическими теориями общего интеллекта [Jensen, 1998]. Таким образом, мы ожидаем увидеть положительную корреляцию между образовательными достижениями и баллами по ICL-тесту. При этом слишком высокая корреляция (выше 0,9) нежелательна, так как результаты ICL-теста должны сообщать об учащихся новую информацию, а не дублировать результаты других методов оценки образовательных достижений.

Опросник частоты обращения к навыкам ИКК состоит из вопросов о том, сколько времени учащийся проводит за компьютером, занимаясь различными видами деятельности. Ожидалось, что частота работы за компьютером не будет связана с ИКК, которая является в первую очередь мерой умения работать с информацией и требует поддержки со стороны учителей, родителей, помощи сверстников. Ожидаемый коэффициент корреляции от $-0,20$ до $0,20$.

Опросник технической грамотности содержал вопросы о конкретных функциях программных пакетов и прочих технических параметрах. Инструмент оценки ИКК создавался таким образом, что он не тестирует специальные технические знания респондентов, поэтому и здесь мы ожидали слабую связь (коэффициент корреляции от $-0,10$ до $0,10$).

В табл. 4 представлены коэффициенты надежности (внутренней согласованности) дополнительных опросников. Надежность в данном случае измерена при помощи коэффициента альфа Кронбаха. Удовлетворительные для психологических тестов значения лежат в интервале от $0,70$ до $0,95$ (бóльшие значения указывают на бóльшую внутреннюю согласованность). Из полученных данных следует, что в рамках каждого из опросников вопросы измеряют один и тот же конструкт. Контроль качества дополнительных опросников проводился также посредством проверки всех шкал на согласие с моделью IRT — Graded Response Model, вопросы с неудовлетворительными параметрами не учитывались при расчете итогового балла. Установлено, что все дополнительные шкалы имеют высокое согласие с моделью, обладают высокой надежностью, элементы политомической оценки наблюдаемых переменных работают корректно.

Как видно из табл. 4, все коэффициенты корреляции находятся в пределах стандартных ошибок от ожидаемых значений. Все коэффициенты корреляции значимы на уровне $0,05$. Умеренная корреляция балла по тесту ИКК с уверенностью в собственной ИКК свидетельствует в пользу конвергентной валидности результатов теста [Campbell, Fiske, 1959]. Близкая к нулю связь с технической грамотностью и частотой обращения к навыкам ИК-компетентности подтверждает дивергентную валидность [Ibid.]. Связь с усредненной оценкой по школьным предметам позволяет судить о текущей валидности результатов теста [Cronbach, 1971].

В одной из московских школ у нас появилась возможность сравнить результаты теста ИК-компетентности с баллами ОГЭ девятиклассников. Преимущество ОГЭ как показателя школьной успеваемости заключается в том, что его результаты не подвержены искажениям, обусловленным субъективизмом оценки учителей. Выборку составили 211 школьников, исследование с помощью ICL-теста прошло за месяц до сдачи ОГЭ. Коэффи-

Таблица 4. Корреляции результатов ICL-теста с внешними переменными и характеристики шкал

	Армения		г. Барнаул	г. Красноярск	Белоруссия	г. Москва
	2013	2014	2014	2014	2014	2014
Размер выборки, учащихся	615	1834	535	511	581	218
Размер выборки, школ	32	94	25	26	32	4
Собственная оценка учениками своей ИК-компетентности (ожидаемые корреляции от 0,20 до 0,40)						
Количество вопросов	24	24	23	23	23	23
Альфа Кронбаха	0,97	0,98	0,97	0,95	0,96	0,95
Корреляция Спирмена (стандартная ошибка)	0,23 (0,04)*	0,30 (0,02)*	0,28 (0,04)*	0,20 (0,05)*	0,26 (0,04)*	0,30 (0,11)*
Частота обращения к навыкам ИК-компетентности (ожидаемые корреляции от -0,20 до 0,20)						
Количество вопросов	16	16	13	13	13	11
Альфа Кронбаха	0,91	0,87	0,81	0,81	0,81	0,76
Корреляция Спирмена (стандартная ошибка)	0,17 (0,04)*	-0,16 (0,02)*	-0,05 (0,04)	-0,10 (0,05)*	-0,13 (0,04)*	-0,04 (0,10)
Техническая грамотность (ожидаемые корреляции от -0,20 до 0,20)						
Количество вопросов	12	12	12	12	12	13
Альфа Кронбаха	0,95	0,93	0,94	0,94	0,89	0,80
Корреляция Спирмена (стандартная ошибка)	0,20 (0,04)*	0,19 (0,02)*	0,04 (0,04)*	0,10 (0,05)*	0,08 (0,04)*	0,04 (0,12)
Средний балл в школе (ожидаемые корреляции 0,20 и выше)						
Количество вопросов	13	13	12	12	14	12
Альфа Кронбаха	0,97	0,97	0,94	0,95	0,96	0,91
Корреляция Спирмена (стандартная ошибка)	0,40 (0,04)*	0,38 (0,02)*	0,40 (0,03)*	0,50 (0,04)*	0,48 (0,03)*	0,32 (0,09)*

* Коэффициенты корреляции значимы на уровне 0,05.

коэффициенты корреляции Спирмена между результатами ОГЭ и баллами по ICL-тесту очень близки к полученным на усредненных оценках по школьным предметам. Корреляция с ОГЭ по математике составила 0,34, а с ОГЭ по русскому языку — 0,37 (обе значимы на уровне 0,05). Как и в случае с усредненными баллами по школьным предметам, умеренная корреляция подтверждает текущую валидность результатов теста. При этом корреляции не настолько высоки, чтобы предполагать, что ICL-тест дублирует другие методы оценки образовательных достижений.

3.4. Анализ дифференцированного функционирования заданий Анализ дифференцированного функционирования заданий (differential item functioning, DIF) позволяет проверить, есть ли в тесте задания, различно действующие на разных подвыборках учащихся. Например, задание для мальчиков может быть проще, чем для девочек. В идеале только уровень ИКК должен влиять на балл респондента.

Значения наблюдаемых переменных были разбиты на две подгруппы по полу, для каждой подгруппы отдельно были рассчитаны IRT-параметры трудности наблюдаемых переменных в рамках одной модели. Затем задача состояла в простом сравнении этих параметров между двумя подгруппами. По классификации Educational Testing Service [Zwick, Thayer, Lewis, 1999] разница трудности в 0,64 логита¹¹ и выше указывает на серьезную дифференцированность функционирования, а разница в 0,32–0,64 логита означает среднюю степень дифференцированности функционирования в данных подгруппах.

По результатам анализа на выборках армянских и московских школьников количество наблюдаемых переменных, демонстрирующих разницу в функционировании, не превысило 13%. При этом серьезное дифференцированное функционирование показали не более 4% наблюдаемых переменных. Таким образом, ни юноши, ни девушки не имеют преимущества при выполнении теста.

3.5. Надежность Необходимой составляющей валидности результатов тестирования является надежность. Чаще всего в качестве меры надежности используют внутреннюю согласованность заданий теста. Наиболее популярный показатель внутренней согласованности теста — альфа Кронбаха. Альтернативным подходом к оценке внутренней согласованности является метод расщепления с коррекцией по формуле Спирмена — Брауна.

Как видно из табл. 5, оба коэффициента внутренней согласованности высоки, но не слишком. Так как ICL-тест измеряет довольно широкий круг компетенций, слишком высокие значения (около 1) были бы признаком того, что многие задания измеряют одно и то же. Значения 0,74–0,84 показывают, что задания теста измеряют разные аспекты одного общего конструкта.

В одном из недавних администрирований нам удалось протестировать одну и ту же выборку учащихся дважды — в начале 9-го класса и на момент окончания 10-го. Исследование проводилось в Ямало-Ненецком автономном округе, размер выбор-

¹¹ Логит — единица измерения уровней подготовленности участников тестирования и трудностей тестовых заданий в рамках логистических моделей тестирования. Если разность между упомянутыми двумя понятиями составляет 1 логит, то вероятность верного выполнения таким испытуемым такого задания равна 0,73.

Таблица 5. Коэффициенты внутренней согласованности заданий теста

	Армения		г. Барнаул	г. Красноярск	Белоруссия	г. Москва
	2013	2014	2014	2014	2014	2014
Альфа Кронбаха	0,95 / 0,84	0,82	0,83	0,78	0,83	0,78
Расщепление с коррекцией	0,97 / 0,80	0,77	0,77	0,77	0,77	0,74

* Вычислялись отдельно по каждому из двух вариантов. Первый из вариантов впоследствии был скорректирован из-за завышенного показателя согласованности.

ки — 429 человек. Учащиеся проходили один и тот же тест с фиксированным набором заданий. Корреляция между результатами тестирования оказалась высоко значимой и составила 0,62, что указывает на надежность результатов теста во времени.

Помимо перечисленных процедур большинство пробных тестирований сопровождалось фокус-групповыми дискуссиями с учащимися и педагогами, в которых обсуждалась понятность, приемлемость, мотивирующая составляющая ICL-теста. Педагоги подтвердили необходимость и полезность данного инструмента, при этом они отмечали, что тест создает высокую нагрузку на учащихся в отношении чтения, что в дальнейшем было исправлено. Отзывы учащихся свидетельствуют о высокой очевидной валидности теста. Например, почти всем (97%) армянским школьникам, принявшим участие в тестировании ($n = 2466$, 2014 г.), понравилось проходить тест, 90% отметили, что отметка за тест справедливо отразила их умение пользоваться компьютером и интернетом, а 81% — что интерфейс инструмента понятен и прост в использовании.

3.6. Очевидная валидность

Данные, полученные с помощью ICL-теста, могут быть использованы в разных целях. Учителя могут применять их при формировании индивидуальных и групповых образовательных технологий для учащихся, ориентируясь при этом не только на общий уровень ИКК учащегося, но и на показатели по семи отдельным компетенциям. Например, если учащийся находится на базовом уровне ИКК в целом, но получил низкие значения по компетенции «создание» и «оценка», при составлении его индивидуальной образовательной траектории можно обратить внимание на создание и/или адаптацию имеющейся информации с учетом конкретной проблемы, а также развивать у него способность формировать собственное мнение о качестве, релевантности, полезности информации и источника ее получения. То же относится и к группе: если у группы учащихся выявлено низкое среднее значение компетенции «управление», то в обучении не-

4. Применение результатов оценки ИКК

обходимо сделать акцент на развитии способности классифицировать и организовывать информацию по заданным критериям. Для выработки более конкретных рекомендаций педагогам предлагаются детальные матрицы для каждого уровня каждой из семи компетенций и значения свидетельств (наблюдаемых переменных).

ICL-тест применяется также для мониторинга метапредметных образовательных результатов у учащихся 8–10-х классов (формирование и развитие компетентности в области использования ИКТ) в ходе освоения основной образовательной программы школы. Для целей мониторинга инструмент поддается небольшой коррекции с учетом особенностей системы образования конкретного субъекта РФ. Тест успешно использовался и в региональных лонгитюдных исследованиях по оцениванию и формированию ИКК.

Тест подходит для сопровождения внедрения в школах инновационных методик, новых педагогических практик и специальных кейсов, направленных на развитие навыков эффективной работы с информацией, решение проблем, принятие решений и достижение других метапредметных результатов (умения определять понятия, создавать обобщения, самостоятельно выбирать основания и критерии для классификации, устанавливать причинно-следственные связи, строить логическое рассуждение и делать выводы).

Результаты ICL-теста могут использоваться в качестве показателей эффективности реализации федеральных, региональных и муниципальных программ, причем направленных как на информатизацию системы образования, так и на модернизацию его содержания и внедрение современных образовательных технологий.

ICL-тест не предназначен для измерения технических навыков по работе с оборудованием и конкретным программным обеспечением и не может быть использован для оценивания знаний по информатике.

5. Заключение В данной работе мы описали основные компоненты и дизайн инструмента для измерения ИКК выпускников основной школы — ICL-теста и продемонстрировали его конструктивную, содержательную, конвергентную, дивергентную и текущую валидность и надежность.

Исследования, основанные на шести разных выборках, подтвердили ряд его преимуществ перед похожими инструментами. ICL-тест:

- оценивает способность эффективно работать с информацией и решать практические задачи, используя возможности ИКТ, а не уровень технологических навыков владения ИКТ;

- тестирует как ИКК в целом, так и ее составляющие;
- содержит тестовые задания, основанные на реальных жизненных или учебных ситуациях;
- может использоваться на любом компьютере, независимо от операционной системы и наличия выхода в интернет;
- подходит для проведения как сравнительных, так и лонгитюдных исследований;
- один из немногих тестов, оценивающих метапредметные результаты с учетом требований новых ФГОС, и пока единственный в России, делающий это в автоматическом режиме.

Получаемые с помощью инструмента данные позволяют не только зафиксировать уровень ИКК выпускников основной школы, но и выявить факторы, оказывающие на нее влияние¹², а также динамику ее формирования при проведении лонгитюдных исследований. Например, результаты тестирования, полученные в Ямало-Ненецком автономном округе, дали возможность местным органам управления образованием разработать комплекс мер, направленных на повышение уровня ИКК учащихся, усовершенствовать подходы к оценке содержания образовательной программы школ и доработать рекомендации по внедрению ФГОС. В Республике Татарстан результаты тестирования входили в перечень индикаторов эффективности при реализации проекта «Электронное образование в Республике Татарстан».

Полученные с помощью ICL-теста результаты стимулируют школы к развитию ИКК учащихся. Самостоятельная работа в классе, домашние задания с поиском информации в интернете, учебная активность в малых группах, проектная деятельность могут быть инициированы учителем в процессе обучения и способны сформировать более высокие уровни ИКК.

ICL-тест можно использовать при проведении национальных и региональных мониторинговых исследований для оценки эффективности реализации программ, направленных на развитие системы образования, особенно полезным этот инструмент будет при оценке результатов внедрения новых ФГОС, информатизации образования, при выявлении образовательных достижений.

1. Авдеева С. М., Барышникова М. Ю., Коваленко С. К., Мельников А. Е. (2009) Об опыте реализации модели оценки ИКТ-компетентности // Информатизация образования и науки. № 2. С. 62–71.

Литература

¹² В настоящее время по результатам исследований с помощью ICL-теста готовится к публикации статья, посвященная факторам, влияющим на формирование ИКК.

2. Авдеева С. М., Худенко Л. А. (2014) Исследование информационно-коммуникационной компетентности учащихся 9-х классов Республики Беларусь: цель и основные результаты // Вестник образования. № 1. С. 3–10.
3. Кондаков А. М., Кезина Л. П. (ред.) (2013) Примерная основная образовательная программа образовательного учреждения. Основная школа. М.: Институт стратегических исследований в образовании РАО.
4. Оборнева И. В. (2006) Автоматизированная оценка сложности учебных текстов на основе статистических параметров. Дисс. канд. пед. наук. М.: Институт содержания и методов обучения.
5. Abbot S. (2014) Hidden Curriculum // S. Abbott (ed.) The Glossary of Education Reform. <http://edglossary.org/hidden-curriculum>
6. American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999) The Standards for Educational and Psychological Testing. Washington, DC: American Psychological Association.
7. Aesaert K., van Nijlen D., Vanderlinde R., van Braak J. (2014) Direct Measures of Digital Information Processing and Communication Skills in Primary Education: Using Item Response Theory for the Development and Validation of an ICT Competence Scale // Computers & Education. Vol. 76. July. P. 168–181.
8. Ainley J., Fraillon J., Freeman C. (2007) National Assessment Program. ICT Literacy Years 6 & 10 Report, 2005. Sydney NSW: Australian Curriculum, Assessment and Reporting Authority.
9. American Library Association (2000) Information Literacy Competency Standards for Higher Education. Chicago, IL: The Association of College and Research Libraries.
10. Australian Library and Information Association, Information Literacy Forum (2001) Statement on Information Literacy for all Australians. Kingston: Australian Library and Information Association. <http://www.alia.org.au/groups/infolit/information.literacy.html>
11. Avdeeva S., Zelman M. (2011) ICT Competence Assessment Instrument and Its Use / The 36th Annual Conference of the International Association for Educational Assessment (IAEA) «Assessment for the Future Generations». Bangkok (Thailand), 2010, 20–27 August.
12. Blömeke S. (2000) Medienpädagogische Kompetenz: theoretische und empirische Fundierung eines zentralen Elements der Lehrerbildung. München: KoPäd-Verlag.
13. Brown J. (1998) Media Literacy Perspectives // Journal of Communication. No 48. Iss. 1. P. 44–57.
14. Cameron L., Wise S. L., Lottridge S. M. (2007) The Development and Validation of the Information Literacy Test // College & Research Libraries. Vol. 68. No 3. P. 229–237.
15. Campbell D. T., Fiske D. W. (1959) Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix // Psychological Bulletin. Vol. 56. No 2. P. 81–105.
16. Celot P., Tornero J. M. P. (eds) (2009) Study on Assessment Criteria for Media Literacy Levels. A Comprehensive View of the Concept of Media Literacy and Understanding of How Media Literacy Levels in Europe Should Be Assessed. Brussels: European Association for Viewers' Interests.
17. Cronbach L. J. (1971) Test Validation // R. L. Thorndike (ed.) Educational Measurement. Washington, DC: American Council on Education. P. 443–507.

18. Educational Testing Service (2002a) Digital Transformation for ICT Literacy: A Report of the International ICT Literacy Panel. http://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf
19. Educational Testing Service (2002b) ETS Standards for Equality and Fairness. Princeton, NJ: ETS.
20. Flesch R. (1974) The Art of Readable Writing. New York: Harper & Row.
21. Flesch R. (1981) Why Johnny Still Can't Read: A New Look at the Scandal of Our Schools. New York: Harper & Row.
22. Fraillon J., Ainley J. (2010) The IEA International Study of Computer and Information Literacy (ICILS). <http://icils.acer.edu.au>
23. Fraillon J., Schulz W., Ainley J. (2013) International Computer and Information Literacy Study: Assessment Framework. Amsterdam: IEA.
24. Hambleton R. K., Swaminathan H., Jane Rogers H. (1991) Fundamentals of Item Response Theory. Newbury Park, CA: SAGE.
25. Hepworth M. (2000) The Challenge of Incorporating Information Literacy into the Undergraduate Curriculum // S. Corral, H. Hathaway (eds) Seven Pillars of Wisdom? Good Practice in Information Skills Development. London: SCONUL. P. 11–21.
26. Huggins A. C., Ritzhaupt A. D., Dawson K. (2014) Measuring Information and Communication Technology Literacy Using a Performance Assessment: Validation of the Student Tool for Technology Literacy (ST 2 L) // Computers & Education. Vol. 77. August. P. 1–12.
27. Jenkins H. (2009) Confronting the Challenges of Participatory Culture: Media Education for the 21st Century. Cambridge, MA; London: MIT Press.
28. Jensen A. R. (1998) The g Factor: The Science of Mental Ability. Westport, CT: Praeger.
29. Johnston B., Webber S. (2003) Information Literacy in Higher Education: A Review and Case Study // Studies in Higher Education. Vol. 28. Iss. 3. P. 335–352.
30. Katz I. R. (2013) Testing Information Literacy in Digital Environments: ETS's iSkills Assessment // Information Technology and Libraries. Vol. 26. No 3. P. 3–12.
31. Katz I. R., Macklin A. S. (2007) Information and Communication Technology (ICT) Literacy: Integration and Assessment in Higher Education // Systemics, Cybernetics and Informatics. Vol. 5. No 4. P. 50–55. [http://www.iiisci.org/Journal/CV\\$/sci/pdfs/P890541.pdf](http://www.iiisci.org/Journal/CV$/sci/pdfs/P890541.pdf)
32. Kim H.-S., Kil H.-J., Shin A. (2014) An Analysis of Variables Affecting the ICT Literacy Level of Korean Elementary School Students // Computers & Education. Vol. 77. August. P. 29–38.
33. Kincaid J. P., Gamble L. G. (1977) Ease of Comprehension of Standard and Readable Insurance Policies as a Function of Reading Ability // Journal of Reading Behavior. Vol. 9. No 1. P. 87–95.
34. Lennon M., Kirsch I., von Davier M., Wagner M., Yamamoto K. (2003) Feasibility Study for the PISA ICT Literacy Assessment. <http://www.oecd.org/dataoecd/35/13/33699866.pdf>
35. Livingstone S., Thumim N. (2003) Assessing the Media Literacy of UK Adults. A Review of Academic Literature Produced for Broadcasting Standards Commission Independent Television Commission NIACE. http://www.ofcom.org.uk/static/archive/itc/uploads/Assessing_the_media_literacy_of_UK_adults1.pdf
36. Love K. G., Hughes F. V. (1994) Relationship of Self-Assessment Ratings and Written Test Score: Implications for Law Enforcement Promotional Systems // Public Personnel Management. Vol. 23. No 1. P. 19–30.

37. Malta Qualifications Council (2007) Descriptors of Key Competences in the National Qualification Framework. Competences of Personal Development. <http://cpd.yolasite.com/key-competences.php>
38. Mislevy R. J., Almond R., Dibello L., Jenkins F., Steinberg, L. (2002) Modeling Conditional Probabilities in Complex Educational Assessments. CSE Technical Report. Los Angeles: The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA. <http://www.cse.ucla.edu/CRESST/Reports/TR580.pdf>
39. Mislevy R. J., Levy R. (2007) Bayesian Psychometric Modeling from an Evidence-Centered Design Perspective // C. R. Rao, S. Sinharay (eds) Handbook of Statistics. Amsterdam: Elsevier. P. 839–865.
40. Neil A., (2007) KS3 Online ICT Test — Members' Responses to the Change in Status. <http://www.naace.co.uk/212>
41. Rader H. B. (2002) Information Literacy 1973–2002: A Selected Literature Review // Library Trends. Vol. 51. No 2. P. 242–259.
42. The European Parliament, the Council (2006) Lifelong Learning — Key Competences. Recommendation (2006/962/EC). <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV:c11090>
43. UNESCO Communication and Information Sector, UNESCO Institute for Statistics (2013) Global Media and Information Literacy Assessment Framework: Country Readiness and Competencies. Paris: United Nations Educational, Scientific and Cultural Organization.
44. Van Joolingen W. (2004) The PISA Framework for Assessment of ICT Literacy. <http://www.ictliteracy.info/rf.pdf/PISA%20framework.ppt>
45. Virkus S. (2003) Information Literacy in Europe: A Literature Review // Information Research. Vol. 8. No 4. Paper no 159.
46. Wang W. C., Wilson M. (2005) The Rasch Testlet Model // Applied Psychological Measurement. Vol. 29. No 2. P. 126–149.
47. Webb N. L. (1997) Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education. Research monograph no 8. Washington, DC: Council of Chief State School Officers.
48. World Economic Forum (2015) New Vision for Education Unlocking the Potential of Technology. Geneva: World Economic Forum.
49. Zwirk R., Thayer D. T., Lewis C. (1999) An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis // Journal of Educational Measurement. Vol. 36. No 1. P. 1–28.

Assessing Information and Communication Technology Competence of Students: Approaches, Tools, Validity and Reliability of Results

Svetlana Avdeeva

PhD in Technical Sciences, Deputy Executive Director, National Training Foundation; Deputy Head, Center of Education Quality Monitoring, Institute of Education, National Research University Higher School of Economics. Address: Bld. 1, 1905 Goda Str., 123022 Moscow, Russian Federation. E-mail: avdeeva@ntf.ru

Authors

Maksim Rudnev

Candidate of Sciences in Sociology, Leading Researcher, Laboratory for Comparative Studies in Mass Consciousness, National Research University Higher School of Economics. Address: 20 Myasnitskaya Str., 101000 Moscow, Russian Federation. E-mail: mrudnev@hse.ru

Georgy Vasin

Psychometrician, Department of ICT Development in Education, National Training Foundation. Address: Bld. 1, 71905 Goda Str., 123022 Moscow, Russian Federation. E-mail: vasin@ntf.ru

Ksenia Tarasova

Candidate of Sciences in Pedagogy, Leading Specialist, National Training Foundation. Address: Bld. 1, 1905 Goda Str., 123022 Moscow, Russian Federation. E-mail: kvtarasova@mail.ru

Daria Panova

Assessment Content Developer, National Training Foundation. Address: Bld. 1, 1905 Goda Str., 123022 Moscow, Russian Federation. Email: panova@ntf.ru

The study describes the Information and Communication Literacy Test (the ICL Test), an instrument to measure information and communication technology competence in middle school students (grades 7–9). An overview of the existing instruments proves that there are no other ICT competence assessment instruments applicable to Russian students today. The assessment was constructed using an innovative systematic approach (evidence-centered assessment design, ECD). The ICL Test results were calculated using Bayesian networks, which are more effective than the Item Response Theory in this case. The ICL Test revealed a high level of construct, content, convergent, divergent and face validity as well as good reliability. The instrument conforms to both Russian and international educational standards according to Webb's criteria for alignment. The ICL Test results were used to assess ICT competence and identify the factors that influence its development in regions of Russia, Armenia and Belarus. The conclusion is that the ICL Test can be applied to both Russian and foreign education systems.

Abstract

information and communication technology competence, tests, Information and Communication Literacy Test, validity, competence, competence assessment, cross-disciplinary educational outcomes.

Keywords

Abbot S. (2014) Hidden Curriculum. *The Glossary of Education Reform* (ed. S. Abbott). Available at: <http://edglossary.org/hidden-curriculum> (accessed 5 November 2017).

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999) *The Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.
- Aesaert K., van Nijlen D., Vanderlinde R., van Braak J. (2014) Direct Measures of Digital Information Processing and Communication Skills in Primary Education: Using Item Response Theory for the Development and Validation of an ICT Competence Scale. *Computers & Education*, vol. 76, July, pp. 168–181.
- Ainley J., Fraillon J., Freeman C. (2007) *National Assessment Program. Literacy Years 6 & 10 Report, 2005*. Sydney NSW: Australian Curriculum, Assessment and Reporting Authority.
- American Library Association (2000) *Information Literacy Competency Standards for Higher Education*. Chicago, IL: The Association of College and Research Libraries.
- Australian Library and Information Association, Information Literacy Forum (2001) *Statement on Information Literacy for all Australians*. Kingston: Australian Library and Information Association. Available at: <http://www.alia.org.au/groups/infolit/information.literacy.html> (accessed 5 November 2017).
- Avdeeva S., Baryshnikova M., Kovalenko S., Melnikov A. (2009) Ob opyte realizatsii modeli otsenki IKT-kompetentnosti [On the Experience of ICT-Competence Evaluation Model Realization]. *Informatizatsiya obrazovaniya i nauki*, no 2, pp. 62–71.
- Avdeeva S., Khudenko L. (2014) Issledovanie informatsionno-kommunikatsionnoy kompetentnosti uchashchikhsya 9-kh klassov Respubliki Belarus': tsel' i osnovnye rezul'taty [Research on the Information and Communication Technology Competence of Ninth-Graders in the Republic of Belarus: The Goal and the Main Findings]. *Vestnik obrazovaniya*, no 1, pp. 3–10.
- Avdeeva S., Zelman M. (2011) *Competence Assessment Instrument and Its Use*. Paper presented at the 36th Annual Conference of the International Association for Educational Assessment (IAEA) "Assessment for the Future Generations". Bangkok (Thailand), 2010, 20–27 August.
- Blömeke S. (2000) *Medienpädagogische Kompetenz: theoretische und empirische Fundierung eines zentralen Elements der Lehrerbildung*. München: KoPäd-Verlag.
- Brown J. (1998) Media Literacy Perspectives. *Journal of Communication*, no 48, iss. 1, pp. 44–57.
- Cameron L., Wise S. L., Lottridge S. M. (2007) The Development and Validation of the Information Literacy Test. *College & Research Libraries*, vol. 68, no 3, pp. 229–237.
- Campbell D. T., Fiske D. W. (1959) Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, vol. 56, no 2, pp. 81–105.
- Celot P., Tornero J. M. P. (eds) (2009) *Study on Assessment Criteria for Media Literacy Levels. A Comprehensive View of the Concept of Media Literacy and Understanding of How Media Literacy Levels in Europe Should Be Assessed*. Brussels: European Association for Viewers' Interests.
- Cronbach L. J. (1971) Test Validation. *Educational Measurement* (ed. R. L. Thorndike), Washington, DC: American Council on Education, pp. 443–507.
- Educational Testing Service (2002a) *Digital Transformation for Literacy: A Report of the International ICT Literacy Panel*. Available at: http://www.ets.org/Media/Tests/Information_and_Communication_Technology_Literacy/ictreport.pdf (accessed 5 November 2017).

- Educational Testing Service (2002b) *ETS Standards for Equality and Fairness*. Princeton, NJ: ETS.
- Flesch R. (1974) *The Art of Readable Writing*. New York: Harper & Row.
- Flesch R. (1981) *Why Johnny Still Can't Read: A New Look at the Scandal of Our Schools*. New York: Harper & Row.
- Fraillon J., Ainley J. (2010) *The IEA International Study of Computer and Information Literacy (ICILS)*. Available at: <http://icils.acer.edu.au> (accessed 5 November 2017).
- Fraillon J., Schulz W., Ainley J. (2013) *International Computer and Information Literacy Study: Assessment Framework*. Amsterdam: IEA.
- Hambleton R. K., Swaminathan H., Jane Rogers H. (1991) *Fundamentals of Item Response Theory*. Newbury Park, CA: SAGE.
- Hepworth M. (2000) The Challenge of Incorporating Information Literacy into the Undergraduate Curriculum. Seven Pillars of Wisdom? Good Practice in Information Skills Development (eds S. Corral, H. Hathaway), London: SCONUL, pp. 11–21.
- Huggins A. C., Ritzhaupt A. D., Dawson K. (2014) Measuring Information and Communication Technology Literacy Using a Performance Assessment: Validation of the Student Tool for Technology Literacy (ST 2 L). *Computers & Education*, vol. 77, August, pp. 1–12.
- Jenkins H. (2009) *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. Cambridge, MA; London: MIT Press.
- Jensen A. R. (1998) *The g Factor: The Science of Mental Ability*. Westport, CT: Praeger.
- Johnston B., Webber S. (2003) Information Literacy in Higher Education: A Review and Case Study. *Studies in Higher Education*, vol. 28, iss. 3, pp. 335–352.
- Katz I. R. (2013) Testing Information Literacy in Digital Environments: ETS's iSkills Assessment. *Information Technology and Libraries*, vol. 26, no 3, pp. 3–12.
- Katz I. R., Macklin A. S. (2007) Information and Communication Technology (ICT) Literacy: Integration and Assessment in Higher Education. *Systemics, Cybernetics and Informatics*, vol. 5, no 4, pp. 50–55. Available at: <http://www.iiisci.org/Journal/CV%2Fsci/pdfs/P890541.pdf> (accessed 5 November 2017).
- Kim H.-S., Kil H.-J., Shin A. (2014) An Analysis of Variables Affecting the ICT Literacy Level of Korean Elementary School Students. *Computers & Education*, vol. 77, August, pp. 29–38.
- Kincaid J. P., Gamble L. G. (1977) Ease of Comprehension of Standard and Readable Insurance Policies as a Function of Reading Ability. *Journal of Reading Behavior*, vol. 9, no 1, pp. 87–95.
- Kondakov A., Kezina L. (eds) (2013) *Primernaya osnovnaya obrazovatel'naya programma obrazovatel'nogo uchrezhdeniya. Osnovnaya shkola* [Approximate Basic Educational Program of Educational Institution. General Secondary Education], Moscow: Institute for Strategy of Education Development of the Russian Academy of Education.
- Lennon M., Kirsch I., von Davier M., Wagner M., Yamamoto K. (2003) *Feasibility Study for the PISA ICT Literacy Assessment*. Available at: <http://www.oecd.org/dataoecd/35/13/33699866.pdf> (accessed 5 November 2017).
- Livingstone S., Thumim N. (2003) *Assessing the Media Literacy of Adults. A Review of Academic Literature Produced for Broadcasting Standards Commission Independent Television Commission NIACE*. Available at: http://www.ofcom.org.uk/static/archive/itc/uploads/Assessing_the_media_literacy_of_UK_adults1.pdf (accessed 5 November 2017).

- Love K. G., Hughes F. V. (1994) Relationship of Self-Assessment Ratings and Written Test Score: Implications for Law Enforcement Promotional Systems. *Public Personnel Management*, vol. 23, no 1, pp. 19–30.
- Malta Qualifications Council (2007) *Descriptors of Key Competences in the National Qualification Framework. Competences of Personal Development*. Available at: <http://cpd.yolasite.com/key-competences.php> (accessed 5 November 2017).
- Mislevy R. J., Almond R., Dibello L., Jenkins F., Steinberg, L. (2002) *Modeling Conditional Probabilities in Complex Educational Assessments. CSE Technical Report*. Los Angeles: The National Center for Research on Evaluation, Standards, Student Testing (CRESST), Center for Studies in Education, UCLA. Available at: <http://www.cse.ucla.edu/CRESST/Reports/TR580.pdf> (accessed 5 November 2017).
- Mislevy R. J., Levy R. (2007) Bayesian Psychometric Modeling from an Evidence-Centered Design Perspective. *Handbook of Statistics* (eds C. R. Rao, S. Sinharay), Amsterdam: Elsevier, pp. 839–865.
- Neil A., (2007) *KS3 Online Test – Members’ Responses to the Change in Status*. Available at: <http://www.naace.co.uk/212> (accessed 5 November 2017).
- Oborneva I. (2006) *Avtomatizirovannaya otsenka slozhnosti uchebnykh tekstov na osnove statisticheskikh parametrov* [Automated Evaluation of Educational Text Complexity Based on Statistical Parameters] (PhD Thesis). Moscow: Institute of Education Content and Teaching Methods.
- Rader H. B. (2002) Information Literacy 1973–2002: A Selected Literature Review. *Library Trends*, vol. 51, no 2, pp. 242–259.
- The European Parliament, the Council (2006) *Lifelong Learning—Key Competences. Recommendation (2006/962/)*. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV:c11090> (accessed 5 November 2017).
- UNESCO Communication and Information Sector, UNESCO Institute for Statistics (2013) *Global Media and Information Literacy Assessment Framework: Country Readiness and Competencies*. Paris: United Nations Educational, Scientific and Cultural Organization.
- Van Joolingen W. (2004) *The Framework for Assessment of ICT Literacy*. Available at: <http://www.ictliteracy.info/rf.pdf/PISA%20framework.ppt> (accessed 5 November 2017).
- Virkus S. (2003) Information Literacy in Europe: A Literature Review. *Information Research*, vol. 8, no 4, paper no 159.
- Wang W. C., Wilson M. (2005) The Rasch Testlet Model. *Applied Psychological Measurement*, vol. 29, no 2, pp. 126–149.
- Webb N. L. (1997) *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education. Research monograph no 8*. Washington, DC: Council of Chief State School Officers.
- World Economic Forum (2015) *New Vision for Education Unlocking the Potential of Technology*. Geneva: World Economic Forum.
- Zwick R., Thayer D. T., Lewis C. (1999) An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis. *Journal of Educational Measurement*, vol. 36, no 1, pp. 1–28.