

Создание русскоязычной версии международного инструмента оценивания ранних навыков чтения

А. Е. Иванова, К. С. Карданова-Бирюкова

Иванова Алина Евгеньевна

научный сотрудник Центра мониторинга качества образования Института образования, Национальный исследовательский университет «Высшая школа экономики». Адрес: 101000, Москва, ул. Мясницкая, 20. E-mail: aeivanova@hse.ru

Карданова-Бирюкова

Ксения Суфьяновна

кандидат филологических наук, доцент Института иностранных языков, МГПУ. Адрес: 105064, Москва, Малый Казенный пер., 5Б. E-mail: Kardanova-BirukovaKS@mgpu.ru

Аннотация. Успешная адаптация первоклассника к школе в значительной степени определяет его дальнейшие академические успехи. В нашей стране и в мире вообще отмечается недостаток качественных стандартизированных инструментов, позволяющих комплексно оценить, что знает и умеет делать ребенок на входе в школу. Особенно трудно провести массовое тестирование, что связано с возрастными характеристиками детей и с форматом оценивания. Цель статьи состоит в описании процедуры локализации для русскоязычной аудитории части международного инструмента iPIPS, разработанной для оценивания ранних навыков чтения у детей на входе в школу и измерения их прогресса за первый год обучения. Под локализацией понимается преобразование инструмента, изначально созданного на другом языке (в на-

шем случае на английском) так, чтобы в нем учитывалась культурная и языковая специфика целевой аудитории. Процедура включала разработку русскоязычной версии инструмента, а также проведение серии исследований с целью сбора свидетельств его конструктивной валидности. В ходе работ по локализации тестового инструмента были изучены лингвистические характеристики оригинальных тестовых заданий, подобраны эквивалентные средства в русском языке и смоделированы русскоязычные задания, содержательно соответствующие англоязычному оригиналу. Для доказательства конструктивной валидности локализованного инструмента осуществлено обоснование психометрического качества шкалы, проведена оценка ее надежности, подтверждено соответствие структуры и иерархичности заданий теоретической рамке. На основании полученных данных сделан вывод, что при проведении массового тестирования на муниципальном или региональном уровне благодаря использованию локализованного инструмента оценивания могут быть получены ценные данные, полезные для анализа общей ситуации и принятия обоснованных решений в рамках образовательной политики.

Ключевые слова: инструмент оценивания чтения, iPIPS, начальная школа, локализация, валидизация.

Статья поступила в редакцию в июне 2019 г.

DOI: 10.17323/1814-9545-2019-4-93-115

Начало школьного обучения является решающим периодом в познавательном и личностном развитии ребенка. Как показали многочисленные исследования, успешная адаптация первоклассника к школе в значительной степени определяет его дальнейшие академические достижения [Безруких, Филиппова, Байдина, 2012; Tymms et al., 2009]. Политики, исследователи и педагоги заинтересованы в использовании качественных инструментов оценивания, позволяющих измерить знания и навыки детей на входе в школу и оценить их прогресс по мере обучения, чтобы иметь доказательную основу для планирования учебного процесса и, в случае необходимости, для раннего вмешательства с целью корректировки образовательных траекторий учащихся.

В системах оценки качества образования используются как национальные инструменты оценивания, так и результаты международных сравнительных исследований, что позволяет повысить эффективность управления образованием в целом [Болотов и др., 2013]. Однако планирование и проведение международных крупномасштабных исследований сопряжено с многочисленными методологическими трудностями, из которых наиболее сложными являются проблемы измерения, поскольку и анализ, и интерпретация результатов исследований зависят от способности используемого инструмента измерить искомый конструкт схожим образом во всех странах-участниках. Когда таким конструктом является чтение, исследователи и разработчики инструментов из разных стран должны решить сложнейшую проблему: создать такие национальные версии инструмента оценивания, которые смогут преодолеть влияние языка страны на чтение, осуществляемое на этом языке. В настоящее время систематически проводятся два международных сравнительных исследования, направленные на оценивание навыка чтения: Международная программа по оценке образовательных достижений учащихся PISA, ориентированная на 15-летних школьников, и Международное исследование качества чтения и понимания текста PIRLS, в котором принимают участие дети, завершающие обучение в начальной школе. Но таких же масштабных исследований, которые позволили бы объективно и достоверно сравнивать в разных странах результаты оценивания учащихся, только начинающих читать, пока не существует. Реализовать подобное исследование чрезвычайно трудно, поскольку язык оказывает особенно сильное влияние именно тогда, когда дети находятся на ранних этапах развития чтения [Ercikan, Roth, Asil, 2015]. Тем не менее ученые предпринимают попытки сопоставить начальные уровни навыка чтения у детей в разных странах, например в англоговорящих [Tymms et al., 2014].

Цель данной статьи состоит в описании процедуры локализации на русский язык международного инструмента, изна-

начально созданного на английском языке для оценивания ранних навыков чтения, на примере инструмента iPIPS¹. Процедура включала разработку русскоязычной версии инструмента, а также проведение серии исследований с целью сбора свидетельств его конструктивной валидности.

Оценить знания и навыки дошкольников и детей на начальном этапе обучения в школе можно двумя способами: 1) напрямую спросить детей, что они знают и умеют делать; 2) использовать не прямые форматы оценки — наблюдение или опрос учителей (родителей). В России созданы авторские методики, нацеленные на оценку отдельных навыков ребенка и часто требующие помощи психолога в проведении и интерпретации, например методика определения уровня произвольности ребенка «Графический диктант» или методика проверки фонематического слуха «Первая буква» [Ковалева и др., 2012]. За рубежом существует ряд известных комплексных исследований, предназначенных для оценки широкого спектра компетенций ребенка на входе в школу. Так, разработанная в Канаде и применяемая сегодня в ряде других стран методика *Early Development Instrument (EDI)* оценивает физическую, социальную, эмоциональную, коммуникационную и когнитивно-языковую сферы развития детей на основании опроса педагогов дошкольных образовательных учреждений в год перед поступлением в школу [Janus, Brinkman, Duku, 2011]. Во многих странах используются инструменты оценки образовательной среды в дошкольном учреждении и в школе, например *Early Childhood Environment Rating Scales (ECERS)* и *School-Age Care Environment Rating Scales (SACERS)*. Эти инструменты основаны на наблюдении и структурированном экспертном оценивании пространства, в котором находится ребенок [Harms, 2013].

Дополнительные сложности в проведении оценивания детей возникают в случае крупномасштабных кампаний по проверке ранней грамотности, поскольку они требуют значительных ресурсов, специального дизайна исследования, стандартизации всех оценочных процедур [Mislevy, Steinberg, Almond, 2003]. Если же осуществляются международные или межкультурные сопоставления, то к вышеперечисленным вызовам добавляются трудности, связанные с необходимостью обеспечения эквивалентности измерений в разных странах, культурах и языках [Rutkowski et al., 2010].

Проектов, аналогичных PISA или PIRLS, но ориентированных на первый год обучения в школе, не существует. Не так

1. Подходы к оцениванию детей

¹ <http://ioe.hse.ru/ipips>

много и инструментов, потенциально применимых для международных сравнений, отвечающих критериям объективности, комплексности и высокого качества. Одним из таких инструментов является iPIPS (*International Performance Indicators in Primary Schools*), предназначенный для стартовой диагностики детей на входе в школу и оценки их прогресса за первый год обучения.

Сегодня в нескольких странах существуют версии инструмента iPIPS, первоначально разработанного в Университете Дарема в Великобритании [Tymms, 1999; Archer et al., 2010; Niklas, Schneider, 2013; Wildy, Styles, 2011]. Оценивание детей проводится в формате адаптивного компьютерного тестирования и предполагает непосредственное выполнение заданий ребенком при поддержке интервьюера (им может быть школьный психолог, методист, учитель, прошедший тренинг). iPIPS проводится в два этапа: когда дети только приходят в школу и когда заканчивают первый год обучения. Совокупность заданий инструмента позволяет оценить развитие ребенка в четырех сферах: ранние навыки чтения и математики, фонематическая грамотность, словарный запас.

Характеристики рассматриваемого инструмента и возможности, которые он предоставляет, дают основания считать весьма перспективной задачу разработки его русскоязычной версии и присоединения к развивающемуся проекту iPIPS. Во-первых, предлагаемый iPIPS формат прямого оценивания детей, хотя и требует определенных ресурсов (привлечения интервьюеров), позволяет сделать вывод о возможностях и затруднениях каждого ребенка. Индивидуальная работа с ребенком гарантирует объективность, редко достижимую при непрямой оценке. Во-вторых, формат компьютерного адаптивного тестирования позволяет оценивать, что знают и умеют делать дети, в щадящем, комфортном для них режиме, не предъявляя им задания, слишком сложные для их актуального уровня. В-третьих, инструмент позволяет измерить прогресс детей за первый год обучения в школе. В-четвертых, доказанная надежность и валидность оригинальной версии инструмента, а также его адаптированных версий в других англоязычных странах [Demetriou, Merrell, Tymms, 2017; Tymms et al., 2009; Wildy, Styles, 2011] открывает перспективы проведения международных сопоставительных исследований развития первоклассников, включая неанглоязычные страны.

2. Модель чтения iPIPS

Исследования раннего чтения показывают, что, во-первых, дети учатся читать поэтапно, во-вторых, язык, на котором проходит обучение, влияет на то, сколько времени потребуется ребенку на овладение чтением, в-третьих, в алфавитных языковых системах овладение чтением проходит в целом одни и те же эта-

пы [Seymour, Aro, Erskine, 2003; Rayner et al., 2012; Gove, Wet-terberg, 2011].

Наиболее общая модель чтения, которую авторы инструмента iPIPS сформировали на основе многолетних исследований и которая согласуется с теоретическими идеями российских педагогов [Егоров, 2006; Корнев, 1995], состоит в том, что дети в овладении чтением проходят несколько важных фаз: развитие общих представлений о том, как работает наш язык, узнавание букв и «схватывание» графической оболочки слова, постепенное развитие навыка декодирования текста и, наконец, появление навыка осмысленного чтения текста [Merrell, Tymms, 2007].

Знание того, каким образом работает наш язык, предполагает понимание базового устройства языка, его форм, включая знания о том, как держать книгу, откуда начинать читать текст, о том, что в европейских языках принято читать слева направо, умение отличать отдельные буквы от целых слов и т. д. [Clay, 1985]. Знание букв — еще один важный предиктор чтения в долгосрочной перспективе, поскольку буквы являются фундаментальным элементом алфавитной письменности [Foulin, 2005]. Распознавание графической оболочки слова часто происходит одновременно со знакомством с буквами. Согласно результатам исследований, многие дети способны узнавать и понимать смысл простых, часто встречающихся им слов еще до появления сформированного навыка чтения [Harrison, 2004].

Ребенку, уже знающему, как работает язык, потребуется применение определенных методов дальнейшего обучения чтению как декодированию и пониманию текста [Harrison, 2004; Merrell, Tymms, 2007; Журова, Эльконин, 1963]. Постепенно практикуясь в чтении с помощью одной или нескольких стратегий, дети овладевают навыком декодирования текста, они набираются опыта и учатся прочитывать все больше слов автоматически, и тогда часть их способности к обработке информации может быть перенаправлена на понимание читаемого текста [Merrell, Tymms, 2007].

Предлагаемая авторами iPIPS модель чтения, описывающая постепенное развитие читательских навыков ребенка, включает следующие этапы:

- базовые идеи о чтении: концепция и структура текста;
- распознавание букв;
- распознавание графической оболочки слова;
- декодирование текста (механическое чтение);
- чтение с пониманием.

Адаптация иноязычного инструмента — это исключительно сложный и ресурсоемкий процесс [Leong et al., 2016]. Кроме того,

3. Адаптация и локализация

даже при соблюдении всех процедур результат далеко не всегда дает возможность проводить международные сопоставительные исследования [Иванова, 2018].

Цель адаптации инструмента состоит в обеспечении валидности интерпретации результатов, полученных с помощью версий инструментов разных стран, языков или культур [Sireci, Patsula, Hambleton, 2005; Leong et al., 2016]. Научно-исследовательские организации, работающие в сфере оценивания, предлагают различные руководства и рекомендации, призванные обеспечить качество адаптации при проведении международных исследований [AERA, APA, NCME, 2014; Leong et al., 2016]. Цель этих процедур состоит в том, чтобы максимально приблизиться к главному условию дальнейшего использования данных международных исследований — обеспечению сопоставимости результатов оценивания. Это возможно только в том случае, если измерения в версиях инструмента оценивания, созданных на базе разных языков и ориентированных на разные культуры, будут эквивалентны. В свою очередь, эквивалентность измерений включает три важных компонента: 1) эквивалентность конструкта; 2) эквивалентность инструмента; 3) эквивалентность процедуры исследования [Ercikan, Roth, Asil, 2015].

При создании русскоязычной версии инструмента iPIPS мы исходили из того, что часть инструмента iPIPS по оцениванию чтения детей может быть только локализована. Локализация — это процесс преобразования некоторого продукта таким образом, чтобы в нем учитывалась культурная и языковая специфика целевой аудитории (страны, региона и т.д.), где он будет использоваться [Esselink, 2000]. В данной работе под локализацией инструмента оценивания мы будем понимать процесс создания версии инструмента на другом языке, в основе которой лежит та же теоретическая модель, что и в оригинальной версии, но учтены культурные особенности страны локализации. Главное отличие локализации от адаптации заключается в отсутствии задачи прямого сравнения индивидуальных результатов тестирования между странами.

Работа над русскоязычной версией инструмента для оценки ранних навыков чтения потребовала значительного времени, больших усилий, создаваемая версия неоднократно изменялась. Наиболее сложные проблемы из тех, с которыми пришлось столкнуться в ходе этой работы, — это значительная, более чем двухлетняя, разница в возрасте между целевой англоязычной аудиторией (4–5 лет) и российскими первоклассниками (6–7 лет), а также ряд существенных различий в структуре русского и английского языков.

Первым шагом на пути создания русскоязычной версии iPIPS стал перевод и экспертиза заданий, созданных для британских детей и предназначенных для оценки их базовых навыков чтения. Перевод (прямой и обратный) осуществлялся в соответствии с рекомендациями Международной тестовой комиссии [Leong et al., 2016]. Оригинальная версия iPIPS в части оценивания навыков чтения ориентирована на проверку особенностей формирования языковой способности ребенка, учащегося в начальной школе и говорящего на английском языке как на родном. В эту часть теста iPIPS вошло несколько блоков заданий, соответствующих разным этапам модели чтения, разработанной авторами iPIPS. Так, понимание структуры текста проверялось с помощью блока заданий, в котором участника тестирования просили указать, с какого элемента начинается прочтение предъявленного текста, где текст заканчивается и т. п. Навык распознавания букв проверялся с помощью заданий, в которых от ребенка требовалось назвать букву или соответствующий ей звук. Для оценивания навыка распознавания графической оболочки слов ребенку предлагались задания, в которых нужно соотнести фонетическую оболочку (озвученную интервьюером) и графическую оболочку, представленную в тесте (например, выбрать в ряду из четырех или пяти слов названное). Навык чтения на этапе декодирования проверялся в ходе выполнения ребенком задания на прочтение короткой истории вслух, при этом задача интервьюера состояла в количественном подсчете правильно прочитанных слов.

При создании русскоязычной версии iPIPS для оценивания навыков чтения эти задания сохранялись фактически в неизменном виде (были переведены формулировки заданий и подобраны эквивалентные буквы, слова и тексты в русском языке с учетом их частотности и степени, в которой этими единицами владеет ребенок на входе в школу). Однако разработка блока заданий для оценки сформированности навыка понимания представляла сложную задачу, поскольку тексты, предъявляемые детям — носителям русского и английского языков, должны быть сопоставимыми по уровню сложности, а содержащиеся в них задания (так называемые ловушки) должны быть направлены на проверку одинаковых навыков. Работа по локализации этой части теста iPIPS проходила в несколько этапов: сначала были изучены лингвистические характеристики оригинального текста, потом были подобраны эквивалентные «ловушки» в русском языке, наконец, был смоделирован русскоязычный текст, содержащий эти «ловушки» и содержательно близкий англоязычному оригиналу. В следующих разделах представлено детальное описание каждого этапа проделанной работы.

Ребенку предлагается прочитать текст, в некоторых частях которого ему требуется выбрать один из трех предложенных вари-

4. Методология

4.1. Локализация теста iPIPS для оценивания навыков чтения

4.1.1. Лингвистическая характеристика оригинального теста для оценивания навыков чтения с пониманием

антов (лингвистическая экспертиза проводилась на материале текстов *Underground*, *Walking to School*). Делая выбор, он сталкивается с так называемыми «ловушками», в которых заложены некоторые из наиболее явных сложностей освоения родного языка (английского): дифференциация между оболочками разных языковых единиц; дифференциация между грамматическими формами (видо-временными формами глагола, личными местоимениями в номинативном и объектном падежах и притяжательными местоимениями, формами единственного и множественного числа имени существительного и проч.); знание артиклей и особенностей их использования; дифференциация значений языковых единиц, знание валентностных и сочетаемостных характеристик слов, знание одиночных и парных предлогов и др.

Единицы, составляющие «ловушку», связаны попарно, при этом один из трех представленных вариантов входит сразу в оба противопоставления. Например, в кластере *creatures—annuals—animal* первая пара *creatures—animal* содержит противопоставление по форме числа существительного (первое слово используется в форме множественного числа, а второе — в форме единственного числа). А в паре *annuals—animal* работает принцип фонетического подобия. При этом общей единицей, задействованной в обеих парах, является слово *animal*, а правильным ответом в заданном контексте является слово *creatures*.

В некоторых «ловушках» используется более сложный механизм вуалирования трудности. Так, в предложении *They can run quickly and are very good at leaping upon to fences, trees and quit/other/offer high places* пара слов *other—offer* связана на основе фонетического/графического подобия, однако в паре слов *quit—other* непосредственная связь отсутствует. «Ловушка» состоит в том, что слово *quit* обладает сходной графической оболочкой со словом *quite*, которое, в свою очередь, может выступать заменой слова *other* в данном контексте: *other high places—quite high places*. Таким образом, ожидается, что ребенок воспользуется вариантом ответа *quit* в том случае, если он не способен дифференцировать оболочки слов *quite* и *quit* (иными словами, ошибка будет вызвана графическим подобием слов).

Для того чтобы эксплицировать логику, определяющую подбор «ловушек» для каждого из трех текстов, необходимо рассмотреть особенности английского языка.

4.1.1.1. Структурные особенности английского языка в сравнении с русским

Современный английский язык является флективным языком аналитического типа, в процессе эволюции которого синтетические формы, представляющие собой объединение в пределах одной фонетической и графической оболочки лексического и грамматического значения (*с-дел-а-л; de-motiv-ate-d*), постепенно вытеснялись (и вытесняются) аналитическими формами,

представляющими собой цепочку самостоятельных фонетических и графических оболочек, в которых дискретировано лексическое и грамматическое значение (*have been asked*). При том что в современном английском языке присутствуют синтетические формы, они становятся все более простыми (содержат меньшее количество компонентов при их сравнительной простоте и часто однозначности — например, *re-do*, где приставка *re-* является носителем только одного значения «повтор действия»).

Другая особенность английского языка состоит в наличии сравнительно большого количества омонимичных форм: омофонов (*of — off; night — knight* и проч.); омографов (*to wind [ai] — wind [i], to tear [εə] — tear [iə]*) и омоформ (*heard (past “to hear”) — herd, left (past “to leave”) — left (лево)*). Наравне с собственно омонимами в английском языке присутствует много сходных по звучанию или написанию форм, не представляющих собой омонимичные пары. Сходство графической и/или фонетической оболочки представлено в следующих примерах из рассмотренных текстов: *wake — work — walk, buy — boy, carried — cared, leaf — leave* и проч.

Центральное положение в системе английского языка занимает глагол [Филлмор, 1981]. Личные и неличные формы глагола, видимо, усваиваются на самых первых этапах освоения родного языка. Это находит отражение в текстах, предлагаемых в англоязычной версии теста. В них содержатся «ловушки» на дифференциацию оболочки и значения глагольных форм, на знание элементов, образующих глагольную форму (вспомогательного глагола и формы смыслового глагола).

Значимой характеристикой английского языка является фиксированный порядок слов, проявляющийся, в частности, в наличии большого количества устойчивых синтаксических моделей (*Complex Object, For-to-the-Infinitive Construction* и проч.). Ребенок начинает усваивать эти модели в виде гештальтных образований уже на самых ранних этапах овладения родным языком (английским), поэтому в рассмотренных текстах присутствует большое количество «ловушек», ориентированных на проверку умения применять синтаксические модели.

Степень сложности текста определяется как количеством, так и качеством содержащихся в нем «ловушек»: в тексте *Walking to School* преобладают «ловушки» на графическое/фонетическое подобие (8 «ловушек»²). В тексте *Underground* содержится только 6 «ловушек» на графическое/фонетическое подобие, однако много «ловушек» на знание грамматики и синтаксиса.

² Статистическая обработка данных представляется несколько затруднительной в силу того, что «ловушки» строятся с опорой на несколько разных оснований, поэтому цифры приведены ориентировочные.

4.1.1.2. Структурные особенности русского языка в сравнении с английским. Учет этих особенностей при создании русскоязычной версии теста

Разработка сопоставимого теста для русскоязычных детей возможна с учетом особенностей русского языка, отличающих его от английского. Основное отличие состоит в том, что русский язык на протяжении своего исторического развития являлся (и является) флективным языком синтетического типа, что проявляется в преобладании синтетических форм, в которых в пределах одной фонетической и графической оболочки сосуществуют несколько лексических и грамматических значений. Так, в форме слова *при-дорож-н-ого* представлены лексические значения («около, у» + «дорога») и грамматические значения (род, число, падеж). В отличие от английского языка, в русском языке не частотны аналитические формы, содержащие несколько самостоятельных компонентов, каждый из которых является носителем какого-то лексического или грамматического значения (ср. *have been working* и *буду готовить*). Единичные примеры аналитизма в русском языке представлены сравнительной и превосходной формами прилагательных (*более удачный/самый важный*) и формами будущего времени глагола (*будем заниматься*). Данное расхождение между языками должно быть учтено при создании «ловушек» для проверки способности дифференцировать грамматическую форму: если уже в самом простом англоязычном тексте *Walking to School* ряд «ловушек» предполагает выбор правильного варианта из перечня форм *Continuous (is/was/were shining* или *was wearing/wear/wore)*, то в русскоязычном тексте «ловушки» могут предусматривать выбор формы лица и/или числа глагола (*светился/светились/светилось* или *было/были/была*).

Другую особенность русского языка составляет его номинативный характер, или ориентация на имя существительное и группу существительного в контексте синтаксической структуры. Иными словами, имя существительное в русском языке является носителем значения, принципиально важного для понимания смысла целостного предложения. В английском же языке эта роль отводится глаголу (в связи с этим принято говорить о вербоцентризме английского языка). Это расхождение между языками проявляется, в частности, в том, что именные части речи в русском языке обладают развернутой системой грамматических категорий (род, число, падеж и проч.), а в английском языке система грамматических категорий имени существительного включает только категорию числа, довольно редуцированную категорию падежа и категорию детерминативности (определенности — неопределенности), а некоторые именные части речи и вовсе обладают только одной категорией (прилагательное — категорией степеней сравнения). В англоязычных текстах фактически не представлены «ловушки», ориентированные на проверку способности первоклассников дифференцировать грамматические формы именных частей речи, однако в русско-

язычные тексты их необходимо внести в силу их релевантности процессу формирования языковой способности ребенка (ср. *животные/животное/животного*, или *высокому/высокого/высокий*, или *ребята/ребятам/ребят*).

При адаптации текстов для тестирования носителя русского языка следует отказаться от многочисленных «ловушек» на идентичные и сходные графические и фонетические оболочки, поскольку таких слов в русском языке во много раз меньше, чем в английском (большое количество омонимов в английском языке объясняется тем, что слова преимущественно являются односложными, а это повышает вероятность совпадения оболочек). Кроме того, в русскоязычном тексте невозможно сохранить «ловушки» на дифференциацию артиклей и вспомогательных глаголов (в русском языке те единичные вспомогательные глаголы, которые используются для образования форм будущего времени (*будет работать*), легко распознаются даже на начальных этапах освоения русского языка как родного).

Наконец, еще одна особенность русского языка заключается в свободном порядке слов. Синтетизм русского языка (наличие грамматических маркеров в пределах одной фонетической и графической оболочки) предоставляет относительную свободу при построении целостной синтаксической конструкции. Используя развернутую систему грамматических маркеров, носитель русского языка устанавливает необходимые логические связи и формулирует законченную мысль без необходимости опираться на аранжировку компонентов (естественно, речь идет об относительной свободе, поскольку в русском языке действуют правила соположения логически связанных элементов текста: прилагательное не может быть оторвано от определяемого слова, а существительное, выполняющее функцию дополнения, обязательно находится в пределах группы глагола). В связи с этим в русском языке, в отличие от английского, отсутствуют устойчивые синтаксические модели (типа *Complex Subject*, *Complex Object*), поэтому подобные синтаксические «ловушки» также не могут быть транслированы в русскоязычный текст.

Исходя из специфики русского языка была предложена типология «ловушек», ориентированных на языковое знание носителя русского языка.

Первую группу составляют «ловушки», основанные на принципе фонетического и графического подобия, в которых три предложенных варианта ответа объединены сходством графической и/или фонетической оболочки: *пока/пора/гора ... вставать*; *Во дворе ее уже ... ждал/сдал/жбан ... Костя* и проч.

Во вторую группу вошли «ловушки», предназначенные для проверки способности дифференцировать грамматические формы слов. Так, в «ловушку» могут входить существительные, различающиеся формой числа/падежа: *ребята/ребятам/ребят*

... пошли к школе; выпей ... молока/молоку/молоком; глаголы в разных видо-временных формах, формах лица и числа: *Они быстро ... бегать/бегает/бегают; Тогда им ... нужен/нужна/нужно ... помощь*; личные и притяжательные местоимения в разных формах: *Но ... она/он/они ... все же ловят мышей*.

Третья группа представлена «ловушками», направленными на проверку способности выбирать правильную единицу в зависимости от контекста. В этих случаях предшествующий контекст допускает использование любого из предложенных вариантов, а ограничения накладываются последующим контекстом либо же выбор единицы может определяться широким контекстом. Так, в предложении *Я буду с ним ... учить/играть/гулять ...* допустим любой из предложенных вариантов, выбор может быть сделан только тогда, когда ребенок прочитает продолжение фразы: команду «Рядом».

Особый подвид смысловых «ловушек» составляют лексические «ловушки», в которых предполагается выбор из нескольких приставочных слов (производных от одной основы). От ребенка требуется не только идентифицировать правильную форму с контекстом, но и дифференцировать собственно предложенные варианты. Так, в предложении *Прежде чем выйти из дома, Юля собралась/набралась/забралась* требуется прежде всего понять значение каждого из глаголов и только потом соотнести требуемое значение с ближайшим контекстом.

Таким образом, учет различий между русским и английским языками позволил нам разработать типологию «ловушек», в которой учитываются структурные особенности «принимающего» русского языка. Готовая версия должна была пройти процедуру последующей апробации и валидации.

4.2. Сбор свидетельств валидности

В октябре 2013 г. было проведено первое пилотное исследование инструмента iPIPS в Великом Новгороде и Новгородской области на выборке, в которую вошли около 300 первоклассников. Оказалось, что ряд заданий демонстрирует потолочный эффект: они оказались слишком легкими для детей. В 2014 г. состоялась встреча участников проекта с экспертами в сфере начального и дошкольного образования, на которой обсуждались вызовы, стоящие перед данным исследованием. С учетом рекомендаций экспертов и в тесном сотрудничестве с партнерами из Университета Дарема задания российской версии iPIPS были пересмотрены, после чего в течение 2014–2018 гг. были проведены еще несколько апробаций инструмента. В частности, были исключены слишком легкие задания и добавлены новые, более трудные. В результате была создана локализованная версия инструмента iPIPS, для которой была проведена серия исследований по сбору свидетельств валидности.

Валидизация инструмента представляет собой необходимый этап в обосновании качества инструмента оценивания, это длительный процесс, требующий проведения большого числа исследований. В качестве первого этапа обоснования валидности данных, получаемых с помощью инструмента iPIPS, были собраны доказательства конструктивной валидности. Для этого было проведено исследование внутренней структуры шкалы чтения — ее размерности, а также оценены ее психометрические характеристики с помощью результатов оценивания, полученных в 2015 г. на выборке учащихся 1-х классов нескольких школ г. Москвы — 1822 ребенка, средний возраст 7,4 года, 51% девочки.

При проведении психометрического анализа русскоязычной шкалы чтения мы использовали базовую дихотомическую модель Раша, созданную для работы с оценочными инструментами, состоящими из заданий дихотомического типа, где есть один правильный и один неправильный ответ [Rasch, 1960]. Эта же модель была использована в анализе оригинальной версии iPIPS на английском языке [Tumms, 1999]. Для анализа данных тестирования было использовано программное обеспечение WINSTEPS [Linacre, 2011].

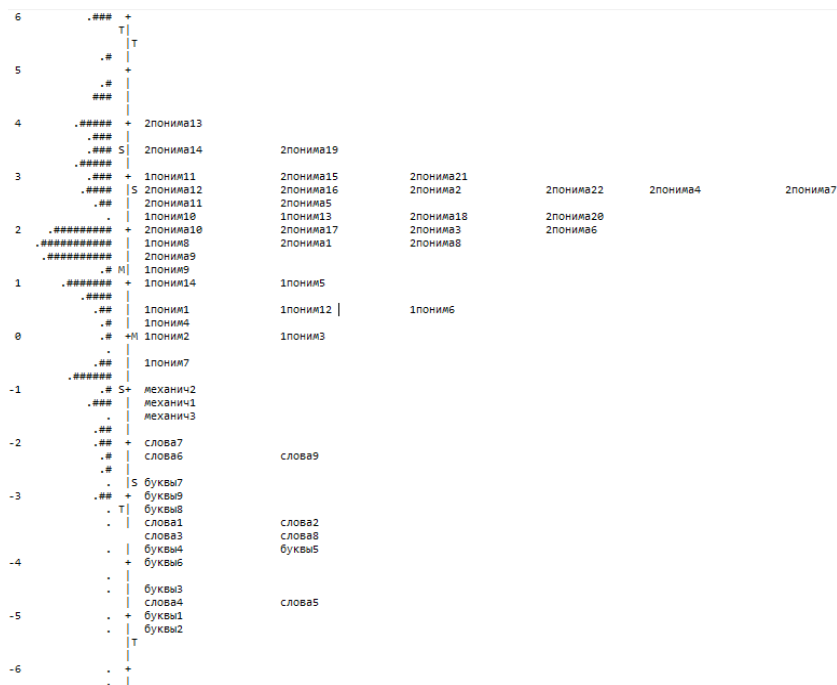
На рис. 1 представлена карта переменных — изображение относительного положения трудностей заданий и уровня подготовленности учащихся на шкале чтения. Слева на рисунке представлена шкала логитов. Справа от вертикальной оси шкалы расположены задания по чтению, их трудность постепенно возрастает от более легких (например, на знание букв — в нижней части карты) к более трудным (например, на чтение с пониманием — задания в верхней части карты). Слева от оси представлено распределение показателей подготовленности учащихся, выполнявших задания.

Эмпирические данные, полученные в ходе тестирования и визуализированные в виде карты, соответствуют нашим теоретическим ожиданиям в отношении того, как должны располагаться задания на шкале чтения. В частности, мы видим, что самым легким оказывается задание, где ребенка просят назвать букву, а самым трудным — задание на чтение с пониманием, задание, представляющее собой смысловую «ловушку». Наконец, на карте хорошо видно, какое широкое распределение имеют и трудности заданий, и уровни подготовленности учащихся. Распределение заданий соответствует оригинальной версии инструмента на английском языке, что позволяет предположить потенциальную возможность проведения в будущем непрямого сопоставления результатов детей из разных стран, используя аппарат современной теории тестирования.

5. Результаты

5.1. Общее описание шкалы чтения

Рис. 1. Карта переменных шкалы чтения



5.2. Согласие данных с моделью

Раш-анализ согласия с моделью работает с «остатками», т. е. различиями между наблюдаемым ответом ученика на задание и его ожидаемым модельным значением [Ludlow, Haley, 1995]. По всем заданиям получено адекватное согласие с используемой моделью измерения, за исключением двух заданий, где ребенка просили назвать букву (что может быть связано с легкостью этих заданий).

5.3. Анализ размерности шкалы

Одним из принципов Раш-подхода к измерению является необходимость единообразия концепта, заложенного в задания инструмента, что часто определяют как одномерность, или существование некоторой единой латентной характеристики, объединяющей данные [Hattie, 1984]. В нашем случае мы измеряем чтение как одномерный конструкт, реализующийся в совокупности заданий нашего инструмента. Проверка требования одномерности осуществлялась с помощью анализа главных компонент остатков, полученных в ходе Раш-моделирования. Результаты анализа показали, что тест является существенно одномерным [Linacre, 2011].

Классическая надежность (альфа Кронбаха) составляет 0,98 и является исключительно высоким показателем согласованности всех заданий теста. При этом разброс значений трудности заданий очень широк — более 9 логитов, разброс показателей уровня подготовленности испытуемых — более 15 логитов.

Таким образом, мы можем заключить, что тест по чтению инструмента iPIPS является качественным измерительным инструментом, который может быть использован для оценивания развития раннего чтения у детей, начинающих обучение в школе. Психометрический анализ, проведенный на результатах оценивания выборки первоклассников г. Казани в 2016 г. [Республиканский центр мониторинга качества образования, 2016] еще раз подтвердил полученные выводы.

5.4. Анализ надежности шкалы

Основная цель работы состояла в описании этапов локализации и валидации русскоязычной версии международного инструмента iPIPS в части оценивания чтения. Инструмент предназначен для детей, только пришедших в систему школьного обучения и, соответственно, только овладевающих навыком чтения.

Чтение является исключительно сложным навыком, ключевым для обучения в школе: он определяет для ребенка дальнейшую возможность учиться [Антипкина, Кузнецова, Карданова, 2017; Stanovich, 2000]. Процесс чтения базируется на целой серии когнитивных процессов, позволяющих обрабатывать лексическую (на уровне понимания слов) и синтетическую (на уровне предложений) информацию, делать выводы, использовать метакогнитивные стратегии (способность направлять собственный учебный процесс, сосредоточить свое внимание и следить за процессом чтения) [Magliano et al., 2007; Stanovich, 2000]. Навыком понимания читаемого текста ребенок овладевает постепенно, последовательно проходя ряд этапов от первого знакомства с текстом на родном языке, узнавания букв и слов до чтения отдельных составляющих и складывания их в слова и, наконец, понимания того, что было прочитано. Предложенная авторами iPIPS модель чтения предполагает именно такое поэтапное усвоение навыков чтения [Merrell, Tymms, 2004].

Характер поставленной в исследовании задачи — локализации и валидации русскоязычной версии инструмента, который тестирует уровень развития такого сложного навыка, как чтение, в такой специфической целевой аудитории, как дети на входе в школу, — не позволяет просто воспользоваться имеющимися практиками адаптации инструментов на языки разных стран и культур, например руководством Международной тестовой комиссии [Leong et al., 2016] или опытом стран — участниц Международного исследования качества чтения и понимания текста PIRLS [Mullis et al., 2009].

6. Обсуждение результатов

В ситуации, когда дети только учатся читать, адаптация международного инструмента означает, что единицами оценивания будут конкретные единицы русского и, в случае iPIPS, английского языка, и обеспечить строгую эквивалентность инструментов [Ercikan, Roth, Asil, 2015] будет невозможно. А значит, необходимо проводить локализацию инструмента с сохранением особенностей принимающего, а именно русского, языка. Оценивание формируемого навыка чтения с пониманием представляет собой исключительно сложную задачу, поскольку тексты, предъявляемые детям — носителям русского и английского языков, должны быть сопоставимыми по уровню сложности, а содержащиеся в них задания должны быть направлены на проверку одних и тех же навыков. Поэтому работа по локализации этой части теста iPIPS проходила в несколько этапов: были изучены лингвистические характеристики оригинальных тестовых заданий, подобраны аналогичные средства в русском языке, наконец, были смоделированы русскоязычные задания, содержательно близкие англоязычному оригиналу.

В силу ряда существенных структурных различий английского и русского языков (среди которых особенно важны номинотцентризм русского языка и вербоцентризм английского языка, несовпадение состава и функционала частей речи, фиксированный порядок слов в английском языке, в результате чего в нем частотны устойчивые синтаксические модели, и нефиксированный порядок слов в русском языке, что хорошо согласуется с развитой системой морфологических маркеров) этапы становления языковой способности у англоговорящих и русскоговорящих детей не совпадают. Это, несомненно, сказывается на процессе формирования навыка чтения.

Для создания настолько идентичных, насколько это возможно, инструментов оценивания навыка чтения у учащихся начальной школы в Великобритании и России было необходимо произвести лингвистический анализ оригинальной версии инструмента iPIPS, идентифицировать соотносимые языковые средства английского и русского языков и создать тестовые задания на русском языке, направленные на оценку равнозначных навыков чтения.

В статье подробно описана процедура локализации данного инструмента iPIPS в части чтения и процесс сбора свидетельств его валидности средствами современной теории тестирования. В частности, приведены анализ структуры и размерности шкалы, анализ функционирования отдельных заданий и шкалы в целом, анализ внутренней согласованности. Таким образом, в ходе серии проведенных процедур было осуществлено обоснование психометрического качества шкалы, проведена оценка его надежности, подтверждена структура и иерархичность заданий, соответствующие изложенной в работе теоретической рамке.

Произведенный и описанный в данной работе сбор свидетельств конструктивной валидности является необходимым, но не последним этапом длительного процесса валидации инструмента измерения. Необходимо проведение дополнительных исследований, в частности сбор свидетельств содержательной, а также предсказательной валидности. Серии данных исследований проведены, но не включены в данную статью и в будущем должны быть описаны отдельно.

Наконец, исследование по доказательству принципиальной возможности проведения непрямого (в связи с неэквивалентностью инструментов по чтению на русском и английском языках, но наличием единых стандартизованных процедур, теоретической рамки и конструкта) сопоставительного исследования результатов по чтению первоклассников из России и Великобритании также могло бы быть рассмотрено как одно из свидетельств валидности инструмента.

Литература

1. Антипкина И. В., Кузнецова М. И., Карданова Е. Ю. (2017) Что способствует и что мешает прогрессу детей в чтении // Вопросы образования/Educational Studies Moscow. № 2. С. 206–233. doi: 10.17323/1814-9545-2017-2-206-233.
2. Безруких М. М., Филиппова Т. А., Байдина В. А. (2012) Диагностика развития детей старшего дошкольного возраста как способ раннего выявления рисков дезадаптации // Новые исследования. № 1 (30). С. 145–157.
3. Болотов В. А., Вальдман И. А., Ковалева Г. С., Пинская М. А. (2013) Российская система оценки качества образования: главные уроки // Качество образования в Евразии. № 1. С. 85–121.
4. Егоров Т. Г. (2006) Психология овладения навыком чтения. СПб.: КАРО.
5. Журова Л. Е., Эльконин Д. Б. (1963) К вопросу о формировании фонематического восприятия у детей дошкольного возраста // А. В. Запорожец, А. П. Усова (ред.) Сенсорное воспитание дошкольников. М.: Изд-во АПН РСФСР. С. 213–227.
6. Иванова А. Е. (2018) Проблема сопоставимости результатов в международных сравнительных исследованиях образовательных достижений // Отечественная и зарубежная педагогика. Т. 1. № 2. С. 68–81.
7. Ковалева Г. С., Даниленко О. В., Ермакова И. В., Нурминская Н. В., Гапонова Н. В., Давыдова Е. И. (2012) О первоклассниках: по результатам исследований готовности первоклассников к обучению в школе // Муниципальное образование: инновации и эксперимент. № 5. С. 30–37.
8. Корнев А. Н. (1995) Дислексия и дисграфия у детей. СПб.: Гиппократ.
9. Республиканский центр мониторинга качества образования (2016) Материалы о мониторинге в 1-х классах iPIPS. Отчет. Весна 2016. <http://rcmko.ru/meropriyatiya/monitoringi/ipips/materialy-o-monitoringe-v-1-klassah-ipips/>.
10. Филлмор Ч. (1981) Дело о падеже // В. А. Звегинцев (ред.) Новое в зарубежной лингвистике. Вып. 10. М.: Прогресс. С. 369–495.
11. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME). (2014) Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association.

12. Archer E., Scherman V., Coe R., Howie S. J. (2010) Finding the Best Fit: The Adaptation and Translation of the Performance Indicators for Primary Schools for the South African Context // *Perspectives in Education*. Vol. 28. No 1. P. 77–88.
13. Clay M. W. (1985) *The Early Detection of Reading Difficulties*. Auckland, New Zealand: Heinemann.
14. Demetriou A., Merrell C., Tymms P. (2017) Mapping and Predicting Literacy and Reasoning Skills from Early to Later Primary School // *Learning and Individual Differences*. Vol. 54. February. P. 217–225.
15. Ercikan K., Roth W. M., Asil M. (2015) Cautions about Inferences from International Assessments: The Case of PISA 2009 // *Teachers College Record*. Vol. 117. No 1. P. 1–28.
16. Esselink B. (2000) *A Practical Guide to Localization*. Vol. 4. Amsterdam; Philadelphia: John Benjamins.
17. Foulon J. N. (2005) Why Is Letter-Name Knowledge Such a Good Predictor of Learning to Read? // *Reading and Writing*. Vol. 18. No 2. P. 129–155.
18. Gove A., Wetterberg A. (2011) *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy*. Research Triangle Park, NC: RTI International.
19. Harms T. (2013) *School-Age Care Environment Rating Scale (SACERS)*. New York, NY: Teachers College.
20. Harrison C. (2004) *Understanding Reading Development*. London: Sage.
21. Hattie J. (1984) An Empirical Study of the Various Indices for Determining Unidimensionality // *Multivariate Behavioral Research*. Vol. 19. No 1. P. 49–78.
22. Janus M., Brinkman S. A., Duku E. K. (2011) Validity and Psychometric Properties of the Early Development Instrument in Canada, Australia, United States, and Jamaica // *Social Indicators Research*. Vol. 103. No 2. P. 283–297.
23. Kuzmina Y., Ivanova A., Kaiky D. (2019) The Effect of Phonological Processing on Mathematics Performance in Elementary School Varies for Boys and Girls: Fixed-Effects Longitudinal Analysis // *British Educational Research Journal*. Vol. 45. No 3. P. 640–661.
24. Leong F. T., Bartram D., Cheung F., Geisinger K. F., Iliescu D. (2016) *The ITC International Handbook of Testing and Assessment*. New York, NY: Oxford University.
25. Linacre J. M. (2011) *Winsteps (Version 3.73) (computer software)*. Chicago, IL: Winsteps. com.
26. Ludlow L. H., Haley S. M. (1995) Rasch Model Logits: Interpretation, Use, and Transformation // *Educational and Psychological Measurement*. Vol. 55. No 6. P. 967–975.
27. Magliano J. P., Millis K., Ozuru Y., McNamara D. S. (2007) A Multidimensional Framework to Evaluate Reading Assessment Tools // D. S. McNamara (ed.) *Reading Comprehension Strategies: Theories, Interventions, and Technologies*. New York: Lawrence Erlbaum Associates. P. 107–136.
28. Merrell C., Tymms P. (2004) Diagnosing and Remediating Literacy Problems Using INCAS Software: Identifying Reading and Spelling Difficulties and Providing Help.
29. <http://www.pipsproject.org/Documents/CEM/publications/downloads/CEMWeb039%20Incas.pdf>.
30. Merrell C., Tymms P. (2007) Identifying Reading Problems with Computer-Adaptive Assessments // *Journal of Computer Assisted Learning*. Vol. 23. No 1. P. 27–35.

31. Mislevy R. J., Steinberg L. S., Almond R. G. (2003) Focus Article: On the Structure of Educational Assessments // *Measurement: Interdisciplinary Research and Perspectives*. No 1. P. 3–62.
32. Mullis I. V., Martin M. O., Kennedy A. M., Trong K. L., Sainsbury M. (2009) *PIRLS 2011 Assessment Framework*. Amsterdam, Netherlands: International Association for the Evaluation of Educational Achievement.
33. Niklas F., Schneider W. (2013) Home Literacy Environment and the Beginning of Reading and Spelling // *Contemporary Educational Psychology*. Vol. 38. No 1. P. 40–50.
34. Rasch G. (1960) *Studies in Mathematical Psychology: I. Probabilistic Models for some Intelligence and Attainment Tests*. Copenhagen: Danmarks Paedagogiske Institut.
35. Rayner K., Pollatsek A., Ashby J., Clifton C., Jr. (2012) *Psychology of Reading*. New York, NY: Psychology Press.
36. Rutkowski L., Gonzalez E., Joncas M., von Davier M. (2010) International Large-scale Assessment Data: Issues in Secondary Analysis and Reporting // *Educational Researcher*. Vol. 39. No 2. P. 142–151.
37. Seymour P. H., Aro M., Erskine J. M. (2003) Foundation Literacy Acquisition in European Orthographies // *British Journal of Psychology*. Vol. 94. No 2. P. 143–174.
38. Sireci S. G., Patsula L., Hambleton R. K. (2005) *Statistical Methods for Identifying Flaws in the Test Adaptation Process. Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. New Jersey: Lawrence Erlbaum Associates.
39. Stanovich K. E. (2000) *Progress in Understanding Reading: Scientific Foundations and New Frontiers*. New York: Guilford.
40. Tymms P. (1999) Baseline Assessment, Value-Added and the Prediction of Reading // *Journal of Research in Reading*. Vol. 22. No 1. P. 27–36.
41. Tymms P., Jones P., Albane S., Henderson B. (2009) The First Seven Years at School // *Educational Assessment, Evaluation and Accountability*. Vol. 21. No 1. P. 67–80.
42. Tymms P., Merrell C., Hawker D., Nicholson F. (2014) Performance Indicators in Primary Schools: A Comparison of Performance on Entry to School and the Progress Made in the First Year in England and Four Other Jurisdictions. <http://dro.dur.ac.uk/23562/1/23562.pdf>.
43. Vasilyeva M., Dearing E., Ivanova A., Shen C., Kardanova E. (2018) Testing the Family Investment Model in Russia: Estimating Indirect Effects of SES and Parental Beliefs on the Literacy Skills of First-Graders // *Early Childhood Research Quarterly*. Vol. 42. P. 11–20.
44. Wildy H., Styles I. (2011) Measuring What High-Achieving Students Know and Can Do on Entry to School: PIPS 2002–2008 // *Australasian Journal of Early Childhood*. Vol. 36. No 2. P. 51–62.

Constructing a Russian-Language Version of the International Early Reading Assessment Tool

Authors **Alina Ivanova**

Research Fellow, Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. Address: 20 Myasnitskaya St, 101000 Moscow, Russian Federation. E-mail: aeivanova@hse.ru

Ksenia Kardanova-Biryukova

Candidate of Sciences in Philology, Associate Professor, Institute of Foreign Languages, Moscow City University. Address: 5B Maly Kazenny Lane, 105064 Moscow, Russian Federation. E-mail: Kardanova-BirukovaKS@mgpu.ru

Abstract Successful adaptation of first-graders to school largely determines their subsequent educational attainment. In Russia as well as across the globe, there are few high-quality standardized assessment instruments providing comprehensive vision of what children know and what they can do when they start school. Large-scale evaluation of reading literacy is particularly challenging due to age-specific characteristics and the assessment format. This article outlines a step-by-step procedure of localizing a part of an international instrument iPIPS designed to measure early reading skills at the start of school and the progress made during the first school year, within the framework of Russian educational paradigm. Localization is understood as transformation of an instrument originally developed in another language (English in this case) so that it accounts for the cultural and linguistic characteristics of the target audience. The procedure included development of a Russian-language version of iPIPS and a series of studies to verify its construct validity. The process involved analyzing the linguistic characteristics of the original tasks, finding equivalent linguistic means in the Russian language, and designing Russian-language tasks identical to the original ones in terms of functionality. To verify construct validity of the localized instrument, we evaluated the psychometric properties of the scale, tested its reliability, and studied compliance of the task structure and hierarchy with the theoretical framework. The findings have revealed that large-scale local or regional tests administered using this localized assessment instrument may yield valuable data which can be further used for analysis of the current situation and informed decision-making in educational policy.

Keywords reading assessment tool, iPIPS, elementary school, localization, validation.

- References**
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, & NCME). (2014) *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Antipkina I., Kuznetsova M., Kardanova E. (2017) Chto sposobstvuet i chto meshaet progressu detey v chtenii [What Factors Help and Hinder Children's Progress in Reading?]. *Voprosy obrazovaniya/Educational Studies Moscow*, no 2, pp. 206–233. doi: 10.17323/1814-9545-2017-2-206-233.
- Archer E., Scherman V., Coe R., Howie S. J. (2010) Finding the Best Fit: The Adaptation and Translation of the Performance Indicators for Primary Schools for the South African Context. *Perspectives in Education*, vol. 28, no 1, pp. 77–88.
- Bezrukh M., Filippova T., Bajdina V. (2012) Diagnostika razvitiya detey starshego doshkolnogo vozrasta kak sposob rannego vyavleniya riskov de-

- zadaptatsii [Developmental Diagnostics of Elder Pre-School Children as the Means of Early Detection of Dysadaptation Risks]. *Novye issledovaniya*, no 1 (30), pp. 145–157.
- Bolotov V., Valdman I., Kovaleva G., Pinskaya M. (2013) Rossiyskaya Sistema otsenki kachestva obrazovaniya: glavnye uroki [Russian Quality Assessment System in Education: Key Lessons]. *Education Quality in Eurazia*, no 1, pp. 85–121.
- Clay M. W. (1985) *The Early Detection of Reading Difficulties*. Auckland, New Zealand: Heinemann.
- Demetriou A., Merrell C., Tymms P. (2017) Mapping and Predicting Literacy and Reasoning Skills from Early to Later Primary School. *Learning and Individual Differences*, vol. 54, February, pp. 217–225.
- Egorov T. (2006) *Psikhologiya ovladeniya navykom chteniya* [The Psychology of Learning to Read]. St. Petersburg: KARO.
- Ercikan K., Roth W. M., Asil M. (2015) Cautions about Inferences from International Assessments: The Case of PISA 2009. *Teachers College Record*, vol. 117, no 1, pp. 1–28.
- Esselink B. (2000) *A Practical Guide to Localization. Vol. 4*. Amsterdam; Philadelphia: John Benjamins.
- Fillmore Ch. (1981) Delo o padezhe [The Case for Case]. *New in Foreign Linguistics* (ed. V. Zvegintsev), Moscow: Progress, iss. 10, pp. 369–495.
- Foulin J. N. (2005) Why Is Letter-Name Knowledge Such a Good Predictor of Learning to Read? *Reading and Writing*, vol. 18, no 2, pp. 129–155.
- Gove A., Wetterberg A. (2011) *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy*. Research Triangle Park, NC: RTI International.
- Harms T. (2013) *School-Age Care Environment Rating Scale (SACERS)*. New York, NY: Teachers College.
- Harrison C. (2004) *Understanding Reading Development*. London: Sage.
- Hattie J. (1984) An Empirical Study of the Various Indices for Determining Unidimensionality. *Multivariate Behavioral Research*, vol. 19, no 1, pp. 49–78.
- Ivanova A. (2018) Problema sopostavimosti rezultatov v mezhdunarodnykh sravnitelnykh issledovaniyakh obrazovatelnykh dostizheniy [Problem of Comparability of Results in International Comparative Studies of Educational Achievements]. *Otechestvennaya i zarubezhnaya pedagogika*, vol. 1, no 2, pp. 68–81.
- Janus M., Brinkman S. A., Duku E. K. (2011) Validity and Psychometric Properties of the Early Development Instrument in Canada, Australia, United States, and Jamaica. *Social Indicators Research*, vol. 103, no 2, pp. 283–297.
- Kornev A. (1995) *Disleksiya i disgrafiya u detey* [Dyslexia and Dysgraphia in Children]. St. Petersburg: Gippokrat.
- Kovaleva G., Danilenko O., Ermakova I., Nurminskaya N., Gaponova N., Davydova E. (2012) O pervoklassnikakh: po rezultatam issledovaniy gotovnosti pervoklassnikov k obucheniyu v shkole [On First-Graders Based on the Results of First-Grade Readiness Assessments]. *Munitsipalnoe obrazovanie: innovatsii i eksperiment*, no 5, pp. 30–37.
- Kuzmina Y., Ivanova A., Kaiky D. (2019) The Effect of Phonological Processing on Mathematics Performance in Elementary School Varies for Boys and Girls: Fixed-Effects Longitudinal Analysis. *British Educational Research Journal*, vol. 45, no 3, pp. 640–661.
- Leong F. T., Bartram D., Cheung F., Geisinger K. F., Iliescu D. (2016) *The ITC International Handbook of Testing and Assessment*. New York, NY: Oxford University.

- Linacre J. M. (2011) *Winsteps (Version 3.73) (computer software)*. Chicago, IL: Winsteps. com.
- Ludlow L.H, Haley S. M. (1995) Rasch Model Logits: Interpretation, Use, and Transformation. *Educational and Psychological Measurement*, vol. 55, no 6, pp. 967–975.
- Magliano J. P., Millis K., Ozuru Y., McNamara D.S. (2007) A Multidimensional Framework to Evaluate Reading Assessment Tools. *Reading Comprehension Strategies: Theories, Interventions, and Technologies* (ed. D. S. McNamara), New York: Lawrence Erlbaum Associates, pp. 107–136.
- Merrell C., Tymms P. (2004) *Diagnosing and Remediating Literacy Problems Using INCAS Software: Identifying Reading and Spelling Difficulties and Providing Help*. Available at: <http://www.pipsproject.org/Documents/CEM/publications/downloads/CEMWeb039%20Incas.pdf> (accessed 1 November 2019).
- Merrell C., Tymms P. (2007) Identifying Reading Problems with Computer-Adaptive Assessments. *Journal of Computer Assisted Learning*, vol. 23, no 1, pp. 27–35.
- Mislevy R. J., Steinberg L. S., Almond R. G. (2003) Focus Article: On the Structure of Educational Assessments. *Measurement: Interdisciplinary Research and Perspectives*, no 1, pp. 3–62.
- Mullis I. V., Martin M. O., Kennedy A. M., Trong K. L., Sainsbury M. (2009) *PIRLS 2011 Assessment Framework*. Amsterdam, Netherlands: International Association for the Evaluation of Educational Achievement.
- Niklas F., Schneider W. (2013) Home Literacy Environment and the Beginning of Reading and Spelling. *Contemporary Educational Psychology*, vol. 38, no 1, pp. 40–50.
- Rasch G. (1960) *Studies in Mathematical Psychology: I. Probabilistic Models for some Intelligence and Attainment Tests*. Copenhagen: Danmarks Paedagogiske Institut.
- Rayner K., Pollatsek A., Ashby J., Clifton C., Jr. (2012) *Psychology of Reading*. New York, NY: Psychology Press.
- Republican Center for Monitoring the Quality in Education (2016) *Materialy o monitoringe v 1-kh klassakh iPIPS. Otchet. Vesna 2016* [Materials on iPIPS First-Grade Monitoring. Spring 2016 Report]. Available at: <http://rcmko.ru/meropriyatiya/monitoringi/ipips/materialy-o-monitoringe-v-1-klassah-ipips/> (accessed 1 November 2019).
- Rutkowski L., Gonzalez E., Joncas M., von Davier M. (2010) International Large-scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Researcher*, vol. 39, no 2, pp. 142–151.
- Seymour P. H., Aro M., Erskine J. M. (2003) Foundation Literacy Acquisition in European Orthographies. *British Journal of Psychology*, vol. 94, no 2, pp. 143–174.
- Sireci S. G., Patsula L., Hambleton R. K. (2005) *Statistical Methods for Identifying Flaws in the Test Adaptation Process. Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. New Jersey: Lawrence Erlbaum Associates.
- Stanovich K. E. (2000) *Progress in Understanding Reading: Scientific Foundations and New Frontiers*. New York: Guilford.
- Tymms P. (1999) Baseline Assessment, Value-Added and the Prediction of Reading. *Journal of Research in Reading*, vol. 22, no 1, pp. 27–36.
- Tymms P., Jones P., Albone S., Henderson B. (2009) The First Seven Years at School. *Educational Assessment, Evaluation and Accountability*, vol. 21, no 1, pp. 67–80.

- Tymms P., Merrell C., Hawker D., Nicholson F. (2014) *Performance Indicators in Primary Schools: A Comparison of Performance on Entry to School and the Progress Made in the First Year in England and Four Other Jurisdictions*. Available at: <http://dro.dur.ac.uk/23562/1/23562.pdf> (accessed 1 November 2019).
- Vasilyeva M., Dearing E., Ivanova A., Shen C., Kardanova E. (2018) Testing the Family Investment Model in Russia: Estimating Indirect Effects of SES and Parental Beliefs on the Literacy Skills of First-Graders. *Early Childhood Research Quarterly*, vol. 42, pp. 11–20.
- Wildy H., Styles I. (2011) Measuring What High-Achieving Students Know and Can Do on Entry to School: PIPS 2002–2008. *Australasian Journal of Early Childhood*, vol. 36, no 2, pp. 51–62.
- Zhurova L., Elkonin D. (1963) K voprosu o formirovanii fonematcheskogo vospriyatiya u detey doshkolnogo vozrasta [On the Development of Phonemic Perception in Preschoolers]. *Sensornoe vospitanie doshkolnikov* [Sensory Development of Preschoolers] (eds A. Zaporozhets, A. Usova), Moscow: RSFSR Academy of Pedagogical Sciences, pp. 213–227.