

УДК 372.881.111.1

КОМПЬЮТЕРНЫЕ АДАПТИВНЫЕ ТЕСТЫ И ДОСТОВЕРНОСТЬ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ ПО АНГЛИЙСКОМУ ЯЗЫКУ

© 2014 г. *М.В. Золотова, Л.К. Делягина, Н.В. Каминская, Т.В. Мартынова*

Нижегородский госуниверситет им. Н.И. Лобачевского

mviazolotova@gmail.com

Поступила в редакцию 12.11.2014

Проанализированы использование компьютерных адаптивных тестов при оценке знания английского языка и применение теории вероятности для подтверждения достоверности результатов такого тестирования.

Ключевые слова: английский язык, компьютерный адаптивный тест, достоверность результатов тестирования.

Иновационные процессы в российском образовании предъявляют новые требования к изучению и преподаванию иностранного языка, поэтому в настоящее время способы и методики адекватной достоверной оценки знаний и компетенций обучающихся особенно актуальны.

Для входного диагностического, промежуточных и итогового тестирования все более широко начинают применяться компьютерные адаптивные тесты [1], которые представляют собой быстрый и достаточно эффективный способ измерения уровня обученности.

В чем коренное отличие адаптивных тестов от традиционных? В адаптивных тестах каждое последующее задание зависит от результатов ответа на предыдущее задание. Поступающие данные обрабатываются компьютером в реальном времени. Это одно из больших достоинств адаптивных компьютерных тестов.

В адаптивном тесте при правильном выполнении задания простого уровня (например, Elementary) автоматически переходят к выполнению заданий более высокого уровня (например, Pre-Intermediate). Тестирование завершается, когда обучаемый выходит на некоторый постоянный уровень сложности (то есть выполняет подряд определенное количество заданий одного уровня сложности).

Еще одним достоинством адаптивных тестов является снижение времени тестирования, так как студенту может быть предъявлено гораздо меньше заданий. При этом диагностические возможности не сокращаются. Достигается также большая точность измерений. Надежность результатов тестирования в данном случае самая высокая [2; 3].

Современное адаптивное тестирование, основанное на Item Response Theory как на современной теории конструирования и применения тестов, «представляет собой итерационный процесс, сходимость которого к оптимальным оценкам латентных параметров испытуемых обеспечивается путем минимизации ошибок измерения, меняющихся после выполнения каждого очередного задания, что позволяет меньшим числом заданий быстрее и точнее оценить уровень подготовленности учащихся» [4, с. 8].

В качестве примера адаптивного теста мы взяли Oxford Online Placement Test. Высокая надежность данного теста определяется в первую очередь тем, что он прошел апробацию в 60 странах мира; более 190000 студентов продавали этот тест; ведется постоянная работа по его усовершенствованию [5–7].

Целью теста является определение грамматических и прагматических знаний второго или иностранного языка и способностей тестируемых использовать эти знания для передачи сообщения, в процессе слушания и чтения.

В результате всестороннего исследования составители теста пришли к решению измерять грамматические и прагматические знания отдельно в The Use of English Section, так как грамматические знания и семантические значения – важнейшая предпосылка способности студента передавать сообщения, то есть вступать в процесс коммуникации наиболее подходящим способом.

Данная часть теста включает четыре задания (Language Knowledge Tasks).

Первое имеет своей целью измерить владение грамматической формой.

Второе направлено на измерение понимания семантического значения.

Третье представляет собой тест на измерение понимания грамматической формы и значения.

Четвертое, включенное в тест в 2010 году, измеряет знание прагматических значений, закодированных в ситуационном общении.

Вторая самостоятельная часть теста предназначена для измерения способности восприятия на слух и включает три части. Задания измеряют способность понимать как прямое (Literal), так и подразумеваемое (Implied) значения. Способность восприятия на слух оценивается с помощью коротких и длинных диалогов и монологов.

В основе создания данного теста лежали идеи Джеймса Пурпуа.

В работе этого автора [5] приведены грамматические формы, которые являются объектом оценки для каждого уровня; лексический материал и темы; аспекты общения для каждого уровня; подробные дескрипторы, определяющие, что обучающийся определенного уровня умеет делать на практике (понимать, использовать, контролировать).

Как любой адаптивный компьютерный тест, Oxford Online Placement Test экономит время, так как баллы подсчитываются автоматически. Тест также учитывает время, затраченное на его выполнение, он помогает определить общий уровень владения языком.

Нас привлекает в оксфордском тесте именно то, что он адаптивный. Тест адаптируется к уровню каждого испытуемого и предлагает ему только те задания, которые соответствуют его уровню. Выбор последующего вопроса зависит от того, как студент ответил на предыдущий вопрос. Это делает адаптивный тест уникальным для каждого.

Oxford Online Placement Test отличается от большинства подобных тестов тем, что оценивает не только знание грамматики и лексики, но и то, как студент использует это знание для понимания и передачи сообщения. Понимание значения является фундаментальным для общения. Это прекрасный индикатор языковой компетентности.

В целом же все специалисты (преподаватели английского языка, использующие адаптивные компьютерные тесты в процессе обучения) отмечают целый ряд преимуществ, которые позволяют выделить адаптивные компьютерные тесты в особую группу. При проведении таких тестов обеспечиваются:

– индивидуализация темпа тестирования для каждого, кто выполняет тест, что становится возможным благодаря соответствующему программному обеспечению;

– ситуация успеха за счет того, что исключается предъявление слабо подготовленным студентам излишне трудных заданий, которые при проведении обычного теста могут вызвать чувство страха и привести к стрессу;

– высокий уровень секретности, когда практически полностью исключается возможность списывания, подсказка и тому подобное;

– незамедлительное сообщение результатов тестирования каждому тестируемому сразу после окончания или работы над индивидуальным подобранным набором заданий;

– минимизация числа заданий, а значит, времени тестирования [4].

Большим достоинством такого тестирования является то, что оно ориентировано на индивидуализированное обучение. Его использование вызывает у студентов устойчивый познавательный интерес, выступает средством дополнительной мотивации [1; 8].

Того, кто занимается составлением тестов и использует их в учебных целях, в первую очередь интересует вопрос достоверности полученных результатов.

Среди студентов и абитуриентов существует мнение, что ничего не зная, можно просто случайно выбрать, например, один из четырех предложенных вариантов ответа (когда речь идет о выполнении заданий Multiple Choice) и получить достаточно высокий результат. Это не соответствует действительности.

Для расчета возможности события через шансы необходимо, чтобы наш тест с заданием типа Multiple Choice, включающий четыре варианта ответа, удовлетворял ряду условий: полноте набора, взаимоисключаемости событий, равновероятности событий.

Следующие задания из лексико-грамматического теста полностью соответствуют вышеперечисленным условиям:

We know that leading is all ... helping people.

a) for b) on c) to d) about

It was time for me to stop and ... at the beauty of the land and nature.

a) stare b) see c) look d) watch

Первый пример с предложениями соответствует всем вышеперечисленным условиям. Приведенные во втором примере события являются равновероятными, так как слова являются довольно близкими синонимами.

Приведем пример неправильного составления испытания:

Buy her flowers to show that you care ... her a lot.

- a) for b) on c) about d) after

В указанном примере два правильных ответа, то есть события не взаимоисключающие.

Еще один пример подобного рода.

She ... in this city for 15 years.

- a) has lived b) has been living c) is living
d) lives

В данном случае как минимум три варианта ответа можно считать правильными.

Перейдем к количественной оценке результатов достоверности наших испытаний. Возьмем тест, состоящий из 20 заданий, каждое из которых имеет 4 варианта ответа. В этом случае вероятность угадывания всех правильных ответов составляет $P = \frac{1}{4^{20}}$, т.е. настолько мала, что

событие можно считать маловероятным.

Вероятность угадывания 50% ответов соответственно также считается незначительной [9].

Различные виды тестирования с 2005 года используются преподавателями, работающими по программе «Переводчик в сфере профессиональной коммуникации» Центра дополнительного профессионального образования ННГУ. Практический опыт работы позволил нам сделать ряд выводов.

1. Количество заданий в тестах может не превышать двадцати, так как при полной достоверности результатов мы экономим время на проведение теста и его оценку.

2. Оптимальное количество вероятных ответов – 4. Это соответствует количеству ответов в тестах по английскому языку на ЕГЭ.

3. Тесты должны быть составлены правильно, т.е. удовлетворять условиям полноты, взаимоисключаемости, равновероятности.

4. Результаты такого тестирования достоверны. Преподаватели имели возможность сравнивать результаты тестирования и реаль-

ные знания, демонстрируемые студентами на протяжении трех лет обучения.

5. Несмотря на валидность и надежность классических тестов, будущее, несомненно, за компьютерными адаптивными тестами в силу отмеченных выше достоинств.

Список литературы

1. Золотова М.В. Эволюция тестов: от прообраза до компьютерного адаптивного теста // Обучение, тестирование и оценка: Материалы тринадцатой международной конференции. Н. Новгород: Изд-во НГЛУ им. Н.А. Добролюбова, 2013. С. 47–51.

2. Wainer H. Computerized Adaptive Testing: A Primer. 2nd Ed. Mahwah, NJ: Erlbaum Associates, 2000. 364 p.

3. Weiss, D.L. (ed.) New Horizons in Testing. N.Y.: Academic Press, 1988. 345 p.

4. Звонников В.И., Челышкова М.Б., Малыгин А.А. Адаптивное тестирование в дистанционном обучении // Высшее образование сегодня. 2012. № 6. С. 7–11.

5. Purpura J.E. Oxford Online Placement Test. What does it measure and how. – URL: <http://www.oxfordenglishtesting.com> (дата обращения 25.01.2014).

6. <http://www.cambridgeassessment.org.uk/images/122751-iaea-2004-paper-let-s-stop-marking-examples-pdf> (дата обращения 21.01.2014).

7. Purpura J.E. Strategy Use and Second Language Test Performance: A structural equation modeling approach. Cambridge: Cambridge University Press, 1999. P. 295.

8. Сидоренков В.А. Адаптивный тест в варианном обучении русскому языку // Русский язык в школе. 2013. № 7. С. 9–13.

9. Якимов А.В., Черепенников В.В. Анализ функциональных и вероятностных связей в высшей математике / Для студентов, обучающихся в рамках подготовки бакалавров/магистрантов по направлению «Биология». Н. Новгород: ННГУ, 1997. 14 с.

COMPUTER ADAPTIVE TESTS AND RELIABILITY OF ENGLISH LANGUAGE TESTING RESULTS

M.V. Zolotova, L.K. Delyagina, N.V. Kaminskaya, T.V. Martianova

An analysis is presented of the use of computer adaptive tests in the evaluation of the knowledge of English and the application of the theory of probability to prove the reliability of the results of this testing.

Keywords: English language, computer adaptive tests, reliability of test results.

References

1. Zolotova M.V. Evoliutsiia testov: ot proobraza do komp'iuternogo adaptivnogo testa // Obuchenie, testirovanie i otsenka: Materialy trinadsatoi mezhdunarodnoi konferentsii. N. Novgorod: Izd-vo NGLU im. N.A. Dobroliubova, 2013. S. 47–51.
2. Wainer H. Computerized Adaptive Testing: A Primer. 2nd Ed. Mahwah, NJ: Erlawrence Erlbaum Associates, 2000. 364 p.
3. Weiss, D.L. (ed.) New Horizons in Testing. N.Y.: Academic Press, 1988. 345 p.
4. Zvonnikov V.I., Chelyshkova M.B., Malygin A.A. Adaptivnoe testirovanie v distantsionnom obuchenii // Vysshee obrazovanie segodnia. 2012. № 6. С. 7–11.
5. Purpura J.E. Oxford Online Placement Test. What does it measure and how. – URL: <http://www.oxfordenglishtesting.com> (data obrashcheniia 25.01.2014).
6. <http://www.cambridgeassessment.org.uk/images/122751-iaea-2004-paper-let-s-stop-marking-examps-.pdf> (data obrashcheniia 21.01.2014).
7. Purpura J.E. Strategy Use and Second Language Test Performance: A structural equatoin modeling approach. Cambridge: Cambridge University Press, 1999. P. 295.
8. Cidorenkov V.A. Adaptivnyi test v variantnom obuchenii russkomu iazyku // Russkii iazyk v shkole. 2013. № 7. С. 9–13.
9. Iakimov A.V., Cherepennikov V.V. Analiz funktsional'nykh i veroiatnostnykh sviazei v vysshei matematike / Dlia studentov, obuchaiushchikhsia v ramkakh podgotovki bakalav-rov/magistrantov po napravleniiu «Biologiya». N. Novgorod: NNGU, 1997. 14 s.