

Цифровое тестирование смыслового чтения обучающихся начальной школы: анализ индикаторов поведения

Бакай Е.А.

ФГАОУ ВО «Научно-исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-7513-5554>, e-mail: ebakay@hse.ru

Юсупова Э.М.

ФГАОУ ВО «Научно-исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0002-8332-681X>, e-mail: eabdurakhmanova@hse.ru

Антипкина И.В.

ФГАОУ ВО «Научно-исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация
ORCID: <https://orcid.org/0000-0003-4865-3433>, e-mail: iantipkina@hse.ru

Представлены результаты исследования, посвященного поиску поведенческих индикаторов, которые могут говорить о результатах применения учениками их метакогнитивных умений во время выполнения теста читательской грамотности. Отмечается, что для этого были разработаны индикаторы тестового поведения, рассчитанные на данные о тайминге и операциях во время прохождения теста. К таким индикаторам относятся: время чтения, время взаимодействия с заданиями, умение вернуться к тексту и заданиям для проверки информации и т.п. На основе анализа поведения 2168 четвероклассников во время выполнения теста смыслового чтения были проанализированы два типа индикаторов тестового поведения, которые могут говорить о степени сформированности метакогнитивных умений: контроль скорости и равномерности деятельности в процессе прохождения теста и вариативность передвижения по тесту (наличие нелинейных переходов по заданиям и возвратов к стимульному тексту). С помощью применяемых индикаторов и кластерного анализа были выделены группы учащихся с типичными паттернами поведения, которые были описаны с привлечением информации о поле обучающегося и его балле за тест. Результаты показали, что наличие нелинейных переходов по тесту позволяет выделить группы обучающихся с трудностями регулятивного характера, однако наличие или отсутствие нелинейных переходов по тесту не связано с тестовыми результатами или с полом респондентов. Более высокие результаты обучающихся связаны с равномерным прохождением теста и умеренно большим временем чтения.

Ключевые слова: тесты смыслового чтения; оценивание читательских умений; тестовое поведение; метакогнитивные умения; читательские стратегии; тестовые стратегии.

Финансирование. Статья подготовлена в рамках гранта, предоставленного Министерством науки и высшего образования Российской Федерации (№ соглашения 075-15-2022-325 от 25.04.2022).

Для цитаты: Бакай Е.А., Юсупова Э.М., Антипкина И.В. Цифровое тестирование смыслового чтения обучающихся начальной школы: анализ индикаторов поведения // Психологическая наука и образование. 2024. Том 29. № 2. С. 32—49. DOI: <https://doi.org/10.17759/pse.2024290203>

Online Reading Comprehension Assessment of Primary School Students: Analysis of Testing Behavior Indicators

Ekaterina A. Bakay

National Research University Higher School of Economics, Moscow, Russia
ORCID: <https://orcid.org/0000-0002-7513-5554>, e-mail: ebakay@hse.ru

Elen M. Yusupova

National Research University Higher School of Economics, Moscow, Russia
ORCID: <https://orcid.org/0000-0002-8332-681X>, e-mail: eabdurakhmanova@hse.ru

Inna V. Antipkina

National Research University Higher School of Economics, Moscow, Russia
ORCID: <https://orcid.org/0000-0003-4865-3433>, e-mail: iantipkina@hse.ru

Test-taking behavior in reading comprehension tests is explained, among other things, by meta-cognitive skills, which can be described by indicators of test behavior. These indicators include reading time, respondent-item interaction time, returning to the text, etc. This study is based on the analysis of the test-taking behavior of fourth graders (N=2168) during a reading comprehension test. Two types of indicators of test behavior were described: 1) timing and action speed control; and 2) nonlinear test navigation from item to item and returns to the text. Based on these indicators and using cluster analysis, we identified groups of students with typical patterns of behavior. These groups were described in the context of student gender and test score in the test. The results showed that the data about nonlinear transitions helps to find out the groups of students with regulatory difficulties, however the presence or absence of nonlinear transitions is not associated with test results. Higher results of students are associated with taking test in a linear pattern and a moderately long time.

Keywords: reading comprehension tests; reading assessment; test-taking behavior; metacognitive skills; reading strategies; test strategies.

Funding. The article was prepared in the framework of a research grant funded by the Ministry of Science and Higher Education of the Russian Federation (grant ID: 075-15-2022-325).

For citation: Bakay E.A., Usupova E.M., Antipkina I.V. Online Reading Comprehension Assessment of Primary School Students: Analysis of Testing Behavior Indicators. *Psikhologicheskaya nauka i obrazovanie = Psychological Science and Education*, 2024. Vol. 29, no. 2, pp. 32—49. DOI: <https://doi.org/10.17759/pse.2024290203> (In Russ.).

Введение

Понимание прочитанного является важной предпосылкой образовательного успеха в дальнейшем обучении [25; 35]. В контексте современных исследований чтения как базового метапредметного конструкта проблема точной оценки читательской грамотности и объективной интерпретации результатов стала еще более актуальной [2].

Комплексность рассматриваемого конструкта объясняется тем, что чтение включает в себя непосредственно читательскую и стратегическую компетентности. Под стратегической компетентностью понимается умение респондента проводить перспективную и ситуативную рефлексию своей читательской деятельности и адаптировать взаимодействие с тестом для улучшения результата [39].

Необходимо оговорить использование терминологии. В российской практике оценивания есть школа, где термин «тест» относится только к инструментам оценивания, состоящим из заданий закрытого типа (с выбором одного или нескольких верных ответов) [9]. Однако вслед за международной практикой мы будем использовать данный термин как «инструмент оценивания» (EFPA), таким образом, в тесте могут присутствовать задания как закрытого типа, так и открытые, со свободно конструируемым ответом [21].

Таким образом, тестовое поведение — это совокупность всех взаимодействий респондента с заданиями с различающимися характеристиками в процессе выполнения теста. При этом действия респондента могут быть осознанными, произвольными либо спонтанными, не основанными на планировании и контроле.

Типичные паттерны действий в ходе тестового поведения мы будем называть тестовыми стратегиями [15]. Осознанные, продуманные, основанные на регулятивных умениях и рефлексии тестовые стратегии являются проявлением метакогнитивных умений, формирование которых является актуальной задачей развития регулятивно-мотивационной сферы обучающихся.

Актуальность исследования связана с тем, что установление дидактической парадигмы

субъектности обучающихся привлекло внимание к изучению и принятию во внимание индивидуальных характеристик обучающихся. Предлагаемое в статье решение основано на использовании процессно-ориентированного подхода, в рамках которого анализируются все операции респондентов в контексте их длительности, описывается картина управления читательской деятельностью, которая позволит сделать выводы об условиях, в которых применялись читательские умения, или условиях, ограничивающих их применение.

Настоящее исследование проведено с целью расширить дидактические возможности тестов чтения и предоставить всем сторонам оценивания информацию не только о том, какой результат был достигнут, но и как он был получен.

Теоретически изменение тестового поведения респондентов в процессе прохождения теста объясняют разными факторами. Так, теория ценности и ожиданий (value expectancy theory) описывает связь между восприятием респондентом тестовой ситуации и его действиями через прогнозирование вероятности успеха и субъективной значимости в случае верного решения, а также через оценку необходимых «затрат» и ожидаемых трудностей [18; 19]. Таким образом, теоретически с поведением респондента связаны два фактора: 1) его убежденность в собственной «предметной» готовности к решению данного задания с данными характеристиками; 2) его представления о значимости решения данной учебной задачи.

Первый фактор в свою очередь состоит из: 1) субъективной оценки характеристик задания, 2) оценки собственной готовности по отношению к данному заданию. Например, факт того, что респондент пропускает задание с большим количеством текста, может говорить о том, что оно показалось слишком сложным. О низкой оценке собственной готовности может говорить факт быстрого гадания в сложных заданиях и его отсутствие в относительно легких (здесь под «сложным» понимается не реальная трудность, а воспринимаемая сложность, абстрактность, большое количество этапов и т.д.).

Второй фактор (представление о значимости учебной задачи) складывается из: 1) эмоционального отношения к результату (близко к внутренней мотивации), 2) восприятия «полезности результата» для ближайшего будущего (близко к внешней мотивации), 3) «важности результата» (например, приближает ли респондента успех в данном задании к образу того учащегося, которым он себя представляет), 4) оценки количества ресурсов, требуемых для решения данного задания (в том числе времени) [18; 27].

Изменчивость взаимодействия учащихся с заданиями теста объясняется на основании концепции языковой компетенции Бахмана и Палмера [11, с. 62]. Эта концепция основана на том, что использование речи обуславливается взаимодействием стратегической компетентности респондента и тестовой ситуации. Комплексность этого взаимодействия включает: 1) непосредственно языковые знания, 2) фоновые знания, 3) индивидуальные особенности и аффективные механизмы, которые ответственны за оценку тестовой ситуации. В прошлых исследованиях было показано, что воспринимаемая учеником трудность задания связана с форматом, наличием в тексте задания трудной лексики, соотносённостью задания с когнитивными операциями, соотносённостью с длиной отрывка, к которому относится задание на чтение, и другими факторами [2; 3].

Важно отметить, что помимо стратегий, основанных на применении предметных и метакогнитивных умений (подробнее о них — ниже), существуют и «стратегии предприимчивости»: это попытки использовать фоновые знания там, где это не предусмотрено заданием, попытки угадывать ответ [15].

Во ФГОС НОО и ФГОС ООО закреплено требование развития универсальных учебных действий (УУД). В научной зарубежной литературе наиболее близким аналогом регулятивного компонента УУД являются метакогнитивные умения. Сформированные регулятивные действия обучающихся отвечают за контроль и оценку учебной деятельности, умение выявлять собственные затруднения и их причины, применять интеллектуальную

рефлексию и т.д. Именно такое «отслеживание обучающимися процесса и результатов собственной познавательной активности в решении задач учения» является метакогнитивной активностью [6].

Изначально термин «метакогнитивные умения» описывал «процесс тщательного, осознанного мышления, когда каждое действие требует предварительного планирования и последующей оценки; где решения и операции одновременно весомы и рискованны» [22]. После многих лет исследований термин также используется для обозначения рефлексии процесса работы с информацией и управления этим процессом [40]. В контексте тестов смыслового чтения он рассматривается как сознательная деятельность респондентов, направленная на контроль понимания прочитанного и комплексное прохождение теста [16]. Метакогнитивные умения являются «ядром стратегической компетенции», потому что помогают выявить неэффективность деятельности ученика и активизировать произвольное применение эффективных стратегий [11].

Метакогнитивные умения значимы при выполнении тестов — применяя разнообразные стратегии, учащиеся исключают неверные варианты, ищут подсказки в формулировках вопросов, возвращаются для коррекции ответов [24]. Вариативность метакогнитивных стратегий обеспечивает разнообразие взаимодействий с тестовыми заданиями различных форматов (например, работа с текстом, рисунками, графиками) [20]. Более того, там, где метакогнитивному развитию обучающихся уделялось больше внимания, где были реализованы специальные практики, обучающиеся использовали различные рефлексивные практики во время взаимодействия с учебными задачами для их рационального решения [28].

В контексте смыслового чтения была доказана связь метакогнитивных умений с пониманием прочитанного [13; 14; 30; 34; 43]. Читатели с высоким уровнем способности понимать его активно и произвольно используют различные стратегии чтения в тех ситуациях, где их метакогнитивные умения позволяют выявить дефицит понимания [44],

опосредуют влияние мотивационных факторов на предметные способности [38].

Структура метакогнитивных умений перекликается с описанными ранее факторами изменчивости тестового поведения. Первый компонент метакогнитивных умений — декларативный, это внеситуативное суждение респондента о собственных способностях и индивидуальных характеристиках при обучении, чтении, запоминании и т.д. [23]. Второй компонент — процедурный (операционный), он характеризует метакогнитивные процессы непосредственно во время работы над конкретной задачей [32].

Декларативный компонент оценивается методом опроса [38]. Для его исследования разработано и апробировано несколько достаточно объемных опросников [23; 41; 44]. Анализ их результатов показал наличие у декларативного компонента трех факторов: планирование (достижение заранее поставленных целей, умение планировать ресурсы, в том числе время); мониторинг (наблюдение за процессом, предсказание результата); контроль (регуляция процесса, например, выделение большего количества времени или повторная работа) [33]. Исследования декларативного компонента показали, что адекватная оценка своих рефлексивных умений в области чтения является предиктором понимания [20; 39]. Ограничением исследований декларативного компонента метакогнитивных умений являются небольшой размер выборок (до 200 человек), а также привлечение только хорошо читающих респондентов: учащихся старших классов или взрослых.

Что касается младших школьников, то есть предположение, что контроль над познавательными умениями появляется в возрасте 8—10 лет и связан с умением принимать решения в нестандартной ситуации, что безусловно повышает уровень функциональной грамотности школьников [10]. Другие исследования также подтверждают, что рост осведомленности о метакогнитивных стратегиях происходит с возрастом [12].

Процедурный компонент специфичен для каждого конкретного конструкта [41]. Некоторые исследования процедурного компо-

нента в чтении были основаны на фиксации движений глаз, и именно тайминг стал в них ключевым параметром [38]. Ограничениями подхода с фиксацией движений глаз являются необходимость применения сложного оборудования, немассовость выборок и неоднозначность выводов вследствие зашумления данных. Другой, более пригодный для массового использования подход снова использует тайминг операций — это фиксация времени взаимодействия с фрагментом текста, в котором исследователями были специально заложены несоответствия, или фиксация времени взаимодействия с вопросом после прочитанного [41].

Читатели с высоким уровнем понимания чаще перечитывают сложные отрывки текста во время чтения [42]. Перечитывание является одним из видов действий, которые связаны с метакогнитивными умениями, потому что именно выявление дефицитов понимания ведет к адаптации читательской деятельности, к использованию новых действий, направленных на повышение эффективности деятельности. Существует неоднозначность в использовании исследователями термина «когнитивные умения» (или «когнитивные стратегии»). Под ними понимаются действия, используемые читателями в тестах читательской грамотности для компенсации дефицита понимания или преодоления затруднений при понимании: например, выявление значимой информации, поиск связей между фрагментами текста, выделение резюмирующих предложений, перечитывание и поиск деталей и т.д. Неоднозначность заключается в том, что одни и те же поведенческие индикаторы (например, перечитывание) являются свидетельством как «когнитивного умения» заставить себя вернуться к трудному фрагменту, так и метакогнитивного умения заметить у себя недостаток понимания. Теоретически когнитивные умения являются проявлением сформированных метакогнитивных умений — это те действия, которые предпринимаются респондентами для реализации «метакогнитивного» плана [11]. Но в практике оценивания невозможно или очень трудно найти индикаторы, которые были бы

строго связаны только с метакогнитивными либо только с когнитивными умениями.

Таким образом, обзор литературы показал, что факторами изменчивости тестового поведения являются характеристики респондента, характеристики заданий и их взаимодействие.

Организация исследования, применяемые методики и гипотезы

Задачей нашего исследования было проанализировать «цифровые следы» (тайминг, переходы по тесту) обучающихся во время прохождения цифровой диагностики смыслового чтения с целью выявления информативных индикаторов индивидуальных особенностей тестового поведения респондентов при взаимодействии с заданиями и стимульным материалом теста.

Выборка: исследование проводилось весной 2022 года, в нем приняли участие 2187 обучающихся 4-х классов школ сибирского города-миллионника. После удаления профилей респондентов, которые пропустили 5 и более заданий, в итоговый анализ были включены 2168 наблюдений. Еще 15 респондентов имели от 1 до 4 пропущенных заданий в конце теста (можно интерпретировать как то, что им не хватило времени), они были оставлены в анализе.

Инструмент оценивания: для оценивания использовался инструмент «Прогресс», разработанный в Центре психометрики и измерений в образовании Института образования НИУ ВШЭ [1]. В качестве стимульного материала использовались специально разработанный для оценивания художественный текст и 23 вопроса на проверку понимания прочитанного. Психометрический анализ теста проводился с применением модели современной теории тестирования Partial credit model, которая относится к семейству моделей Раша, в ПО Winsteps [26; 29]. Тест показал существенную одномерность: собственное значение первого контраста при анализе остатков составило 1,5, что меньше конвенционального критерия 2,0 [36]. Надежность теста Person reliability составила 0,78/0,81 (реальная/модельная), что позволяет выделить три различающиеся по

уровню способности группы респондентов. Надежность альфа Кронбаха составила 0,8, что говорит о высокой внутренней согласованности заданий. Все задания теста находятся в согласии с моделью — взвешенные среднеквадратичные статистики согласия всех заданий меньше конвенционального критерия 1,3 [26]. Задания имеют диапазон трудностей от $-2,32$ логита до $2,13$ логитов, что позволяет оценивать респондентов с низкими, средними и высокими результатами. Распределение ответов респондентов схоже с нормальным.

Процедура оценивания: Работа над тестом была ограничена 40 минутами. Оценивание проходило в компьютеризированной форме в компьютерных классах школ. Деятельность учеников во время выполнения тестов была полностью самостоятельной. Респонденты имели возможность пропускать задания без ввода ответа или пролистывать текст, возвращаться к любому этапу, корректировать ответы.

Гипотезы исследования:

1. Индикаторы тестового поведения (индикаторы равномерности прохождения заданий теста, индикаторы линейности выполнения заданий теста и возвраты к чтению текста) позволяют выделить группы учащихся со схожими паттернами тестового поведения.
2. Паттерны тестового поведения связаны с предметными (читательскими) результатами учащихся.

Процесс определения типичных паттернов

Для работы по выявлению типичных паттернов мы использовали две группы индикаторов, удобных для доступа в лог-журнале инструмента.

Индикаторы линейности

Это данные о возвратах к тексту и заданиям. Линейным прохождением теста мы называем последовательное (без возвратов, пропусков) прохождение теста с однократным обращением к тексту и последующим заданиям. Нелинейное прохождение теста — это произвольный пропуск заданий,

непоследовательный переход по заданиям (переход с первого задания на пятое и т.п.), возвраты к предыдущим заданиям, а также возвраты к тексту.

Нелинейность при прохождении теста может свидетельствовать о применении метакогнитивных умений (например, оценить предстоящий объем работы перед началом выполнения теста, пропустить слишком трудные задания, чтобы вернуться к ним потом или вернуться к тексту для уточнения прочитанного).

Для целей анализа из переменной, описывающей количество возвратов к тексту, мы выделили три фиктивных переменные, закодированные 1/0 (1 — респондент относится к этой переменной, 0 — не относится к этой переменной): 1) ноль возвратов к тексту (1469 респондентов); 2) 1 возврат к тексту (373 респондента); 3) 2 и более возвратов к тексту (329 респондентов).

На основании переменной, описывающей количество возвратов к заданиям, также были созданы три фиктивных переменные, закодированные 1/0: 1) количество нелинейных переходов по заданиям равно 0 (825 человек); 2) количество нелинейных переходов по заданиям составляет 1—5 (952 человек); 3) количество нелинейных переходов по заданиям составляет 6 и более (394 человек). Выбор критериев упрощения данных определялся целями достижения контрастности групп.

Индикаторы равномерности

Мы предположили, что важными характеристиками поведения респондентов будут как абсолютные, так и относительные показатели времени. Инструмент разрабатывался таким образом, чтобы чтение текста не занимало более трети отведенного на весь тест времени. Это было проверено на когнитивных лабораториях. Мы также предполагали, что в процессе работы над тестом трудность заданий и значимость учебной задачи теста могут варьироваться в зависимости от усталости и ситуативной мотивации участников теста. Именно динамичность вовлеченности диктовала необходимость использовать в анализе три переменных времени: 1) время (длитель-

ность) чтения текста (стимульного материала); 2) время прохождения первой половины заданий теста (это этап знакомства с форматом заданий, оценка и адаптация к трудности заданий, к тестовой ситуации, к самостоятельной форме работы); 3) время прохождения второй половины заданий теста (во второй половине теста было больше заданий на более высокие уровни когнитивных умений: поиск информации, не данной в явном виде, анализ и синтез). Конечно, вопросы теста не могут быть одинаковы по предполагаемому времени выполнения, но в целом первая и вторая половина теста должны быть сбалансированы по времени выполнения. Одно задание, которое было заданием открытого типа и могло требовать существенно большего времени, было исключено из анализа. Таким образом, первая и вторая половина теста состояли из 11 заданий закрытого типа.

Если первые три переменные из этой группы индикаторов описывают абсолютное время работы учащихся в «начале, середине и конце теста», то следующие два показателя характеризуют динамику длительности взаимодействия с заданиями. Поскольку не существует нормативов по времени взаимодействия с заданиями теста смыслового чтения, мы использовали для анализа длительность взаимодействия с заданиями относительно среднего значения по выборке.

Мы выделяем очень быстрое относительно всей выборки время выполнения задания (для удобства названное «быстрые всплески») и очень долгое время взаимодействия с заданиями теста («медленные всплески»). Если время выполнения конкретного задания ученика меньше -1 стандартного отклонения времени по всей выборке, то это «быстрый всплеск». Если время выполнения конкретного задания ученика больше $+1$ стандартного отклонения времени по всей выборке, то это «медленный всплеск» — задание выполнялось очень долго.

«Быстрые» и «медленные» всплески анализировались отдельно по первой и второй половине теста.

В табл. 1 представлено описание переменных равномерности выполнения теста:

времени взаимодействия с текстом, первой половиной заданий и второй половиной заданий и количество «быстрых» и «медленных» всплесков в первой и второй половинах заданий.

Для выделения тестовых стратегий был проведен иерархический кластерный анализ [31]. Достоинствами этого метода является устойчивость результатов. Анализ проходил в два этапа: работа с группой индикаторов линейности и работа с группой индикаторов равномерности. Индикаторы линейности включали в себя дихотомические переменные, поэтому для расчета расстояний между наблюдениями использовалось бинарное расстояние, которое рассчитывается как доля элементов, которые являются общими для двух переменных. Для расчета расстояний между наблюдениями второй группы индикаторов равномерности использовалась метрика Евклида. Метод Варда, который позволяет выделять кластеры так, чтобы минимизировать внутригрупповую дисперсию, использовался как метод агломерации.

Различия в средних баллах по тесту между группами проверялись с помощью критерия Краскела-Уоллиса, а попарные различия — с помощью рангового критерия Вилкоксона с поправкой Бонферрони на множественные сравнения. Решение использовать непараметрические методы было основано на том, что требования для использования однофакторного дисперсионного анализа не выполнялись во всей строгости. Проверка равенства распределения учащихся по полу в кластерах проверялась с помощью двухпропорционального z-теста.

Предварительно все количественные переменные, включенные в анализ, были стандартизированы со средним 0 и стандартным отклонением 1.

Анализ был проведен с использованием пакета stats (версия 4.2.2) и rstatix (версия 0.7.2.) в среде программирования R (версия 4.2.2).

Результаты

Для более однозначной интерпретации результатов обе группы индикаторов анализировались отдельно.

Этап 1. Выделение групп учащихся с использованием индикаторов равномерности

На рис. 1 представлена дендрограмма, полученная по итогам кластерного анализа для группы индикаторов равномерности.

На основании баланса расстояний было выбрано решение из шести кластеров.

В табл. 2 представлены средние кластеров.

На рис. 2 показано, как распределены предметные результаты оценивания чтения (z-баллы) по тесту по кластерам.

Критерий Краскела-Уоллиса показал наличие статистически значимых различий в тестовых баллах между группами (Хи-квадрат=246,78, df=5, p-value<0,001). Результаты попарного сравнения не выявили различий в тестовых баллах только между кластерами 1 и 3 (p=0,437), 2 и 6 (p=0,198) (рис. 3).

В табл. 3 показано, как распределились наблюдения по полу. Оценивание кластеров с привлечением информации о поле уче-

Таблица 1
Время чтения текста, время выполнения первой и второй половины заданий

Переменная	Среднее	Станд. откл.	Медиана	Мин.	Макс.	Межквартильный размах
Общее время теста	21,31	6,27	20,87	1,62	43,05	8,28
Время чтения текста	6,54	3,09	6,58	0,03	24,10	3,40
Время заданий (1 часть)	8,04	3,15	7,42	0,93	23,18	3,77
Время заданий (2 часть)	6,73	2,43	6,49	0,57	18,78	3,00
«Быстрые всплески» (1 часть)	0,58	0,97	0,00	0,00	9,00	1,00
«Медленные всплески» (1 часть)	1,01	1,37	1,00	0,00	9,00	2,00
«Быстрые всплески» (2 часть)	0,70	1,65	0,00	0,00	11,00	1,00
«Медленные всплески» (2 часть)	1,21	1,55	1,00	0,00	10,00	2,00

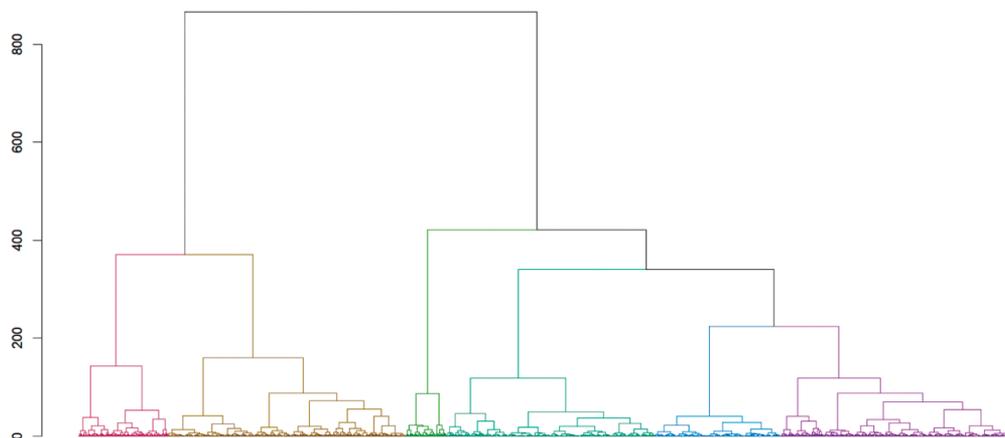


Рис. 1. Дендрограмма иерархического кластерного анализа

Таблица 2

Средние кластеров

Кластер	Число наблюдений	Время (мин)			Всплески			
		Текст	Задания (1 часть)	Задания (2 часть)	Быстрые (1 часть)	Медленные (1 часть)	Быстрые (2 часть)	Медленные (2 часть)
1	482	5,60	5,50	5,42	1,43	0,21	0,46	0,38
2	536	5,96	8,92	6,09	0,24	1,37	0,77	0,69
3	294	6,46	6,14	5,58	0,00	0,01	0,25	0,17
4	554	8,55	8,43	8,89	0,28	0,95	0,09	2,44
5	95	2,99	5,63	2,42	2,52	0,43	6,71	0,20
6	207	6,64	14,41	9,29	0,24	3,80	0,57	3,17

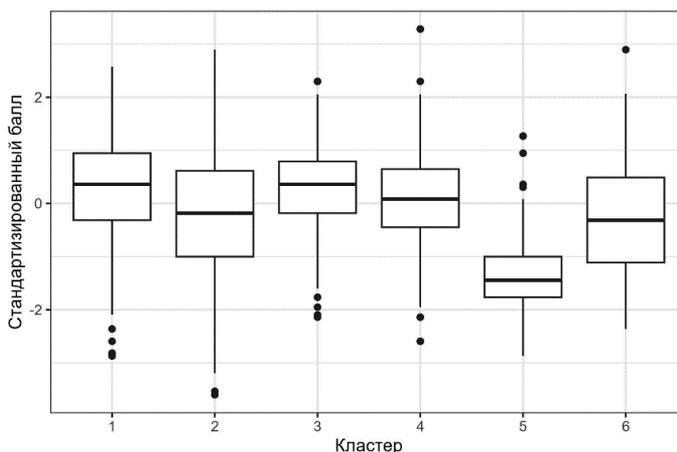


Рис. 2. Распределение тестовых баллов учащихся по кластерам, выделенным по индикаторам равномерности

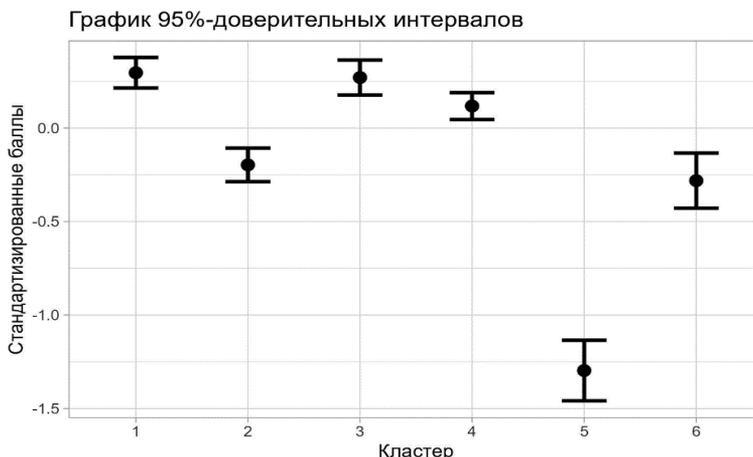


Рис. 3. График 95%-доверительных интервалов средних баллов в кластерах, выделенны с помощью индикаторов равномерности

ников проводилось, поскольку в контексте изучения индивидуальных особенностей ранее было показано, что мальчики более нетерпеливы, нуждаются в смене деятельности, активном характере действий [8], они также склонны оценивать свои интеллектуальные качества и успехи в учебной деятельности выше, чем девочки [5]. Например, в PIRLS-2018, несмотря на повышения качества чтения среди мальчиков по сравнению с предыдущими исследованиями, девочки все же показали более высокие результаты.

Эти же выводы подтвердили и российские исследования. Например, показано, что «мальчики 9—10 лет превосходят девочек технически», то есть читают быстрее, однако уступают в понимании прочитанного [4].

В кластерах 3 и 6 девочек статистически значимо больше, чем мальчиков (хи-

квадрат=4,40, $df=1$, $p<0,05$ и хи-квадрат=15,7, $df=1$, $p<0,05$), а в кластере 5 — наоборот, мальчиков статистически значимо больше (хи-квадрат=6,05, $df=1$, $p<0,05$). В остальных кластерах нет различий по распределению по полу.

Этап 2. Выделение групп учащихся с использованием индикаторов линейности

На рис. 4 представлена дендрограмма, полученная по итогам кластерного анализа для группы индикаторов линейности.

Для анализа переменных линейности было выбрано решение из четырех кластеров. В табл. 5 представлены средние кластеров.

На боксплот-графике (рис. 5) показано, как распределены оценки в тесте чтения (z-баллы) по кластерам.

Таблица 3

Доля девочек в кластерах, выделенных с помощью индикаторов равномерности

Кластер	Доля девочек
1	0,49
2	0,52
3	0,59
4	0,54
5	0,33
6	0,61

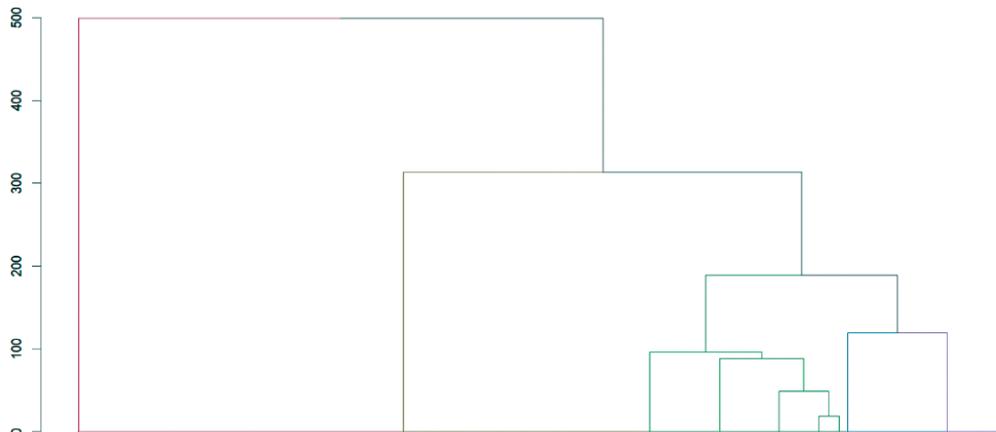


Рис. 4. Дендрограмма иерархического кластерного анализа, проведенного на индикаторах линейности

Таблица 5

Средние кластеров по индикаторам линейности

Кластер	Число наблюдений	Нет нелинейных переходов по заданиям	От 1 до 5 нелинейных переходов по заданиям	6 и более нелинейных переходов по заданиям	Нет возвратов к тексту	1 возврат к тексту	2 и более возвратов к тексту
1	378	0	1	0	0	0,61	0,39
2	573	0	1	0	1	0	0
3	460	0,14	0	0,85	0,3	0,3	0,40
4	757	1	0	0	1	0	0

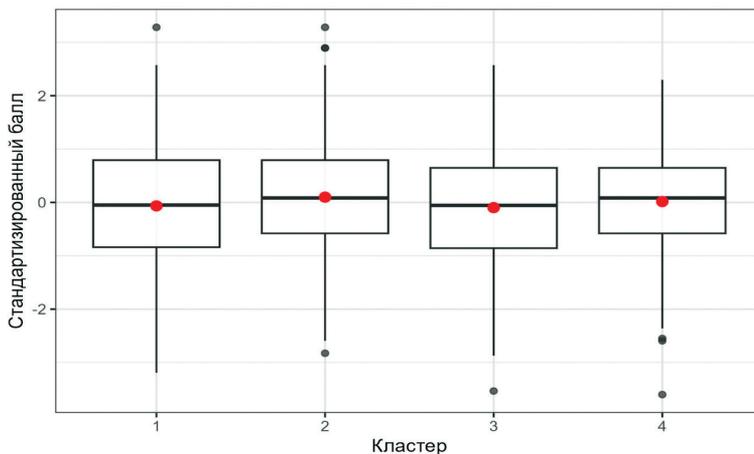


Рис. 5. Распределение тестовых баллов учащихся по кластерам, выделенным по индикаторам линейности

Критерий Краскела-Уоллиса показал наличие статистически значимых различий между группами ($\chi^2=10,05$, $df=3$, $p\text{-value}<0,05$). Результаты попарного сравнения выявили различия в тестовых баллах только между кластерами 2 и 3 ($p=0,018<0,05$) (рис. 3). На уровне тенденции ($p<0,1$) обнаружены различия в тестовых баллах в кластерах 1 и 2, а также 3 и 4 ($p=0,078$).

В табл. 6 показано, как распределились наблюдения по полу.

Анализ пропорций показал, что нет статистически значимых различий в распределении по полу в четырех кластерах.

Интерпретация результатов

На основании анализа кластеров получены следующие результаты.

Кластеры равномерности прохождения

При анализе групп учащихся в контексте линейности переходов были выделены 6 групп.

Первый кластер характеризуется относительно быстрым и равномерным прохождением и наличием «быстрых всплесков» в первой половине теста: учащиеся не испытывали затруднения при взаимодействии с тестом, легко отвечали на начальные, сравнительно менее трудные задания.

Второй кластер характеризуется длительным прохождением и «медленными всплесками» именно в первой половине теста. Можно предположить, что у этих детей возникли затруднения с первоначальным входением в оценивание, адаптацией к тому, как выполнять тест.

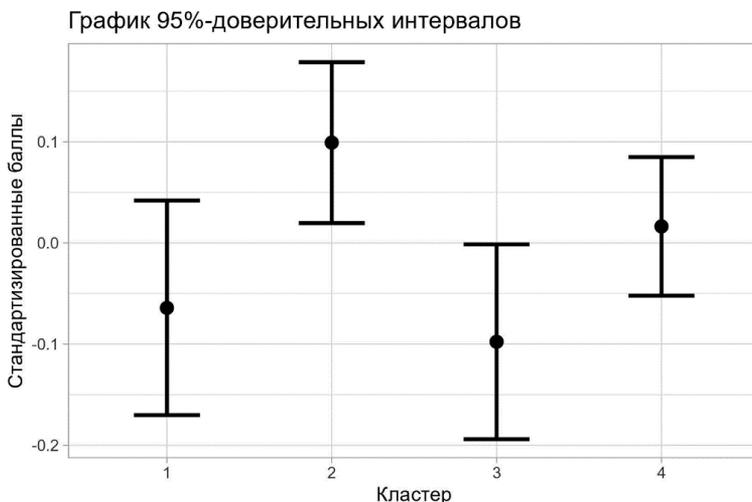


Рис. 7. График 95%-доверительных интервалов средних баллов в кластерах, выделенных с помощью индикаторов равномерности

Таблица 6
Доля девочек в кластерах, выделенных с помощью индикаторов линейности

Кластер	Доля девочек
1	0,51
2	0,53
3	0,52
4	0,54

Третий кластер отличается очень высокой степенью равномерности: в нем собраны учащиеся с меньшим количеством «быстрых» и «медленных» всплесков.

В четвертом кластере при относительно длительном времени выполнения трех частей теста замедление произошло во второй половине теста (возможная интерпретация: усталость).

Пятый кластер включает в себя респондентов с очень маленьким временем выполнения трех частей теста и наибольшим количеством «быстрых всплесков» — это «кликеры», учащиеся, которые не включились в работу над заданием и демонстрировали низкую вовлеченность, отвечали формально, оценивать их читательские умения по результатам теста нельзя, поскольку они не приняли учебную задачу.

Шестой кластер — это учащиеся с максимальным временем работы и «медленными всплесками» в 1 и 2 половине теста, это также учащиеся с самыми низкими (не считая «кликеров») результатами.

Кластеры линейности

При анализе групп респондентов по индикаторам линейности переходов были выделены 4 группы по переходам. Наиболее интересные кластеры 1, 2 и 4. Кластер 4 — это полностью линейный способ выполнения заданий. Кластер 2 описывает паттерн с небольшим количеством переходов по заданиям, но полным отсутствием возвратов к тексту. В кластере 1 собраны учащиеся, которые относительно умеренно перемещались по заданиям и при этом возвращались к тексту. Кластер включает всех остальных учащихся — с достаточно большим количеством переходов по разным категориям («стихийные серферы» — возможно, этим учащимся было трудно разобраться в интерфейсе или в заданиях и их переходы говорят о более или менее неструктурированных попытках сориентироваться с помощью переходов на доступные страницы).

Мы ожидали, что отсутствие возвратов к тексту будет связано с более низкими предметными результатами, и, наоборот,

активная работа с тестовыми заданиями и возвращения к тексту будут связаны с более высоким баллом. Однако полностью линейное прохождение в целом связано с положительными результатами, а у учащихся, возвращавшихся к тексту, в среднем тестовый балл ниже, чем у тех, кто не возвращался.

Выводы и обсуждение результатов

В исследовании были предложены индикаторы тестового поведения, которые могут обогатить обратную связь по результатам предметного оценивания. На примере теста читательской грамотности мы показали, как индикаторы равномерности и индикаторы линейности прохождения теста позволяют сгруппировать учащихся в интерпретируемые кластеры, что является подтверждением гипотезы 1.

Исследование показало наличие дефицитов метапредметных умений (что соответствует возрасту детей). Так, комбинация индикаторов равномерности может быть надежным критерием для выявления «кликеров» — учащихся, предметные результаты которых нет смысла сообщать, поскольку они не выполняли задания.

В начальной школе значительное количество учащихся не возвращаются к тексту, а если возвращаются, то это не связано с более высокими предметными результатами. Наша гипотеза 2, таким образом, не находит подтверждения.

Неожиданным оказался результат о том, что полностью линейное прохождение теста не связано с более низкими предметными результатами. Возможно, в начальной школе сравнительно небольшой объем текста позволяет успешно выполнять задания, опираясь только на однократное прочтение. Однако отсутствие привычки возвращаться к тексту может создать трудности при чтении в средней школе. У тех учащихся, кто возвращались к тексту, средний тестовый балл ниже, что контринтуитивно. Вероятно, в большинстве случаев это были возвраты, не повышающие эффективность чтения.

Более длительное время выполнения теста может говорить не о вдумчивости, а

о трудностях, причем значимо более низкие результаты выявлены и у тех, кто «долго сидел» на первой половине заданий (предположительно: трудности с ориентацией в заданиях) и на второй половине теста (предположительно: усталость, демотивация). И хотя предметные результаты у этих двух кластеров одинаково низкие, работа по преодолению их метакогнитивных трудностей, скорее всего, должна различаться.

Респонденты с самым большим количеством нелинейных переходов («серферы») характеризуются также самым низким тестовым баллом. Это группа риска среди испытуемых. Даже если их активная навигация по тесту вызвана осознанием собственного неуспеха в чтении, то их когнитивные стратегии по адаптации тестового поведения и изменению тестовой траектории не связаны с повышением результатов. Причиной полной линейности выполнения теста или «серфинга» может быть как недостаточный опыт работы с цифровыми инструментами, так и отсутствие осознанной стратегии выполнения теста.

Перспективы исследования

Важной задачей следующего этапа авторы считают исследование стабильности тестового поведения (повторяются ли одни и те же паттерны у одних и тех же учащихся

из теста в тест). Кроме того, очень важно посмотреть на функционирование предложенных индикаторов у учащихся разных возрастов. В этом исследовании также показаны некоторые гендерные различия в тестовом поведении учащихся, согласующиеся с результатами предыдущих исследований. В целом мы считаем важным продолжать поиски интерпретируемых индикаторов тестового поведения, которые могут описывать регуляторную сферу развития учащихся.

Ограничения и допущения исследования

У нас нет возможности жестко связать индикаторы тестового поведения со степенью сформированности метакогнитивных умений, мы можем только предполагать, что эти индикаторы можно интерпретировать в терминах метакогнитивных умений.

К ограничениям исследования относится также то, что мы стремились к упрощению сложных данных о поведении детей при выполнении тестов (например, нелинейные переходы по заданиям оценивались в целом по тесту, независимо от места заданий в тесте). Требуется дополнительная валидизация индикаторов тестового поведения, более детальная их оценка, прежде чем их можно будет рекомендовать как надежную меру, которая может лечь в основу педагогических интервенций.

Литература

1. Бакай Е.А., Юсупова Э.М., Антипкина И.В. Читают или делают вид? Анализ поведения учащихся начальных классов при выполнении заданий теста читательской грамотности // Вопросы образования. 2023. № 1. С. 8—28.
2. Иванова А.Е., Антипкина И.В. Декомпозиция трудностей заданий теста читательской грамотности // Вопросы образования. 2023. № 3. С. 92—112.
3. Иванова А.Е., Антипкина И.В., Федеряки Д.А. Декомпозиция читательских действий в инструменте измерения навыка чтения // В кн.: Информатизация образования и методика электронного обучения: цифровые технологии в образовании. Ч. 1. Красноярск: Сибирский федеральный университет, 2021. С. 189—193.
4. Крещенко О.Ю., Хромова С.К. Гендерные особенности речевого развития и сформированность навыков письма и чтения у детей 9—10 лет [Электронный ресурс] // Новые

- исследования. 2011. № 29. URL: <https://cyberleninka.ru/article/n/gendernye-osobennosti-rechevogo-razvitiya-i-sformirovannost-navykov-pisma-i-chteniya-u-detey-9-10-let> (дата обращения: 19.03.2024).
5. Талакова Е.А. Гендерные особенности я-концепции в процессе личностного становления детей младшего школьного возраста: автореф. дисс. ... канд. психол. наук. Нижний Новгород, 2012. 23 с.
6. Фомин А.Е. Метакогнитивный мониторинг решения учебных задач: психологические механизмы и условия развития в образовательном процессе: дисс. ... докт. психол. наук. М., 2020.
7. Цукерман Г.А., Ковалева Г.С., Кузнецова М.Т. Становление читательской грамотности, или новые похождения тяни-толкая [Электронный ресурс] // Вопросы образования. 2015. № 1. URL: <https://cyberleninka.ru/article/n/stanovlenie-chitatelskoy-gramotnosti-ili-novye-pohozhdeniya-tyani-tolkaya> (дата обращения: 24.08.2023).

8. Шешукова О.В. Гендерные особенности формирования социального интеллекта младших школьников: автореф. дисс. ... канд. психол. наук. М., 2008. 18 с.
9. Шмелев А.Г. Практическая тестология. Тестирование в образовании, прикладной психологии и управлении персоналом. М.: Маска, 2013. 792 с.
10. Akaydin B., Alper, Çokçalışkan H. Investigation of Primary School Students' Metacognitive Awareness and Decision-Making Skill // International Journal of Progressive Education. 2020. P. 16. DOI:10.29329/ijpe.2020.268.10
11. Bachman L.F., Palmer A.S. Language testing in practice. Oxford University Press, 2010.
12. Bakkağolu S. Analysis of Metacognitive Awareness of Primary and Secondary School Students in Terms of Some Variables // Journal of Education and Learning. 2020.
13. Brown A. Metacognition, executive control, self-regulation, and other more mysterious mechanisms / Hillsdale, NJ: Lawrence Erlbaum F. Weinert & R. Kluwe (Eds.) // Metacognition, motivation and understanding. 1987. P. 65—116.
14. Carrell P.L. Metacognitive awareness and second language reading. // The Modern Language Journal. 1989. Vol. 73(2). P. 121—134.
15. Cohen A.D. The coming of age of research on test-taking strategies // Language Assessment Quarterly. 2006. Vol. 3. № 4. P. 307—331.
16. Cohen A.D., Upton T.A. "I want to go back to the text": Response strategies on the reading subtest of the new // TOEFL. Language Testing. 2007. Vol. 24(2). P. 209—250.
17. Diken E.H. Cognitive and metacognitive strategies of 6th-grade students to answer multiple-choice questions on "human body systems" // International Journal of Curriculum and Instruction. 2020. Vol. 12(2). P. 436—456. URL: <http://ijci.wcci-international.org/index.php/IJCI/article/view/435>
18. Eccles J.S., Adler T.E., Futterman R., Goff S.B., Kaczala C.M., Meece J.L. et al. Expectancies, values, and academic behaviors / In J.T. Spence (ed.) // Achievement and Achievement Motivation. 1983. P. 75—146.
19. Eccles J.S., Wigfield A. Motivational beliefs, values, and goals // Annu. Rev. Psychol. 2002. Vol. 53. P. 109—132. DOI:10.1146/annurev.psych.53.100901.135153
20. Edossa A.K., Neuenhaus N., Artelt C., Lingel K., Schneider W. Developmental relationship between declarative metacognitive knowledge and reading comprehension during secondary school // European Journal of Psychology of Education. 2018. Vol. 34. P. 397—416.
21. Evers A., Hagemester C., Høstmælingen A., Lindley P., Muñiz J., Sjöberg A. EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. 2013. Version 4.2, 6.
22. Flavell J.H. Metacognition and cognitive monitoring: A new area of cognitive—developmental inquiry // American Psychologist. 1979. Vol. 34. P. 906—911. DOI:10.1037/0003-066X.34.10.906
23. Flavell J.H., Wellman H. Metamemory / In R. Krail Jr., J. Hegen (Eds.) // Perspectives on the development of memory and cognition. Hillsdale, NJ: Lawrence Erlbaum, 1977. P. 3—33.
24. Hawker M.J., Dysleski L., Rickey D. Investigating general chemistry students' metacognitive monitoring of their exam performance by measuring postdiction accuracies over time // Journal of Chemical Education. 2016. Vol. 93(5). P. 832—840. DOI:10.1021/acs.jchemed.5b00705
25. Korhonen J., Linnanmäki K., Aunio P. Language and mathematical performance: A comparison of lower secondary school students with different level of mathematical skills // Scandinavian Journal of Educational Research. 2012. Vol. 56. P. 333—344. DOI:10.1080/00313831.2011.599423
26. Linacre J.M. A user's guide to WINSTEPS® MINISTEP: Rasch-model computer programs // (No Title). 2016.
27. Lindner M.A., Lüdtke O., Nagy G. The onset of rapid-guessing behavior over the course of testing time: a matter of motivation and cognitive resources // Frontiers in Psychology. 2019. Vol. 10. DOI:10.3389/fpsyg.2019.01533
28. Liu S., Liu M. The impact of learner metacognition and goal orientation on problem-solving in a serious game environment // Computers in Human Behavior. 2020. Vol. 102. P. 151—165. DOI:10.1016/j.chb.2019.08.021
29. Masters G.N. Partial credit model // Chapman and Hall/CRC Handbook of item response theory. 2016. P. 137—154.
30. Mokhtari K., Sheorey R. Measuring ESL students' awareness of reading strategies // Journal of Developmental Education. 2002. Vol. 25(3). P. 2—11.
31. Murtagh F., Contreras P. Algorithms for hierarchical clustering: an overview // WIREs Data Mining Knowl Discov. 2012. P. 86—97. DOI:10.1002/widm.53
32. Nelson T.O., Narens L. Metamemory: A theoretical framework and new findings / Academic Press G. Bower (Ed.) // The psychology of learning and motivation: Advances in research and theory. 1990. Vol. 26. P. 125—141.
33. Paris S.G., Winograd P. How metacognition can promote academic learning and instruction / Hillsdale, NJ: Erlbau B.F. Jones & L. Idol, 1990 (Eds.) // Dimensions of thinking and cognitive instruction. P. 15—51.
34. Phakiti A. A closer look at the relationship of cognitive and metacognitive strategy use to EFL

reading achievement test performance // Language Testing. 2003. Vol. 20(1). P. 26—56.

35. Savolainen H., Ahonen T., Aro M., Tolvanen A., Holopainen L. Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education // Learning and Instruction. 2008. Vol. 18. P. 201—210. DOI:10.1016/j.learninstruc.2007.09.0174

36. Smith Jr.E.V. Detecting and Evaluating the Impact of Multidimensionality using Item Fit Statistics and Principal Component Analysis of Residuals // Journal of Applied Measurement. 2002. Vol. 3. № 2. P. 205—231.

37. Tadayon F., Ravand H. Using grounded theory to validate Bachman and Palmer's (1996) strategic competence in EFL graph-writing // Language Testing in Asia. 2016. Vol. 6. DOI:10.1186/s40468-016-0031-y

38. Tibken C., Richter T., von der Linden N., Schmiedeler S., Schneider W. The role of metacognitive competences in the development of school achievement among gifted adolescents // Child Development. 2022. Vol. 93(1). P. 117—133. DOI:10.1111/cdev.13640

39. Van Kraayenoord C.E. The role of metacognition in reading comprehension // Focal points of the research

and development of pedagogically-psychological perspectives. 2010. P. 277—302.

40. Vandergrift L., Goh C.C.M. Teaching and learning second language listening // Routledge. 2012. DOI:10.4324/9780203843376

41. Van der Stel M., Veenman M.V.J. Metacognitive skills and intellectual ability of young adolescents: A longitudinal study from a developmental perspective // European Journal of Psychology of Education. 2014. Vol. 29. P. 117—137. DOI:10.1007/s10212-013-0190-5

42. Zargar E., Adams A.M., Connor C.M. The relations between children's comprehension monitoring and their reading comprehension and vocabulary knowledge: An eye-movement study // Reading and Writing. 2019. Vol. 33. P. 511—545. DOI:10.1007/s11145-019-09966-3

43. Zhang L.-J. A dynamic metacognitive systems account of Chinese university students' knowledge about EFL reading // TESOL Quarterly. 2010. Vol. 44. P. 320—353.

44. Zhang L. Metacognitive and Cognitive Strategy Use in Reading Comprehension: A Structural Equation Modelling Approach. // Springer Nature Singapore. 2018. Pte Ltd. DOI:10.1007/978-981-10-6325-1

References

1. Bakay E.A., Usupova E.M., Antipkina I.V. Chitayut ili delayut vid? Analiz povedeniya uchashchikhsya nachal'nykh klassov pri vypolnenii zadaniy testa chitatel'skoi gramotnosti [Are they reading or pretending? Analysis of the behavior of primary school students when performing tasks of the reading literacy test]. *Voprosy obrazovaniya = Education issues*, 2023, no. 1, pp. 8—28 (in Russ.).

2. Ivanova A.E., Antipkina I.V. Dekompozitsiya trudnosti zadaniy testa chitatel'skoj gramotnosti [Decomposition of the difficulty of reading literacy test tasks]. *Voprosy obrazovaniya = Education issues*, 2023, no. 3, pp. 92—112 (in Russ.).

3. Ivanova A.E., Antipkina I.V., Federyakin D.A. Dekompozitsiya chitatel'skih deystvij v instrumente izmereniya navyka chteniya [Decomposition of reading actions in the reading skill measurement tool]. *Informatizatsiya obrazovaniya i metodika elektronnoy obucheniya: cifrovye tekhnologii v obrazovanii = Informatization of education and e-learning methods: digital technologies in education*, 2021, pp 189—193 (in Russ.).

4. Kreshchenko O.U., Khromova S.K. Gendernye osobennosti rechevogo razvitiia i sformiro-vannost navykov pisma i chteniia u detei 9—10 let [Gender features of speech development and the formation of writing and reading skills in children aged 9—10 years] [Electronic resource]. *Novye issledovaniia = New research*. 2011, no. 29. URL: <https://cyberleninka.ru>

article n gendernye-osobennosti-rechevogo-razvitiya-i-sformiro-vannost-navykov-pisma-i-chteniya-u-detey-9-10-let (Accessed 19.03.2024) (in Russ.).

5. Talakova E.A. Gendernye osobennosti iakontseptsii v protsesse lichnostnogo stanovleniia detei mladshego shkolnogo vozrasta avtoreferat dis kandidata psikhologicheskikh nauk [Gender features of the self-concept in the process of personal development of primary school children Avtoref. PhD (Psychology) Thesis]. Nizhniy Novgorod, 2013, p. 23. (in Russ.).

6. Fomin A.E. Metakognitivnyj monitoring resheniya uchebnykh zadach: psikhologicheskie mekhanizmy i usloviya razvitiya v obrazovatelnom protsesse [Metacognitive monitoring of solving educational problems: Psychological mechanisms and conditions of development in the educational process]. Dr. Hab. Dis. Moscow, 2020. (in Russ.).

7. Tsukerman G.A., Kovaleva G.S., Kuznetsova M.T. Stanovlenie chitatel'skoi gramotnosti, ili novye pokhozhdeniya tyani-tolkaya [The formation of reading literacy, or new adventures of pulling and pushing]. *Voprosy obrazovaniya = Education issues*, 2015, no. 1. URL: <https://cyberleninka.ru/article/n/stanovlenie-chitatel'skoy-gramotnosti-ili-novye-pokhozhdeniya-tyani-tolkaya> (Accessed 24.08.2023) (in Russ.).

8. Sheshukova O.V. Gendernye osobennosti formirovaniia sotsialnogo intellekta mladshikh shkolnikov avtoreferat dis kandidata psikhologicheskikh nauk [Gender characteristics of the formation of social

- intelligence in primary school children Avtoref. PhD (Psychology) Thesis]. Moscow, 2008, p. 18. (in Russ.).
9. Shmelev A.G. Prakticheskaya testologiya. Testirovanie v obrazovanii, prikladnoi psikhologii i upravlenii personalom [Practical testology. Testing in education, applied psychology and human resources management]. Moscow, 2013. 792 p. (in Russ.).
10. Akaydin B., Alper, Çokçalışkan H. Investigation of Primary School Students' Metacognitive Awareness and Decision-Making Skill. *International Journal of Progressive Education*, 2020. P. 16. DOI:10.29329/ijpe.2020.268.10
11. Bachman L.F., Palmer A.S. Language testing in practice. Oxford University Press, 2010.
12. Bakkaloğlu S. Analysis of Metacognitive Awareness of Primary and Secondary School Students in Terms of Some Variables. *Journal of Education and Learning*, 2020.
13. Brown A. Metacognition, executive control, self-regulation, and other more mysterious mechanisms. Hillsdale, NJ: Lawrence Erlbaum F. Weinert & R. Kluwe (Eds.). *Metacognition, motivation and understanding*, 1987, pp. 65—116.
14. Carrell P.L. Metacognitive awareness and second language reading. *The Modern Language Journal*, 1989. Vol. 73(2), pp. 121—134.
15. Cohen A.D. The coming of age of research on test-taking strategies. *Language Assessment Quarterly*, 2006. Vol. 3, no.4, pp. 307—331.
16. Cohen A.D., Upton T.A. "I want to go back to the text": Response strategies on the reading subtest of the new. *TOEFL. Language Testing*, 2007. Vol. 24(2), pp. 209—250.
17. Diken E.H. Cognitive and metacognitive strategies of 6th-grade students to answer multiple-choice questions on "human body systems". *International Journal of Curriculum and Instruction*, 2020. Vol. 12(2), pp. 436—456. URL: <http://ijci.wcci-international.org/index.php/IJCI/article/view/435>
18. Eccles J.S., Adler T.E., Futterman R., Goff S.B., Kaczala C.M., Meece J.L. et al. Expectancies, values, and academic behaviors // *Achievement and Achievement Motivation*, ed J.T. Spence. 1983. P. 75—146.
19. Eccles J.S., Wigfield A. Motivational beliefs, values, and goals. *Annu. Rev. Psychol.*, 2002. Vol. 53, pp. 109—132. DOI:10.1146/annurev.psych.53.100901.135153
20. Edossa A.K., Neuenhaus N., Artelt C., Lingel K., Schneider W. Developmental relationship between declarative metacognitive knowledge and reading comprehension during secondary school. *European Journal of Psychology of Education*, 2018. Vol. 34, pp. 397—416.
21. Evers A., Hagemester C., Høstmælingen A., Lindley P., Muñiz J., Sjöberg A. EFPA review model for the description and evaluation of psychological and educational tests. Test review form and notes for reviewers. 2013. Version 4.2, 6.
22. Flavell J.H. Metacognition and cognitive monitoring: A new area of cognitive—developmental inquiry. *American Psychologist*, 1979. Vol. 34, pp. 906—911. DOI:10.1037/0003-066X.34.10.906
23. Flavell J.H., Wellman H. Metamemory. R. Krail Jr., J. Hegen (Eds.). Perspectives on the development of memory and cognition Hillsdale. NJ: Lawrence Erlbaum, 1977, pp. 3—33.
24. Hawker M.J., Dysleski L., Rickey D. Investigating general chemistry students' metacognitive monitoring of their exam performance by measuring postdiction accuracies over time. *Journal of Chemical Education*, 2016. Vol. 93(5), pp. 832—840. DOI:10.1021/acs.jchemed.5b00705
25. Korhonen J., Linnanmäki K., Aunio P. Language and mathematical performance: A comparison of lower secondary school students with different level of mathematical skills. *Scandinavian Journal of Educational Research*, 2012. Vol. 56, pp. 333—344. DOI:10.1080/00313831.2011.599423
26. Linacre J.M. A user's guide to WINSTEPS® MINISTEP: Rasch-model computer programs.(No Title). 2016.
27. Lindner M.A., Lüdtke O., Nagy G. The onset of rapid-guessing behavior over the course of testing time: a matter of motivation and cognitive resources. *Frontiers in Psychology*, 2019. Vol. 10. DOI:10.3389/fpsyg.2019.01533
28. Liu S., Liu M. The impact of learner metacognition and goal orientation on problem-solving in a serious game environment. *Computers in Human Behavior*, 2020. Vol. 102, pp. 151—165. DOI:10.1016/j.chb.2019.08.021
29. Masters G.N. Partial credit model. Chapman and Hall/CRC Handbook of item response theory, 2016, pp. 137—154.
30. Mokhtari K., Sheorey R. Measuring ESL students' awareness of reading strategies. *Journal of Developmental Education*, 2002. Vol. 25(3), pp. 2—11.
31. Murtagh F., Contreras P. Algorithms for hierarchical clustering: an overview. *WIREs Data Mining Knowl Discov*, 2012, pp. 86—97. DOI:10.1002/widm.53
32. Nelson T.O., Narens L. Metamemory: A theoretical framework and new findings. Academic Press G. Bower (Ed.). The psychology of learning and motivation: Advances in research and theory, 1990. Vol. 26, pp. 125—141.
33. Paris S.G., Winograd P. How metacognition can promote academic learning and instruction. Hillsdale, NJ: Erlbau B.F. Jones & L. Idol (Eds.). *Dimensions of thinking and cognitive instruction*, 1990, pp. 15—51.
34. Phakiti A. A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing*, 2003. Vol. 20(1), pp. 26—56.

35. Savolainen H., Ahonen T., Aro M., Tolvanen A., Holopainen L. Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction*, 2008. Vol. 18, pp. 201—210. DOI:10.1016/j.learninstruc.2007.09.0174
36. Smith Jr.E.V. Detecting and Evaluating the Impact of Multidimensionality using Item Fit Statistics and Principal Component Analysis of Residuals. *Journal of Applied Measurement*. 2002. Vol. 3, no. 2, pp. 205—231.
37. Tadayon F., Ravand H. Using grounded theory to validate Bachman and Palmer's (1996) strategic competence in EFL graph-writing. *Language Testing in Asia*, 2016. Vol. 6. DOI:10.1186/s40468-016-0031-y
38. Tibken C., Richter T., von der Linden N., Schmiedeler S., Schneider W. The role of metacognitive competences in the development of school achievement among gifted adolescents. *Child Development*, 2022. Vol. 93(1), pp. 117—133. DOI:10.1111/cdev.13640
39. Van Kraayenoord C.E. The role of metacognition in reading comprehension. *Focal points of the research and development of pedagogically-psychological perspectives*, 2010, pp. 277—302.
40. Vandergrift L., Goh C.C.M. Teaching and learning second language listening. Routledge. 2012. DOI:10.4324/9780203843376
41. Van der Stel M., Veenman M.V.J. Metacognitive skills and intellectual ability of young adolescents: A longitudinal study from a developmental perspective. *European Journal of Psychology of Education*, 2014. Vol. 29, pp. 117—137. DOI:10.1007/s10212-013-0190-5
42. Zargar E., Adams A.M., Connor C.M. The relations between children's comprehension monitoring and their reading comprehension and vocabulary knowledge: An eye-movement study. *Reading and Writing*, 2019. Vol. 33, pp. 511—545. DOI:10.1007/s11145-019-09966-3
43. Zhang L.J. A dynamic metacognitive systems account of Chinese university students' knowledge about EFL reading. *TESOL Quarterly*, 2010. Vol. 44, pp. 320—353.
44. Zhang L. Metacognitive and Cognitive Strategy Use in Reading Comprehension: A Structural Equation Modelling Approach. Springer Nature Singapore, 2018. DOI:10.1007/978-981-10-6325-1

Информация об авторах

Бакай Екатерина Анатольевна, стажер-исследователь Центра психометрики и измерений в образовании Института образования, ФГАОУ ВО «Научно-исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-7513-5554>, e-mail: ebakay@hse.ru

Юсупова Элен Магомедовна, младший научный сотрудник Центра психометрики и измерений в образовании Института образования, ФГАОУ ВО «Научно-исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-8332-681X>, e-mail: eabdurakhmanova@hse.ru

Антипкина Инна Вениаминовна, научный сотрудник Центра психометрики и измерений в образовании Института образования, ФГАОУ ВО «Научно-исследовательский университет «Высшая школа экономики» (ФГАОУ ВО «НИУ ВШЭ»), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0003-4865-3433>, e-mail: iantipkina@hse.ru

Information about the authors

Ekaterina A. Bakay, Researcher at the Center for Psychometrics and Measurements in Education, Institute of Education, National Research University Higher School of Economics, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-7513-5554>, e-mail: ebakay@hse.ru

Elen M. Yusupova, Junior Research Fellow, at the Center for Psychometrics and Measurements in Education, Institute of Education, National Research University Higher School of Economics, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-8332-681X>, email: eabdurakhmanova@hse.ru

Inna V. Antipkina, Research Fellow, at the Center for Psychometrics and Measurements in Education, Institute of Education, National Research University Higher School of Economics, Moscow, Russia, ORCID: <https://orcid.org/0000-0003-4865-3433>, e-mail: iantipkina@hse.ru

Получена 31.08.2023

Принята в печать 29.04.2024

Received 31.08.2023

Accepted 29.04.2024