

А. Е. Иванова¹, И. В. Антипкина², Д. А. Федерякин³

¹aeivanova@hse.ru; ²iantipkina@hse.ru; ³dafederiakin@hse.ru

Национальный исследовательский университет «Высшая школа экономики»,
Москва, Россия

ДЕКОМПОЗИЦИЯ ЧИТАТЕЛЬСКИХ ДЕЙСТВИЙ В ИНСТРУМЕНТЕ ИЗМЕРЕНИЯ НАВЫКА ЧТЕНИЯ *

Данное исследование посвящено глубокому анализу компьютерного инструмента оценки навыков чтения и оцениванию связи уровня сложности его заданий с рядом когнитивных читательских действий, осуществляемых испытуемыми в процессе чтения и выполнения заданий в компьютерной форме. Предлагаемый аналитический подход (линейно-ограниченные экспланаторные модели современной теории тестирования) позволяет лучше понять природу смыслового чтения в условиях цифровой среды.

Ключевые слова: чтение, читательские действия, LLTM, разработка инструмента оценивания, когнитивная трудность заданий.

Alina E. Ivanova¹, Inna V. Antipkina², Denis A. Federiakin³

¹aeivanova@hse.ru; ²iantipkina@hse.ru; ³dafederiakin@hse.ru

National Research University Higher School of Economics, Moscow, Russia

DECOMPOSING READING ACTIVITIES USING A LINEAR LOGISTIC TEST MODEL

This study is devoted to an in-depth analysis of a computer-based instrument for assessing reading skills and an investigation of the relationship between the level of difficulty of test tasks and a set of cognitive reading actions carried out by students in the process of reading and performing test in a computer form. As an analytical approach, we used a family of explanatory latent linear test models. The proposed approach allows to make the assessment more accurate and better understand the nature of reading in a digital environment.

Keywords: reading, reading activities, LLTM, assessment design, cognitive difficulty of assignments.

Раннее смысловое чтение – это важный навык, привлекающий внимание как исследователей в области образования, так и политиков. Этот интерес основан на повышающейся роли грамотности в современном об-

© Иванова А. Е., Антипкина И. В., Федерякин Д. А., 2021

* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-29-14110 «Использование контекстной информации и информации из цифровой среды оценивания при измерении индивидуального прогресса учащихся начальной школы с помощью цифровых технологий».

шестве в целом и с тем, что именно первые годы обучения критически важны для более позднего развития. Сам по себе процесс чтения весьма сложен, он включает в себя целый комплекс когнитивных компонентов, читательских действий, осуществляемых ребенком, который способен прочесть текст и извлечь из него смысл. Некоторые исследователи, базирующиеся на так называемых новых исследованиях грамотности [1], подчеркивают необходимость учета тех новых способов, которые учащиеся используют при овладении грамотностью в целом и чтением в частности, в результате растущего влияния цифровых технологий на нашу повседневную жизнь.

Важно также отметить, что сегодня изучение в том числе навыка чтения в определенной степени базируется на результатах и новшествах, привносимых международными сравнительными исследованиями, в том числе, например, идеями Международного исследования грамотности чтения (PIRLS), который постепенно переходит в электронный формат – ePIRLS.

Данное исследование посвящено глубокому анализу компьютерного инструмента оценки навыков чтения и оцениванию связи уровня сложности его заданий с рядом когнитивных читательских действий, осуществляемых испытуемыми в процессе чтения и выполнения заданий в компьютерной форме. В своей работе мы опираемся на предположения теории валидности [2], требующей аргументации выводов, осуществляемых по результатам оценивания, в т. ч. на основе объяснения тех когнитивных процессов, которые происходят при решении заданий. Для осуществления такой аргументации необходим учет тех групп читательских действий и тех компонентов заданий, который определяет сложность задания.

Так, PIRLS выделяет четыре группы читательских действий, необходимых для извлечения информации в процессе чтения: 1) поиск информации в тексте; 2) простые выводы; 3) интерпретация и обобщение информации; 4) оценка и критическая интерпретация информации [3]. Также предыдущие исследования показали, что существует ряд дополнительных факторов, влияющих на сложность тестовых заданий [4; 5]. Во-первых, это сложность текста для восприятия, которая включает в себя необычный стиль изложения, сложную грамматику, редкую или сложную лексику, различные образные средства – эти характеристики могут затруднить понимание ребенком прочитанного текста. Во-вторых, размер отрывка в тексте, из которого необходимо извлечь информацию (конкретное предложение или пара предложений, абзац, весь текст). В-третьих, это дополнительные элементы (графики, таблицы, иллюстрации). В-четвертых, если в формулировке вопроса используются те же слова, что и в тексте, это может значительно упростить возможность верного ответа, а если в формулировке используются синонимы, то, наоборот, усложнить. Наконец, для ответа на вопрос ребенку может потребоваться только прочитанный текст, или же, дополнительные бэкграундные знания.

Также, на сложность вопроса может оказывать определенные эффекты среда, в которой осуществляется оценивание, т. е. важно учитывать особенности компьютерного формата тестирования и используемых типов заданий (задания с выбором одного правильного ответа из нескольких предложенных, задания со свободным конструируемым ответом, задания на выделение текста и другие форматы (сопоставление в форме таблицы, работа с рисунком)).

В данном исследовании мы пытаемся оценить эффекты различных групп читательских действий, заложенных в инструменте оценивания чтения, на трудность заданий теста, с тем чтобы сделать оценивание более точным и лучше понять природу смыслового чтения ребенком в условиях цифровой среды.

Выборка. Участники исследования включали более 2 300 учеников третьего класса г. Красноярска. Выборка не была репрезентативной, но охватывала разные районы города и школы с разным социально-экономическим положением и образовательными ресурсами.

Инструмент. Для измерения смыслового чтения учащихся в конце начальной школы – в третьем классе – мы используем инструмент ПРОГРЕСС. Для разработки инструмента были проанализированы существующие подходы к пониманию и оценке читательской грамотности и смыслового чтения. Смысловое или семантическое чтение – один из основных метакогнитивных навыков, которым все ученики должны овладеть на достаточном уровне к концу начальной школы. Учитывая международный опыт в области крупномасштабных образовательных исследований на уровне начальной школы, как концептуальная основа для оценки смыслового чтения была использована теоретическая рамка PIRLS (Progress in International Reading Literacy Study). В соответствии с принципами измерений Джорджа Раша [6] сложность заданий для текстов теоретически планировалась заранее и изначально включалась в задания на основе изучаемых в научной литературе прогнозируемых выше характеристик. В результате этой работы был создан инструмент ПРОГРЕСС, где тест по оценке смыслового чтения включал один информационный текст (рассказ о том, как ребенок обнаруживает упавшую птицу и пытается выяснить и применить способы помощи ей) и 19 вопросов, которые задаются по частям после каждого нового экрана, содержащего часть текста, предъявляемого ребенку на компьютере.

Процедура. Учащиеся третьих классов оценивались во фронтальном формате, находились в тихой комнате за персональным компьютером с клавиатурой в присутствии одного взрослого, который был специально проинструктирован оказывать им любую необходимую техническую поддержку, а также предотвращать шумное поведение во время тестирования. Оценивание длилось около 40 минут (включая инструкции).

Оценка читательских действий осуществлялась группой экспертов в сфере начального образования. Два эксперта независимо друг от друга определили (когнитивные) читательские действия (сформулированные

на основе анализа литературы), необходимые для каждого задания данного теста, а затем согласовали свои результаты с третьим экспертом.

В качестве *аналитического подхода* мы использовали семейство линейно-ограниченных экспланаторных моделей (Latent Linear Test Models, LLTM) современной теории тестирования [7]. Эти модели подразумевают, что параметры трудности заданий являются результатом линейной комбинации параметров трудности атрибутов заданий – например, когнитивных операций, необходимых для решения задания, формата ответа и презентации информации, и т. д. Использование этой информации в психометрическом моделировании адресует к использованию коллатеральной информации – побочной информации о измерениях, использование которой в психометрическом моделировании не меняет интерпретации оценок параметров, но уменьшает ошибку их измерения [8]. Соответственно, коллатеральная информация о заданиях позволяет более точно вычислить параметр трудности задания.

Мы использовали вариацию классической модели LLTM – LLTM+e – которая может декомпозировать трудность задания не только на трудности его атрибутов, но и на некоторый индивидуальный для задания компонент, не объясняемый атрибутами. Эта модель допускает, что этот индивидуальный компонент распределен нормально со средним, равным нулю и оцениваемой дисперсией.

Результаты. Описание рассчитанных моделей приведено в табл. 1.

Таблица 1

Рассчитанные модели современной теории тестирования

Параметр модели		Раш-модель	LLTM+e1	LLTM+e2	LLTM+e3	
Дисперсия заданий (e)		1.261	0.758	0.545	0.346	
Легкость атрибута	Уровни рамки	Ур. 1		0.274	-0.932	
		Ур. 2		1.329**	0.254	
		Ур. 3		-0.367	-0.369	
		Ур. 4		-0.790	-0.280	
	Формат ответа	Выбор ответа			1.581***	1.925***
		Открытый ответ			1.103*	1.312**
		Выделение			1.390***	1.874***
	Сложность лексики				-0.386	-0.220
	Объем текста для ответа				-0.391**	-0.352
	AIC		48819.6	48818.0	48813.7	48813.1
BIC		48837.1	48870.3	48874.7	48909.0	
Объясненная дисперсия заданий		0%	40%	57%	73%	

Уровень значимости: 0 < *** < 0.001 < ** < 0.01 < * < 0.05

Обсуждение результатов. В этом исследовании была сделана попытка декомпозировать трудность заданий теста на оценку чтения и объяснить ее через эффекты групп читательских действий, заложенных в концептуальную рамку инструмента. Было выдвинуто предположение,

что трудность заданий теста определяется следующими характеристиками: формат заданий компьютерного инструмента оценки, уровни читательских действий согласно концептуальной рамке инструмента, а также дополнительные параметры, такие как трудность лексики в вопросах теста и объем текста, к которому должен обращаться ребенок при ответе на задания теста. Анализ показал, что наибольшую трудность заданию дает больший объем текста, к которому обращается ребенок при ответе, однако этот эффект перестает быть значимым при контроле уровней концептуальной рамки. В то же время формат задания в форме выбора одного ответа позволяет сделать задание проще. Иерархия уровней читательских действий согласно рамке PIRLS не соответствует заявленной иерархии.

Выявление процессов, лежащих в основе понимания прочитанного, необходимо для объяснения природы чтения и помощи учителям в использовании наиболее эффективных методов обучения. LLTM может предоставить аргументированное объяснение, почему тестируемые определенным образом отвечают на задания той или иной трудности. Данный тип моделирования позволяет декомпозировать стратегии, процессы и действия, необходимые для успешного выполнения каждого задания теста [9]. Определив компоненты, которые усложняют или упрощают задания на понимание прочитанного, исследователи и разработчики теста могут корректировать трудность, а также создавать инструменты с заданным уровнем сложности и спецификой функционирования.

Список литературы

1. Validation of automated scoring of science assessments / Liu O.L. [et al.] // *Journal of Research in Science Teaching*. 2016. № 53(2). P. 215–233.
2. Kane M.T. Current concerns in validity theory // *Journal of Educational Measurement*. 2001. №38(4). P. 319-342.
3. Martin M.O., Mullis I.V.S., Foy P. Assessment design for PIRLS, PIRLS literacy, and ePIRLS in 2016 // *PIRLS 2016 Assessment Framework*. 2nd Ed. Chestnut Hill, 2015. Ch. 3. P. 55–88.
4. Mullis I.V.S., Mullis I.V.S., Hooper M. (Eds.). Prendergast C.O. Developing the PIRLS 2016 achievement items // *Methods and procedures in PIRLS 2016*. Chestnut Hill, 2017. Ch. 1. P. 1–29.
5. Kuznetsova M.I. Features of reading literacy of Russian primary students obtained from PIRLS results // *SHS Web of Conferences*. EDP Sciences. 2016. № 29. P. 1042.
6. Bond T.G., Fox C.M. *Applying the Rasch model: Fundamental Measurement in the Human Sciences* (2nd Ed.). New York: Psychology Press, 2007. 360 p.
7. Fischer G.H. Logistic latent trait models with linear constraints // *Psychometrika*. 1983. №48(1). P. 3–26.
8. Федерякин Д.А., Угланова И.Л., Скрябин М.А. Новые источники информации в компьютерном тестировании // *Вестник Томского государственного университета* (в печати).
9. Embretson S.E. Construct validity: Construct representation and nomothetic span // *Psychological Bulletin*. 1983. № 93. P. 179–197.