

ПСИХОЛОГИЯ СПЕЦИАЛЬНОГО И ИНКЛЮЗИВНОГО ОБРАЗОВАНИЯ PSYCHOLOGY OF SPECIAL AND INCLUSIVE EDUCATION

Применение универсального дизайна оценивания для справедливого тестирования детей с нарушениями чтения

Брагинец Е.И.,

аспирант, стажер-исследователь Центра мониторинга качества образования Института образования, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия, ebraginet@hse.ru

Описаны принципы разработки психологических тестов и тестов достижений под названием «универсальный дизайн оценивания». Освещены и раскрыты основные элементы данного подхода. Введено понятие «аккомодации» как изменение процедуры тестирования без внесения изменений в измеряемый конструктор. В качестве иллюстрации описано создание теста в парадигме универсального дизайна оценивания через разработку аккомодаций процедуры тестирования для детей с нарушением чтения.

Ключевые слова: нарушения чтения, справедливое оценивание, универсальный дизайн обучения, аккомодации.

Для цитаты:

Брагинец Е.И. Применение универсального дизайна оценивания для справедливого тестирования детей с нарушениями чтения [Электронный ресурс] // Современная зарубежная психология. 2018. Том 7. № 3. С. 28—37. doi: 10.17759/jmfp.2018070303

For citation:

Braginets E.I. The application of universal design evaluation for fair testing children with reading disorders [Elektronnyi resurs]. *Journal of Modern Foreign Psychology*, 2018, vol. 7, no. 3, pp. 28—37. doi: 10.17759/jmfp.2018070303 (In Russ.; Abstr. in Engl.).

Проведение измерений в психологии, образовании и других социальных науках ставит своей целью объективно описать реальность. Однако существует ряд трудностей, снижающих точность измерения. В этой статье рассмотрена проблема оценивания детей и взрослых с нарушениями чтения, описаны способы повышения валидности измерений этой группы респондентов, а также предложены практические рекомендации для нивелирования нарушений чтения при прохождении тестов.

Нарушения чтения

Уметь хорошо читать значит точно и эффективно извлекать смысл из написанного текста.

Обучаясь чтению, дети развивают параллельно два набора навыков. Первый направлен на распознавание написанных букв и слов. Для этого развиваются:

- 1) Фонологическая осведомленность: произносимые слова состоят из более мелких элементов — звуков;
 - алфавитный принцип: звуки выражаются определенными символами — буквами;
 - декодирование: между буквами и звуками есть связь;
 - запоминание слов для более быстрого их распознавания.

Второй набор навыков ориентирован на понимание смысла прочитанного текста. Обучение здесь направлено на увеличение объема знаний ученика об окружа-

ющем мире, расширение его осведомленности о возможных контекстах, встречаемых в различных текстовых источниках, и включает в себя:

- общие знания о различных фактах и понятиях;
- расширение и углубление лексического аппарата;
- ознакомление с семантическими и синтаксическими конструкциями предложений, а также с конвенциональными нормами в грамматике [19].

Согласно одной из первых моделей нарушения чтения [29], такие нарушения можно разделить на четыре группы.

- **Дефект** — невозможность чтения без устранения какого-либо нарушения (например, плохое зрение) или предъявления текста в другой модальности.
- **Дефицит** — недостаточность навыков или возможностей для осуществления чтения.
- **Нарушение** — чтению мешают некоторые особенности психики или физиологии (например, тревожность или гиперактивность).
- **Новизна** — используемые в текстах паттерны речи неизвестны читающему.

В данной статье под нарушениями чтения будут пониматься первые три группы данной модели. Четвертая группа модели не будет рассматриваться нами как целевая, поскольку не является нарушением: ознакомление читающего с новыми паттернами речи решает проблему чтения [7].

Актуальность оценивания детей с нарушениями чтения

Для того чтобы обучаться в любой системе образования, необходимо уметь читать; это создает трудности для обучения и оценивания детей с нарушениями чтения.

Согласно отчету Американской национальной оценки прогресса в области образования (NAEP) 2017 г. [21], среди детей с нарушениями развития только 9% учащихся восьмого класса имеют хорошо развитые навыки чтения. В отчете американского Национального центра образовательной статистики 2001 г. [8] показано, что 37% детей 9—10 лет не обладают достаточно развитыми навыками чтения для использования обучающих материалов, рассчитанных на этот возраст.

Численность учащихся с нарушениями чтения в России точно неизвестна, поскольку массовая диагностика данного параметра в школе не проводится.

Согласно исследованию 1977 г., проведенному М.С. Певзнером, численность детей с когнитивными нарушениями, приводящими к затруднениям в чтении, составляет около 9% [2]. Этот показатель ниже, чем в странах Европы и Америки, однако по нему можно судить о том, что группа учащихся с нарушениями чтения в отечественной системе образования достаточно многочисленна. Важно отметить, что указанный выше источник является достаточно отдаленным от сегодняшнего дня по времени, однако более современных статистических данных о числе детей с нарушениями речи в России найдено не было, поэтому в статье используется наиболее цитируемая в отечественной научной литературе работа.

Эмпирически доказано, что нарушения в овладении навыками чтения оказывают долгосрочный негативный эффект на социальные и образовательные аспекты жизни человека [10; 22].

Для эффективного обучения детей с особыми образовательными потребностями, к которым относятся нарушения чтения, используют подход, который называется «универсальный дизайн обучения» [17; 20].

Универсальный дизайн обучения предполагает гибкость трех составляющих образовательного процесса: постановки целей обучения, поддержки ученика в достижении поставленных целей, обеспечения справедливого оценивания в процессе обучения [23].

В данной статье рассмотрен один из аспектов универсального дизайна обучения — универсальный дизайн оценивания.

Этот аспект крайне важен, так как качественные инструменты оценивания и адекватное их применение позволяют создать единую легитимную доказательную базу образовательных практик, на основе которой могут реализоваться действительно эффективные научные коммуникативные связи специалистов.

Универсальный дизайн оценивания

Свое начало универсальный дизайн оценивания берет из архитектуры. Термин «универсальный дизайн» впервые ввел Рон Мэйс, архитектор на инвалидной коляске, который стремился к тому, чтобы здания и улицы городов становились все более доступными для людей с особыми потребностями — физическими, сенсорными и психическими [26].

Рон Мэйс и его коллеги сформулировали основные принципы универсального дизайна: равное использование, гибкость в использовании, простой и интуитивно понятный дизайн, легко воспринимаемая информация, толерантность к ошибкам, минимальные физические усилия, размер и пространство для доступа и использования.

Универсальный дизайн оценивания базируется на принципах универсального дизайна в целом, но видоизменяет их под свои цели:

- справедливая оценка групп респондентов с особыми потребностями;
- ясное обозначение конструкта измерения;
- одинаковое функционирование заданий для всех групп респондентов;
- процедура тестирования, допускающая аккомодации;
- простые, ясные и интуитивно-понятные инструкции и процедура теста;
- максимальное упрощение восприятия текста и его смысла;
- максимальная смысловая и визуальная ясность.

Далее перечисленные выше принципы будут подробно рассмотрены.

Справедливая оценка групп респондентов с особыми потребностями

При разработке теста один из первых шагов — это установление потенциальных групп респондентов теста [24]. В случае создания теста, направленного на отбор и обеспечение конкурса, допустимо ограничение целевой аудитории теста по некоторым характеристикам (например, соискатель с высшим инженерным образованием, опыт работы не менее 3 лет).

Однако ограничение целевой аудитории невозможно при создании теста для использования в массовом образовании: проведение политики инклюзивного образования [1] не позволяет предугадать, дети с какими особыми образовательными потребностями окажутся в классе общеобразовательной школы.

В соответствии с Приказом Минобрнауки 2016 г., проводимые в общеобразовательных школах измерения должны отвечать потребностям детей с особыми образовательными потребностями, а также предоставлять сравнимые результаты по итогам измерения. Тестирование должно одинаково справедливо измерять знания и характеристики детей с различными возможностями здоровья, знанием языка, культурным опытом.

Такие требования к тестированию могут приводить к изменению конструкта измерения или упрощению заданий теста для респондентов с особыми образовательными потребностями, что может ухудшить сопоставимость результатов тестирования между детьми из разных групп.

Однако разработка инструментов измерения в парадигме универсального дизайна оценивания позволяет избежать таких искажений и обеспечить максимально справедливое оценивание респондентов с различными образовательными потребностями.

Основным способом осуществления справедливой оценки групп респондентов с особыми потребностями является использование подгрупп респондентов с различными ограничениями возможностей здоровья при проведении пилотных исследований.

Таким образом, у разработчиков заданий появляется возможность узнать о функционировании заданий до выхода в массовое тестирование.

Ясное обозначение конструкта измерения

Универсальный дизайн среды создавался для нивелирования физических барьеров при взаимодействии со средой; целью универсального дизайна оценивания считают устранение нерелевантных конструкту измерения барьеров: в процедуре тестирования, формате заданий, форме подачи материала.

Для понимания того, какие элементы процедуры тестирования являются нерелевантными, важна формулировка определения измеряемого конструкта.

Например, такой конструкт, как «понимание текста» может быть определен по-разному. Разработчики могут рассматривать его как понимание прочитанного текста и тогда изменения процедуры тестирования, касающиеся формы предъявления стимульного материала, будут снижать валидность измерений. Определение же конструкта «понимание текста» безотносительно формата его предъявления позволяет в дальнейшем применять к тесту различные голосовые модификации, которые повысят доступность теста для людей с нарушениями чтения.

Описываемый принцип демонстрирует, что для создания теста, справедливо оценивающего различные группы респондентов, необходимо учитывать разнообразие целевой аудитории еще до начала разработки заданий и процедуры тестирования.

Более того, определение измеряемого конструкта должно быть ясно определено в документах, сопровождающих инструмент, для того чтобы пользователи могли понять, насколько инструмент подходит под их выборку и цели тестирования.

Одинаковое функционирование заданий для всех групп респондентов

Задания могут по-разному оценивать респондентов из групп с разным культурным капиталом [16], жизненным опытом [28], нарушениями развития [4].

Примером дискриминации одной из групп респондентов может быть использование специфического

контекста задания: мальчики, решающие математические задачи, в которых речь идет о работе парикмахера, с большей вероятностью получают более низкие баллы за решение задачи, чем девочки. Проблема такого задания будет в новизне контекста для решающих задачу мальчиков, поскольку им понадобится дополнительное время на то, чтобы разобраться в предложенной ситуации.

Причиной различающегося функционирования заданий теста могут выступать особенности их элементов: инструкций к тесту, процедуры проведения, текстов заданий. Использование слов в инструкции и заданиях, которые могут быть неизвестны группе детей с низким социокультурным статусом или новый для одной группы детей формат заданий будет приводить к более низким результатам теста при равных способностях.

Основным способом проверки справедливого функционирования заданий теста является DIF-анализ [30]. Дифференцированное функционирование заданий (DIF) возникает, когда респонденты с равным уровнем способностей имеют разную вероятность верно ответить на задание теста [14]. Если у заданий обнаруживается DIF, то эти задания измеряют не только заявленный конструкт, а значит, их результаты не могут сравниваться между данными группами респондентов [28].

Процедура тестирования допускает аккомодации

Аккомодации процедуры тестирования и формата заданий являются одним из наиболее эффективных способов повысить доступность теста для различных групп респондентов без изменения измеряемого конструкта.

Важным элементом при разработке теста в парадигме универсального дизайна оценивания является необходимость принимать во внимание возможность внесения аккомодаций в процедуру тестирования при оценке особых групп респондентов.

Важно обозначить различия между понятиями аккомодации и модификации. В работе 1998 г. Холленбек, Тиндал и Алмонд формулируют эти понятия следующим образом.

Аккомодация — изменение процедуры тестирования без внесения изменений в измеряемый конструкт. Целью аккомодаций является создание возможности сопоставления результатов между респондентами с различными возможностями здоровья и образовательными потребностями. Аккомодации могут затрагивать формат заданий, форму подачи материала, администрирование или формат ответа и предоставляются специальным группам респондентов. Модификации — это аккомодации процедуры тестирования, происходящие для всех групп респондентов [3].

Аккомодации могут касаться различных аспектов процедуры тестирования [20]:

- формат представления материала (шрифт Брайля или буклеты с крупными буквами, использование языка жестов или текстовое написание инструкций);

- формат ответа (записывание ответа, подчеркивание участка текста, компьютерная форма, голосовой ответ);
- продолжительность тестирования и перерывы (увеличенное время выполнения теста, дополнительные перерывы);
- условия тестирования (индивидуальное администрирование теста, условия освещения или другие физические условия прохождения тестирования).

Более подробно аккомодации рассмотрены в следующем разделе данной статьи.

Простые, ясные и интуитивно-понятные инструкции и процедура теста

С целью обеспечения максимальной доступности теста для различных групп респондентов отдельное внимание должно быть уделено инструкции теста. Стандарты тестирования обозначают необходимость формулировки инструкции таким образом, чтобы респондент после ознакомления с ней вел себя именно тем способом, который предполагает разработчик теста.

Для проверки функционирования инструкции возможно проведение когнитивных лабораторий, направленных на понимание потенциальными респондентами инструкции. Включение в когнитивные лаборатории респондентов с особыми образовательными потребностями позволит выделить потенциально сложные или непонятно сформулированные элементы инструкции.

Также для облегчения понимания инструкции и процедуры тестирования рекомендуется использовать примеры задач, таким образом, респондент может потренироваться в выполнении задания или работе с интерфейсом до перехода к основному тестированию. Предварительное размещение критериев оценки также может повысить доступность теста для специальных групп респондентов.

Максимальное упрощение восприятия текста и его смысла

Обеспечение максимально понятного текста инструкций и задания важно для валидного измерения заявленного конструкта. Разработчики тестов в парадигме универсального дизайна оценивания советуют использовать руководство, предложенное Гастером и Кларком в 1995 г. [11]:

- при написании текстов используйте простые, понятные, часто используемые слова, избегайте использования избыточных слов;
- при использовании терминов необходимо включить в текст инструкции или задания по их определению;
- составные сложные предложения рекомендуется разделить на несколько коротких предложений, первым разместить то, которое содержит основную мысль;
- представляйте одну идею или факт за раз, далее логично развивайте и объясняйте их;

- соотнесение между существительным и относящимся к нему местоимением должно быть очевидно;
- если обстоятельства времени и места важны для понимания предложения, размещайте их в начале предложения;
- при написании инструкции используйте описание действий респондента в той последовательности, которая необходима для выполнения теста;
- если используется описание процессов, последние должны быть просто проиллюстрированы (примерами или изображениями), подписаны и расположены рядом с текстом, который они иллюстрируют.

Максимальная смысловая и визуальная ясность

При использовании текстовых материалов в тесте особое внимание должно быть уделено смысловой и визуальной ясности текста. Визуальная четкость достигается с помощью максимально удобных для чтения шрифтов, соотношением цвета и контрастности текста и фона, межстрочных интервалов. Сложный для визуального восприятия текст задания или инструкции вносит в полученные результаты искажения, поскольку не относится к измеряемому конструкту.

Смысловая четкость предполагает использование максимально простых символов для обозначения элементов на графике, подписи и легенды должны располагаться максимально близко к обозначаемому элементу. Разработчики теста должны стремиться к снижению нагрузки на краткосрочную память респондента, если она не релевантна измеряемому конструкту.

Немаловажно учесть при создании теста в рамках универсального дизайна оценивания и формат ответа респондента. Создание альтернативных форм ответа повышает доступность инструмента: например, правильным ответом может засчитываться подчеркивание, выделение и выписывание нужного слова в тексте.

Таким образом, снижается вред от ошибки респондента при неверно понятой или забытой инструкции. По этой же причине исследователи отмечают, что наиболее эффективной является форма ответа на листе с заданием, а не вынесение полей ответа на отдельный бланк [15].

Аккомодации процедуры тестирования

Первые стандарты для создания специальных условий для прохождения тестов были созданы Американской ассоциацией образовательных исследований, Американской психологической ассоциацией и Национальным советом по измерениям в образовании.

В 1985 г. в сотрудничестве ими были выпущены «Стандарты для образовательного и психологического тестирования» [25], одна глава которых была посвящена тестированию людей, испытывающих трудности при выполнении тестов в обычных условиях.

В этой главе были перечислены восемь стандартов, касающихся создания и импликации аккомодаций к тестовым процедурам.

В дальнейшем стандарты тестирования развивались в сторону вопроса валидности измененной процедуры стандартизированного тестирования — можно ли интерпретировать результаты тестирования с использованием аккомодаций так же, как и без них.

В стандартах выделяются следующие угрозы валидности:

- информация об эффекте ограниченных возможностей респондента на результаты тестирования может быть недостаточна;
- выборки с каждым отдельным нарушением для пилотных исследований могут быть труднодоступны;
- эмпирические данные об эффекте произведенных аккомодаций на валидность измерения могут быть недостаточны.

Несмотря на это, аккомодации являются очень важным элементом для обеспечения справедливого тестирования, и при аккуратном и ответственном обращении они могут выступать эффективным инструментом повышения доступности инструмента оценивания специальных групп респондентов.

Встает вопрос о том, как должна быть выстроена процедура проведения аккомодаций при тестировании групп респондентов с особыми потребностями.

В случае со взрослыми респондентами обычно используется предварительная форма, которую участник заполняет при регистрации на экзамен, где указывает, какие нарушения у него есть и какие аккомодации ему нужно получить во время экзамена (например, GMAT), также респондент должен приложить заключение специалиста о наличии у него указанных нарушений.

Но существует трудность в обеспечении доступа к аккомодациям респондентов младшего возраста: как определить, что ребенок должен получить особые условия прохождения того или иного тестирования?

Возможно использовать подход, описанный выше, однако не все ограниченные возможности здоровья диагностируются в детстве.

В некоторых школах США используется система, в которой учитель указывает особые образовательные потребности ребенка, и в соответствии с этим тест, который предлагается ребенку, уже включает в себя все необходимые аккомодации.

Такой подход имеет свои ограничения, поскольку оценка особых потребностей ребенка учителем может быть субъективна.

Помимо этого, существует ряд сложностей в разработке и психометрическом анализе инструментов с использованием аккомодаций.

Во-первых, это малые размеры выборок. Труднодоступность сбора выборки усложняет проведение пилотных исследований и сбор данных о валидности инструмента после изменения процедуры.

Также трудность для разработчиков теста представляет разнородность нарушений развития у респондентов. Стандарты тестирования требуют, чтобы адапти-

ровать процедуры тестирования к каждой специфической группе респондентов. Однако разнообразие характеристик респондентов так велико, что классификация респондентов на отдельные группы зачастую затруднительна.

Разработчики тестов сталкиваются с проблемой разнообразия доступных аккомодаций: подачи материала, формата заданий, формы ответа, средовые аккомодации. Поскольку для различных нарушений развития существуют отличающиеся наборы аккомодаций, большую трудность представляет проведение исследований с их применением.

Эти ограничения в применении аккомодаций поднимают ряд дискуссионных вопросов, ответа на которые нет в научном сообществе:

- Справедливо ли изменять процедуру тестирования только для специальных групп респондентов?
- Как должен быть организован доступ респондента к получению аккомодаций?
- Как валидно интерпретировать полученные результаты и принимать решения на их основе?

Поднятые вопросы не имеют однозначного ответа, поэтому при анализе данных мониторинговых образовательных исследований и написании отчетов данные учеников, получавших аккомодации во время тестирования, иногда исключаются [27].

Как было сказано выше, впервые о возможности ослабления стандартизации процедуры теста было написано в стандартах тестирования 1985 г.

Особое внимание оценке респондентов с особыми потребностями было уделено в обновленной версии 1999 г., поскольку число таких респондентов увеличилось.

Согласно стандарту, в процедуру тестирования могут вноситься изменения, если внесенные изменения не релевантны измеряемому конструкту. Определение релевантности произведенных аккомодаций измеряемому конструкту должно происходить на этапе пилотных исследований качества теста.

В 2000 г. ряд авторов [2] предложили модель, в которой аккомодации теста могут быть отнесены к одной из трех групп.

Группа 1: аккомодации, не вносящие изменений в интерпретации тестового балла. Например, участник выполняет тест, находясь в комнате один, а не вместе с группой.

Группа 2: аккомодации, способные внести изменения в интерпретацию тестового балла. Например, добавление дополнительного времени к тестам, предполагающим выполнение задачи в строго определенное время.

Группа 3: аккомодации, вносящие изменения в интерпретацию тестовых баллов. В таком случае балл должен быть интерпретирован с осторожностью. Например, использование калькулятора на тесте по математике может внести искажения в измеряемый конструкт.

Реализация принципов УД для детей с нарушениями чтения

Универсальный дизайн оценивания может быть применен и при разработке тестов, в целевую группу которых входят дети с нарушениями чтения. Одна из задач разработчиков — учесть возможность аккомодаций процедуры тестирования, которые используются для данной группы респондентов.

Поскольку нарушения чтения могут проявляться среди детей и взрослых очень разнообразно, стандартизация аккомодаций для этого вида нарушений не может быть произведена [9].

В дополнение к проявлениям нарушения, компенсаторные техники респондентов могут также серьезно отличаться.

Таким образом, введение единого списка обязательных аккомодаций для респондентов с нарушением чтения может оказаться избыточным или недостаточным для некоторых респондентов.

Роль учителя или сопровождающего ребенка с нарушениями чтения становится критически важной для преодоления этих ограничений и определения аккомодаций, необходимых респонденту.

Распространены несколько видов аккомодаций для респондентов с нарушениями чтения: автоматизированное чтение вслух [12], аккомодации написания, выделение строки или слова для облегчения чтения, изменение характеристик текста (фон, шрифт, цвет).

Поскольку для валидности измерения важно, чтобы введенные аккомодации не были релевантны конструкту измерения, проводятся специальные исследования.

Проверка валидности изменения процедуры тестирования должна подбираться и проводиться для каждого отдельного теста.

Например, в одном исследовании [6] рассматривалось функционирование аккомодации «полное чтение вслух» для детей с нарушениями чтения. Авторы показали, что независимо от того, есть ли у ребенка-респондента нарушения чтения или нет, баллы детей повышаются. В описываемой работе такой результат связывается с тем, что аккомодация теста привела к изменению измеряемого конструкта — вместо смыслового чтения, респонденты занимались смысловым слушанием.

По результатам данного исследования можно сделать вывод, что аккомодация «полного чтения вслух» не подходит для используемого в исследовании теста. Возможно, она может быть заменена на аккомодации с частичным озвучиванием инструкций к заданиям, а также потенциально сложных для чтения слов.

В то же время ряд исследований показывает, что применение такой же аккомодации, напротив, повышает валидность теста, поскольку делает процедуру теста более доступной для учащихся с нарушениями чтения [5; 13]. Эти исследования показали, что «полное чтение вслух» значительно повышает баллы детей в тесте по математике. Такая аккомодация не создает угрозы измерению заявленного конструкта, поскольку полностью не релевантна ему.

Также при разработке теста, который будет использован респондентами с нарушениями чтения, очень важно уделить особое внимание текстам заданий и инструкций. Поскольку респондент испытывает затруднения при чтении, используемые в тесте тексты должны быть максимально простыми и иметь логичную структуру.

Заключение

Валидные измерения детей и взрослых с нарушениями чтения невозможны без учета их особых потребностей.

Использование универсального дизайна оценивания при разработке тестов позволяет повысить объективность оценки способностей и характеристик респондентов с такими нарушениями.

Изменение процедуры тестирования с помощью проверенных в исследованиях аккомодаций способствует повышению валидности измерений, а также включению результатов особых групп респондентов в анализ наравне с остальными участниками тестирования.

Следование принципам универсального дизайна оценивания позволяет валидно оценивать психологические характеристики и знания групп респондентов с ограниченными возможностями здоровья с помощью опросников, анкет и тестов.

Несмотря на то, что в статье были рассмотрены аспекты применения этого подхода только для респондентов с нарушениями чтения, следует сказать, что универсальный дизайн оценивания применим для широкого спектра особых потребностей: нарушения слуха и зрения, поведенческие расстройства, физические ограничения.

В России описанный подход к разработке измерительных инструментов только начинает свое развитие, однако в условиях становления инклюзивного образования он является остро актуальным и востребованным.

Универсальный дизайн оценивания может получить широкое распространение в психологии, образовании, социологии и тех сферах, где необходимо использование стандартизированных тестов и опросников.

ЛИТЕРАТУРА

1. Об утверждении состава Совета Министерства образования и науки Российской Федерации по вопросам образования лиц с ограниченными возможностями здоровья и инвалидов: Приказ Минобрнауки России от 16 ноября 2016 г. № 1430 [Электронный ресурс]// Министерство образования и науки Российской Федерации. URL: <https://минобрнауки.рф/документы/9134> (дата обращения: 13.07.2018).

2. Певзнер М.С., Явкин В.М. Результаты обследований неуспевающих учащихся начальных классов общеобразовательных школ // Дефектология. 1977. № 2. С. 19—23.
3. Accommodating students with disabilities on large-scale tests: An empirical study / G. Tindal [et al.] // *Exceptional Children*. 1998. Vol. 64. № 4. P. 439—450. doi:10.1177/001440299806400401
4. Bolt S. E., Ysseldyke J. Accommodating students with disabilities in large-scale testing: A comparison of differential item functioning (DIF) identified across disability types // *Journal of Psychoeducational Assessment*. 2008. Vol. 26. № 2. P. 121—138. doi:10.1177/0734282907307703
5. Calhoun M.B., Fuchs L., Hamlett C. Effects of computer-based test accommodations on mathematics performance assessments for secondary students with learning disabilities // *Learning Disability Quarterly*. 2000. Vol. 23. № 4. P. 271—282. doi:10.2307/1511349
6. Crawford L., Tindal G. Effects of a read-aloud modification on a standardized reading test // *Exceptionality*. 2004. Vol. 12. № 2. P. 89—106. doi:10.1207/s15327035ex1202_3
7. Cromer W. The difference model: A new explanation for some reading difficulties // *Journal of Educational Psychology*. 1970. Vol. 61. № 6. P. 471—483. doi:10.1037/h0030288
8. Digest of Education Statistics 2001 [Электронный ресурс] / National Center for Education Statistics. Washington, DC: U.S. Department of Education. Office of Educational Research and Improvement, 2002. 597 p. URL: <https://nces.ed.gov/pubspubs2002/2002130.pdf> (дата обращения: 13.07.2018).
9. Enhancing Accessibility for Students With Decoding Difficulties on Large-Scale Reading Assessments / A.G. Gandhi [et al.] // *Journal of learning disabilities*. 2018. Vol. 51. № 6. P. 540—551. doi:10.1177/0022219417714774
10. Evidence-based reading and writing assessment for dyslexia in adolescents and young adults / K. Nielsen [et al.] // *Learning Disabilities: A Multidisciplinary Journal*. 2016. Vol. 21. № 1. P. 38—56. doi:10.18666/LDMJ-2016-V21-11-6971
11. Gaster L., Clark C. A guide to providing alternate formats [Электронный ресурс]. Washington, DC: Center for Rehabilitation Technology Services, 1995. 24 p. URL: <https://eric.ed.gov/?id=ED405689> (дата обращения: 13.07.2018).
12. Gross J. The long term costs of literacy difficulties [Электронный ресурс]. 2d ed. London: Every Child a Chance Trust. 2009. 52 p. URL: http://readingrecovery.org/images/pdfs/Reading_Recovery/Research_and_Evaluation/long_term_costs_of_literacy_difficulties_2nd_edition_2009.pdf (дата обращения: 13.07.2018).
13. Harker J.K., Feldt L.S. A comparison of achievement test performance of nondisabled students under silent reading plus listening modes of administration // *Applied Measurement*. 1993. Vol. 6. № 4. P. 307—320. doi:10.1207/s15324818ame0604_4
14. Holland P. W., Wainer H. Differential item functioning. New York; London: Routledge, 2012. 437 p.
15. Hollenbeck K., Tindal G., Almond P. Teachers' knowledge of accommodations as a validity issue in high-stakes testing // *Journal of Special Education*. 1998. Vol. 32. № 3. P. 175—183. doi:10.1177/002246699803200304
16. Jones R.N., Gallo J.J. Education and sex differences in the mini-mental state examination: effects of differential item functioning // *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2002. Vol. 57. № 6. P. P548—P558. doi:10.1093/geronb/57.6.P548
17. Ketterlin-Geller L.R. Knowing what all students know: Procedures for developing universal design for assessment [Электронный ресурс] // *The Journal of Technology, Learning and Assessment*. 2005. Vol. 4. № 2. P. 4—22. URL: <https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1649/1491> (дата обращения: 13.07.2018).
18. Ketterlin-Geller L.R., Johnstone C. Accommodations and Universal Design: Supporting Access to Assessments in Higher Education // *Journal of Postsecondary Education and Disability*. 2006. Vol. 19. № 2. P. 163—172. URL: <https://files.eric.ed.gov/fulltext/EJ844632.pdf> (дата обращения: 13.07.2018).
19. McCardle P., Scarborough H.S., Catts H.W. Predicting, explaining, and preventing children's reading difficulties // *Learning Disabilities Research & Practice*. 2001. Vol. 16. № 4. P. 230—239. doi:10.1111/0938-8982.00023
20. Pitoniak M.J., Royer J.M. A review of psychometric, legal, and social policy issues // *Special Educational Needs and Inclusive Education: Assessment and teaching strategies*. 2004. Vol. 3. № 1. P. 53—104.
21. Reading Results [Электронный ресурс]: The National Assessment of Educational Progress (NAEP) measures the progress of our nation's fourth- and eighth-graders in reading, mathematics, and various other subjects. 2017 // *The Nation's Report Card*. URL: https://www.nationsreportcard.gov/reading_math_2017_highlights/files/infographic_2018_reading.pdf (дата обращения: 13.07.2018).
22. Reynolds M., Wheldall K., Madeline A. Components of effective early reading interventions for young struggling readers // *Australian Journal of Learning Difficulties*. 2010. Vol. 15. № 2. P. 171—192. doi:10.1080/19404150903579055
23. Rose D.H., Meyer A. Teaching every student in the digital age [Электронный ресурс]: Universal design for learning. Alexandria, VA: ASCD, 2002. 11 p. URL: <https://eric.ed.gov/?id=ED466086> (дата обращения: 13.07.2018).
24. Standards for educational and psychological testing / American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Washington, DC: American Psychological Association, 1985. 100 p.
25. Standards for educational and psychological testing / American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Washington, DC: American Psychological Association, 1999. 194 p.

26. *Story M.F., Mueller J.L., Mace R.L.* The universal design file: Designing for people of all ages and abilities [Электронный ресурс]. Raleigh, NC: NC State University, The Center for Universal Design, 1998. 172 p. URL: <https://eric.ed.gov/?id=ED460554> (дата обращения: 13.07.2018).
27. *Thompson S.J., Johnstone C.J., Thurlow M.L.* Universal Design Applied to Large Scale Assessments. Synthesis Report [Электронный ресурс]. Minneapolis: National Center on Educational Outcomes, 2002. 38 p. URL: <https://files.eric.ed.gov/fulltext/ED467721.pdf> (дата обращения: 13.07.2018).
28. *Van de Vijver F., Tanzer N.K.* Bias and equivalence in cross-cultural assessment: An overview // *European Review of Applied Psychology*. 2004. Vol. 54. № 2. P. 119—135. doi:10.1016/j.erap.2003.12.004
29. *Wiener M., Cromer W.* Reading and reading difficulty: A conceptual analysis / *Harvard Educational Review*, 1967. Vol. 37. № 4. P. 620—643. doi:10.17763/haer.37.4.7701k0020072hh81
30. *Zumbo B.D.* A handbook on the theory and methods of differential item functioning (DIF) [Электронный ресурс]. Ottawa: National Defense Headquarters, 1999. 57 p. URL: <http://faculty.educ.ubc.ca/zumbo/DIF/handbook.pdf> (дата обращения: 13.07.2018).

The application of universal design evaluation for fair testing children with reading disorders

Braginets E.I.,

post-graduate student, Research Intern at the Center for Education quality monitoring of National Research University
«Higher School of Economics», Moscow, Russia,
ebraginets@hse.ru

The article describes the principles of development of psychological tests and achievement tests that are referred to as «universal design evaluation». It reveals the main elements of this approach. The authors introduce the concept of accommodation as a shift in testing procedure without changing the measurable construct. They illustrate how to create a test in the paradigm of the universal evaluation design through the development of accommodation of testing procedures for children with reading difficulties.

Keywords: reading, fair assessment, universal design training, accommodation.

REFERENCES

1. Ob utverzhdenii sostava Soveta Ministerstva obrazovaniya i nauki Rossiiskoi Federatsii po voprosam obrazovaniya lits s ogranichennymi vozmozhnostyami zdorov'ya i invalidov [Elektronnyi resurs] [On the approval of the Council of the Ministry of Education and Science of the Russian Federation on the education of persons with disabilities and disabled people]: Prikaz Minobrnauki Rossii ot 16 noyabrya 2016 g. № 1430. In *The Ministry of education and science of the Russian Federation*. Available at: <https://минобрнауки.рф/документы/9134> (Accessed 13.07.2018). (In Russ.).
2. Pevzner M.S., Yavkin V.M. Rezul'taty obsledovaniya neuspevayushchikh uchashchikhsya nachal'nykh klassov obshcheobrazovatel'nykh shkol [The results of inspections of underachieving primary school pupils in general education schools]. *Defektologiya [Defectology]*, 1977, no. 2, pp. 19—23. (In Russ.).
3. Tindal G. et al. Accommodating students with disabilities on large-scale tests: An empirical study. *Exceptional Children*, 1998, vol. 64, no. 4, pp. 439—450. doi:10.1177/001440299806400401
4. Bolt S. E., Ysseldyke J. Accommodating students with disabilities in large-scale testing: A comparison of differential item functioning (DIF) identified across disability types. *Journal of Psychoeducational Assessment*, 2008, vol. 26, no. 2, pp. 121—138. doi:10.1177/0734282907307703
5. Calhoun M.B., Fuchs L., Hamlett C. Effects of computer-based test accommodations on mathematics performance assessments for secondary students with learning disabilities. *Learning Disability Quarterly*, 2000, vol. 23, no. 4, pp. 271—282. Available at: doi:10.2307/1511349
6. Crawford L., Tindal G. Effects of a read-aloud modification on a standardized reading test. *Exceptionality*, 2004, vol. 12, no. 2, pp. 89—106. doi:10.1207/s15327035ex1202_3
7. Cromer W. The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, 1970, vol. 61, no. 6, pp. 471—483. doi:10.1037/h0030288
8. National Center for Education Statistics. Digest of Education Statistics 2001 [Elektronnyi resurs]. Washington, DC: U.S. Department of Education. Office of Educational Research and Improvement, 2002. 597 p. Available at: <https://nces.ed.gov/pubs2002/2002130.pdf> (Accessed 13.07.2018).
9. Gandhi A.G. et al. Enhancing Accessibility for Students With Decoding Difficulties on Large-Scale Reading Assessments. *Journal of learning disabilities*, 2018, vol. 51, no. 6, pp. 540—551. doi:10.1177/0022219417714774
10. Nielsen K. et al. Evidence-based reading and writing assessment for dyslexia in adolescents and young adults. *Learning Disabilities: A Multidisciplinary Journal*, 2016, vol. 21, no. 1, pp. 38—56. doi:10.18666/LDMJ-2016-V21-I1-6971
11. Gaster L., Clark C. A guide to providing alternate formats [Elektronnyi resurs]. Washington, DC: Center for Rehabilitation Technology Services, 1995. 24 p Available at: <https://eric.ed.gov/?id=ED405689> (Accessed 13.07.2018).
12. Gross J. The long term costs of literacy difficulties [Elektronnyi resurs]. 2d ed. London: Every Child a Chance Trust. 2009. 52 p. Available at: http://readingrecovery.org/images/pdfs/Reading_Recovery/Research_and_Evaluation/long_term_costs_of_literacy_difficulties_2nd_edition_2009.pdf (Accessed 13.07.2018).
13. Harker J.K., Feldt L.S. A comparison of achievement test performance of nondisabled students under silent reading plus listening modes of administration. *Applied Measurement*, 1993, vol. 6 no. 4, pp. 307—320. doi:10.1207/s15324818ame0604_4
14. Holland P. W., Wainer H. Differential item functioning. New York; London: Routledge, 2012. 437 p.
15. Hollenbeck K., Tindal G., Almond P. Teachers' knowledge of accommodations as a validity issue in high-stakes testing. *Journal of Special Education*, 1998, vol. 32, no. 3, pp. 175—183. doi:10.1177/002246699803200304
16. Jones R.N., Gallo J.J. Education and sex differences in the mini-mental state examination: effects of differential item functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 2002, vol. 57, no. 6, pp. P548—P558. doi:10.1093/geronb/57.6.P548

17. Ketterlin-Geller L.R. Knowing what all students know: Procedures for developing universal design for assessment [Elektronnyi resurs]. *The Journal of Technology, Learning and Assessment*, 2005, vol. 4, no. 2, pp. 4—22. Available at: <https://files.eric.ed.gov/fulltext/EJ844632.pdf> (Accessed 13.07.2018).
18. Ketterlin-Geller L.R., Johnstone C. Accommodations and Universal Design: Supporting Access to Assessments in Higher Education [Elektronnyi resurs]. *Journal of Postsecondary Education and Disability*, 2006, vol. 19, no. 2, pp. 163—172. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.835.8321&rep=rep1&type=pdf> (Accessed 13.07.2018).
19. McCardle P., Scarborough H.S., Catts H.W. Predicting, explaining, and preventing children's reading difficulties. *Learning Disabilities Research & Practice*, 2001, vol. 16, no. 4, pp. 230—239. doi:10.1111/0938-8982.00023
20. Pitoniak M.J., Royer J.M. A review of psychometric, legal, and social policy issues. *Special Educational Needs and Inclusive Education: Assessment and teaching strategies*, 2004, vol. 3, no. 1, pp. 53—104.
21. Reading Results [Elektronnyi resurs]: The National Assessment of Educational Progress (NAEP) measures the progress of our nation's fourth- and eighth-graders in reading, mathematics, and various other subjects. 2017. In *The Nation's Report Card*. Available at: https://www.nationsreportcard.gov/reading_math_2017_highlights/files/infographic_2018_reading.pdf (Accessed 13.07.2018).
22. Reynolds M., Wheldall K., Madelaine A. Components of effective early reading interventions for young struggling readers. *Australian Journal of Learning Difficulties*, 2010, vol. 15, no. 2, pp. 171—192. doi:10.1080/19404150903579055
23. Rose D.H., Meyer A. Teaching every student in the digital age [Elektronnyi resurs]: Universal design for learning. ASCD, 2002. 11 p. Available at: <https://eric.ed.gov/?id=ED466086> (Accessed 13.07.2018).
24. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Psychological Association, 1985. 100 p.
25. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Psychological Association, 1999. 194 p.
26. Story M.F., Mueller J.L., Mace R.L. The universal design file: Designing for people of all ages and abilities [Elektronnyi resurs]. Raleigh, NC: NC State University, The Center for Universal Design, 1998. 172 p. Available at: <https://eric.ed.gov/?id=ED460554> (Accessed 13.07.2018).
27. Thompson S.J., Johnstone C.J., Thurlow M.L. Universal Design Applied to Large Scale Assessments. Synthesis Report [Elektronnyi resurs]. Minneapolis: National Center on Educational Outcomes, 2002. 38 p. Available at: <https://files.eric.ed.gov/fulltext/ED467721.pdf> (Accessed 13.07.2018).
28. Van de Vijver F., Tanzer N.K. Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 2004, vol. 54, no. 2, pp. 119—135. doi:10.1016/j.erap.2003.12.004
29. Wiener M., Cromer W. Reading and reading difficulty: A conceptual analysis. *Harvard Educational Review*, 1967, vol. 37, no. 4, pp. 620—643. doi:10.17763/haer.37.4.7701k0020072hh81
30. Zumbo B.D. A handbook on the theory and methods of differential item functioning (DIF) [Elektronnyi resurs]. Ottawa: National Defense Headquarters, 1999. 57 p. Available at: <http://faculty.educ.ubc.ca/zumbo/DIF/handbook.pdf> (Accessed 13.07.2018).