

УДК 37.022+004

Специфика компьютерного тестирования в образовании и его формы

Петрова Н.П.

Ключевые слова: педагогические измерения, инновационное тестирование, тестовый контроль, компьютерное тестирование, формы тестирования, ответственная деятельность.

Педагогический контроль представляет собой единую дидактическую и методическую систему проверочной деятельности, которая протекает при руководящей и организующей роли педагогов, носит совместный характер, объединяя преподавателей и студентов, и направлена на оценку результатов учебного процесса. С помощью контроля можно оценить достижения студентов и выявить пробелы в их знаниях, установить взаимосвязь между планируемыми, реализуемыми и достигнутыми уровнями образования, понять достоинства и недостатки новых методов обучения, сравнить работу преподавателей, дать руководителю учебного заведения объективную информацию для принятия управленческих решений и выполнить ряд других, не менее важных задач [1].

При тестовом контроле операционализация понимается как процедура перехода от понятийных индикаторов к эмпирическим референтам, в роли которых выступают задания теста. Модель результатов предельно стандартизируется и задается в виде правил оценивания и ключа ответов, сличение же результатов тестируемых с ключом ответов проводится с минимизацией субъективного фактора, автоматизировано либо с помощью экспертов.

На протяжении всей истории развития тестов отношение к ним не было однозначным. В различных странах в

ориентации общественного мнения за или против тестов побеждают то их противники, то их сторонники.

Критика тестов, если она не сводится лишь к отрицанию, а носит конструктивный характер, оказывает позитивное влияние на их развитие. Она порождает теоретические и прикладные исследования, способствует развитию аппарата теории педагогических измерений. Если сопоставить упрощенные подходы к конструированию тестов начала XX в. и современные методы компьютерного моделирования тестов из банка калиброванных заданий, то можно заметить большой прогресс, благодаря которому ликвидированы многие недостатки тестовых методов, повышена их объективность и технологичность. Даже наиболее убежденные противники тестов вынуждены согласиться с тем, что сегодня, когда приходится принимать множество управленческих решений в образовании, полный отказ от тестов невозможен.

Многие преподаватели далеки от абсолютизации роли тестирования в учебном процессе, поскольку всегда анализируют обоснованность тестовых оценок и правомерность их применения для различных ситуаций в обучении, что вполне оправданно. Во-первых, тесты – это только инструмент, средство осуществления педагогического контроля, и, как любое средство, они могут приносить пользу, если применяются по назначению, или быть

неуместными, когда их функциональное назначение не адекватно ситуации применения. Во-вторых, тесты могут быть сделаны хорошо или плохо. В последнем случае они не обеспечивают ни высокой объективности, ни сопоставимости, поскольку требования теории педагогических измерений не выполняются. В-третьих, даже очень качественные тесты при неумелом их использовании представляют опасность. Необходимы специальные знания для корректного выбора теста из числа имеющихся – в соответствии с целями измерения; инструктирования тестируемых; подбора адекватных методов шкалирования результатов студентов; выравнивания шкал по отдельным вариантам и правильной интерпретации тестовых баллов при использовании результатов в учебном процессе. Негативные последствия неумелого применения тестов нередко отмечаются в тех случаях, когда при сопоставлении результатов студентов не учитываются важные факторы, влияющие на результаты тестирования, например индивидуальные особенности конкретного студента и дополнительная информация о нем, релевантная целям измерения.

В педагогических инновациях появилось отдельное направление – компьютерное тестирование, при котором предъявление тестов, оценивание результатов студентов и выдача им результатов осуществляются с помощью компьютера.

Этап генерации тестов технологически может протекать по-разному, в том числе путем ввода в компьютер бланковых тестов. На сегодняшний день по компьютерному тестированию имеются многочисленные публикации, разработаны программно-инструментальные средства для генерации и предъявления тестов [2].

Хотя компьютерное тестирование значительно облегчает работу преподавателя при предъявлении и оценивании результатов выполнения тестов, его рас-

пространение во многом не более чем дань моде, все негативные последствия которого до сих пор не выявлены в полной мере. Выбор компьютерного формата экзамена должен основываться на более важных и обоснованных предположениях, чем просто увлечение инновациями, поскольку он порождает множество проблем и ставит студентов в неравные условия. Обращаться к компьютерному тестированию следует в тех случаях, когда есть настоятельная потребность в отказе от традиционных бланковых тестов.

Например, компьютерное тестирование необходимо при проведении ЕГЭ в труднодоступных районах России. Сбор выпускников школ отдаленных районов в обозначенное время проведения ЕГЭ становится настолько сложным и дорогостоящим мероприятием, что обойтись без компьютерного тестирования и современных средств коммуникации просто невозможно. Компьютерное тестирование целесообразно также применять при проведении экзаменов для детей с ограниченными возможностями, имеющих серьезные нарушения зрения или слуха. С помощью компьютера можно использовать большие по размерам шрифты, аудиозаписи, дополнительные устройства для ввода данных тестирования и другие приспособления, компенсирующие на экзаменах потенциальное отставание детей с ограниченными возможностями [3].

Компьютерное тестирование может проводиться в различных формах, различающихся по технологии объединения заданий в тест (рис. 1). Часть из них пока не получила специального названия в литературе по тестовой проблематике.

Первая форма – самая простая. Готовый тест, стандартизованный или предназначенный для текущего контроля, вводится в специальную оболочку, функции которой могут различаться по степени



Рис. 1. Формы компьютерного тестирования

полноты. Обычно при итоговом тестировании оболочка позволяет предъявлять задания на экране, оценивать результаты их выполнения, формировать матрицу результатов тестирования, обрабатывать ее и шкалировать первичные баллы испытуемых путем перевода в одну из стандартных шкал для выдачи каждому испытуемому тестового балла и протокола его оценок по заданиям теста.

Вторая форма компьютерного тестирования предполагает автоматизированную генерацию вариантов теста, осуществляемую с помощью инструментальных средств. Варианты создаются перед экзаменом или непосредственно во время его проведения из банка калиброванных тестовых заданий с устойчивыми статистическими характеристиками. Калибровка достигается благодаря длительной предварительной работе по формированию банка, параметры заданий которого получают на репрезентативной выборке студентов, как правило, на протяжении 3–4 лет с помощью бланковых тестов. Содержательная валидность и параллельность вариантов обеспечиваются за счет строго регламентированного отбора заданий каждого варианта в соответствии со спецификацией теста.

Третья форма – компьютерное адаптивное тестирование – базируется на специальных адаптивных тестах. В основе идей адаптивности лежат сообра-

жения о том, что студенту бесполезно давать задания теста, которые он выполнит наверняка правильно без малейших затруднений или гарантированно не справится с ними в силу высокой трудности. Поэтому предлагается оптимизировать трудность заданий, адаптируя ее к уровню подготовленности каждого испытуемого, и сократить за счет исключения части заданий длину теста.

Компьютерное тестирование имеет определенные преимущества по сравнению с традиционным бланковым тестированием, которые проявляются особенно заметно при массовых проверках, например при проведении национальных экзаменов типа ЕГЭ. Предъявление вариантов теста на компьютере позволяет сэкономить средства, расходуемые обычно на печать и транспортировку бланковых тестов.

Благодаря компьютерному тестированию можно повысить информационную безопасность и предотвратить рассекречивание теста за счет высокой скорости передачи информации и специальной защиты электронных файлов. Упрощается также процедура подсчета результирующих баллов в тех случаях, когда тест содержит только задания с выбором ответов [4].

Другие преимущества компьютерного тестирования проявляются в текущем контроле, при самоконтроле и

самоподготовке студентов; благодаря компьютеру можно незамедлительно выдать тестовый балл и принять неотложные меры по коррекции усвоения нового материала на основе анализа протоколов по результатам выполнения корректирующих и диагностических тестов. Это отражено в электронных учебниках, разработанных в опытно-экспериментальной лаборатории профессионально-информационной педагогики и методики преподавания технологий ЮФУ. Возможности педагогического контроля при компьютерном тестировании значительно увеличиваются за счет расширения спектра измеряемых умений и навыков в инновационных типах тестовых заданий, использующих многообразные возможности компьютера при включении аудио- и видеофайлов, интерактивности, динамической постановки проблем с помощью мультимедийных средств и др.

Благодаря компьютерному тестированию повышаются информационные возможности процесса контроля, появляется возможность сбора дополнительных данных о динамике прохождения теста отдельными студентами и для осуществления дифференциации пропущенных и не достигнутых заданий теста.

Помимо неоспоримых достоинств, компьютерное тестирование имеет ряд существенных недостатков, которые представлены на рис. 2.

Обычно психологические и эмоциональные реакции студентов на компьютерное тестирование носят позитивный характер. Студентам нравится незамедлительная выдача тестовых баллов, протокола тестирования с результатами по каждому заданию, а также сам инновационный характер контроля в том случае, когда привлекаются современные гипермедийные технологии для выдачи теста.

Инновационные задания, использующие возможности компьютерного тестирования, на сегодняшний день являются наиболее перспективным направлением развития автоматизации педагогических измерений. Основной причиной этого является большой потенциал инновационных заданий для повышения информативности педагогических измерений и увеличения содержательной валидности тестов.

Основная цель разработки инновационных заданий для компьютерного тестирования состоит в оценивании тех когнитивных умений, функциональной грамотности и коммуникативных умений, которые остаются не выявленными при традиционном контроле или использовании бланковых тестов.

Предметом оценивания при инновациях может быть уровень аналитико-синтетической деятельности обучаемого, скорость обобщения новой информации, гибкость мыслительного процесса и многие другие показатели умственной

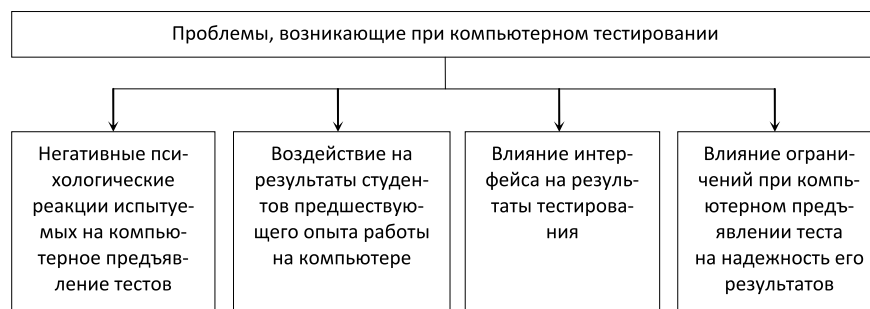


Рис. 2. Проблемы, возникающие при компьютерном тестировании

деятельности, сформировавшиеся в процессе обучения и не поддающиеся оцениванию с помощью обычных тестов.

В использовании инновационных заданий можно выделить два аспекта: дидактический и психолого-педагогический. Первый предполагает развернутую содержательную интерпретацию результатов тестирования в контексте освоенных на момент предъявления теста когнитивных, учебных и общеучебных умений, а второй позволяет оценить уровень развития мыслительных процессов у студента и выявить особенности усвоения им новых знаний. Большинство инновационных заданий, разработанных к настоящему времени в ЮФУ, обеспечивают совершенствование измерений в обоих направлениях. Таким образом, инновационные задания позволяют расширить возможности самого педагогического измерения за счет получения результатов в новых, недоступных ранее направлениях оценивания качества подготовленности студентов. Например, для оценивания уровня сформированности функциональной грамотности экзаменуемым можно предложить отрывок текста, в котором есть ошибки, а затем попросить идентифицировать их и исправить путем перепечатывания разделов текста.

Инновационные задания способствуют сокращению влияния случайного угадывания за счет увеличения числа возможных ответов без нарастания громоздкости заданий теста. Например, при оценивании понимания прочитанного текста можно попросить студента выбрать ключевое предложение в тексте и указать на него щелчком мыши. Таким образом, каждое предложение в текстовом отрывке становится опцией для выбора вместо 4–5 ответов в традиционных заданиях с готовыми ответами. Для совершенствования формы заданий используют сложный рисунок, динами-

ческие элементы, включая изображения, мультипликацию или видео; тем самым сокращается время чтения условия. Расширение возможностей тестирования происходит при включении звука, что позволяет вести диалог со студентом, оценивать фонетические особенности его произношения при тестировании по иностранному языку, проверять правильность интерпретации различных звуков.

Инновации при разработке заданий для компьютерного тестирования охватывают пять связанных между собой направлений. К ним относятся: форма задания, действия испытуемого при ответе, уровень использования мультимедийных технологий, уровень интерактивности и методика подсчета баллов [5].

Нововведения в форме задания включают визуальный и звуковой информационные ряды или их сочетание. Визуальная информация может носить реалистический (фото, кино) и синтезированный (рисунок, анимация) характер. Тип информации в сочетании с тестовой формой определяет формат ответа, выбираемого или создаваемого экзаменуемым. При использовании фотографий или рисунков информация, содержащаяся в тестовых заданиях, носит статический характер. Анимация вносит динамику в выполнение теста.

Действия студента при ответе на задания зависят от тех инновационных средств, которые включены в тест. При включении в задания звуковой информации, предполагающей голосовой ответ студента, для ответа используются клавиатура, мышь или микрофон. Значительное место при ответах отводится интерактивным процессам. Интерактивный режим работы студентов при компьютерном тестировании означает поочередную выдачу аудиовизуальной информации, при которой каждое новое высказывание со стороны студента или

компьютера строится с учетом предыдущей информации с той и другой стороны. При организации интерактивного режима в компьютерном тестировании используется в основном экранное меню, в котором студент для ответа на тестовые задания выбирает, создает или перемещает объекты – компоненты ответа. Реже в интерактивном режиме применяют голосовой ввод ответа.

В целом уровень интерактивности, обеспеченный в компьютерном тестировании, характеризует степень, в которой определенная форма задания реагирует или отвечает на ввод информации со стороны экзаменуемого. Этот уровень варьируется от простейшего случая, когда совершается один шаг, до сложных, многошаговых заданий с разветвлением после каждого очередного ответа студента.

Задания повышенной трудности всегда требуют больше времени для ответов, вне зависимости от того, предъявляются ли они с помощью компьютерного моделирования виртуальной реальности, имеют ли форму лабораторной работы, эссе или используют мультимедийные технологии. Из-за временных затрат число сложных заданий должно быть незначительно – не более 10–15%, в отдельных случаях – 20–25%. Многообразие звуковых и зрительных образов в компьютерном тестировании приводит к возникновению у школьников усталости, поэтому при включении в тест даже небольшого количества трудных инновационных заданий приходится значительно уменьшать длину теста, что негативно сказывается на содержательной валидности, надежности и информационной безопасности педагогического измерения [6].

Несмотря на преимущества инновационных форм заданий, предъявляемых с помощью компьютера, к ним нужно относиться с осторожностью, тщательно

анализировать их адекватность целям измерения и уместность в тесте. Обычно инновационные задания высокой трудности выделяют в отдельный блок и помещают в конце теста. Их выполнение не должно отнимать времени у наиболее слабых студентов, которые, скорее всего, не дойдут до конца теста.

Если в компьютерном тестировании не используются мультимедийные и интерактивные технологии, то подсчет первичных баллов студентов проводится традиционно, путем суммирования оценок по отдельным заданиям. Привлечение мультимедийных технологий приводит к многомерности результатов выполнения теста, поскольку оценивание целого спектра творческих, коммуникативных, общепредметных и других умений с помощью инновационных форм заданий всегда связано с несколькими переменными измерения. Появление интерактивности еще больше усложняет процедуру подсчета баллов студентов, она становится зависимой от ответа экзаменуемого на каждом шаге выполнения заданий теста и требует полигамических оценок.

Проверка результатов выполнения заданий с конструируемым регламентированным ответом осуществляется путем сравнения ответа экзаменуемого с эталоном, хранящимся в памяти компьютера, и включает различные синонимы правильного ответа с приемлемыми орфографическими ошибками.

Намного сложнее автоматизированный подсчет баллов в заданиях со свободно конструируемым ответом (типа эссе) в гуманитарных дисциплинах. На сегодняшний день зарубежными тестологами разработаны специальные программы для автоматизированной проверки эссе. Критерии оценивания в этих программах довольно разнообразны: от рассмотрения поверхностных характеристик эссе типа длины и степе-

ни полноты ответа до сложных случаев анализа с использованием достижений компьютерной лингвистики. Обычно все эти различные автоматизированные программы подсчета баллов требуют участия экспертов только на момент начала работы, когда квалифицированным педагогам необходимо «обучить» компьютерную программу оцениванию любых развернутых ответов [7; 8].

Согласно современным воззрениям, учебный процесс – динамичная управляемая система, в которой управляющей подсистемой является педагог, объектом управления – обучающиеся, предметом управления – их учебно-познавательная деятельность. Как известно, субстанциальную основу управления составляет получение достоверной своевременной информации об управляемой системе, поэтому вооружение педагога методами и средствами организации педагогического контроля учебно-познавательной деятельности обучающихся является важным фактором повышения эффективности педагогического управления. Функции педагогического контроля направлены прежде всего на получение информации, необходимой для педагогического управления, а соблюдение его принципов обеспечивает успешную реализацию основных функций [9].

Библиография

1. Бондаревская Е.В. Гуманистическая парадигма лично-ориентированного образования // Педагогика. 1997. № 4.
2. Звонников В.И., Чельшкова М.Б. Современные средства оценивания результатов обучения: учеб. пособие для студ. вузов. 3-е изд., стереотип. М.: Академия, 2009.
3. Михайлычев Е.А., Карпова Г.Ф., Леонова Е.Е. Педагогическая диагностика: история, теория, современность. Ростов н/Д, 2002.
4. Современные методы и средства оценки обученности: учеб. пособие / С.Б. Полянская [и др.]. Славянск-на-Кубани: ИЦ СГПИ, 2008.
5. Звонников В.И., Чельшкова М.Б. Указ. соч.

6. Ковалева Г.С. Зарубежный опыт построения и актуальные проблемы развития тестовых систем // Российский и зарубежный опыт построения систем образовательного тестирования: материалы к семинару «Актуальные проблемы построения системы национальных образовательных стандартов и тестирования». М., 2000.
7. Соколов В.М. Роль и место тестов достижений в диагностике качества образования // Вестник Нижегородского университета. 2006. Вып. 1.
8. Van der Linden, W.J. and C.A.W. Glas (Ed.), 2003. Computerized Adaptive Testing: Theory and Practice. London: Kluwer academic publishers, 2003.
9. Du Toit, M., 2003. IRT from SSI: Bilog-mg Multilog. Parscale Testfast. Lincolnwood: Scientific Software International.
10. Современные методы и средства оценки обученности: учеб. пособие / С.Б. Полянская [и др.].

Bibliography

1. Bondarevskaya, E.V., 1997. Humanist paradigm personality oriented education. Pedagogy, 4. (rus)
2. Zvonnikov, V.I. and M.B. Chelyshkova, 2009. Modern means of evaluating learning outcomes: manual. 3rd ed. Moscow. (rus)
3. Mikhaylychev, Ye.A., G.F. Karpova and E.E. Leonova, 2002. Pedagogical diagnostics: history, theory, modernity. Rostov-on-Don. (rus)
4. Polyanskaya, S.B. et al., 2008. Modern methods of training and assessment tools: manual. Slavyansk-on-Kuban. (rus)
5. Zvonnikov, V.I. and M.B. Chelyshkova. Op. cit.
6. Kovaleva, G.S., 2000. Foreign experience of building and urgent problems of development of test systems. In: Russian and foreign experience building systems educational testing: materials for the seminar "Actual problems of the construction of national educational standards and testing". Moscow. (rus)
7. Sokolov, V.M., 2006. Role and place of achievement tests in the diagnosis of the quality of education. Bulletin of the Nizhny Novgorod University, 1. (rus)
8. Van der Linden, W.J. and C.A.W. Glas (Ed.), 2003. Computerized Adaptive Testing: Theory and Practice. London: Kluwer academic publishers.
9. Du Toit, M., 2003. IRT from SSI: Bilog-mg Multilog. Parscale Testfast. Lincolnwood: Scientific Software International.
10. Polyanskaya, S.B. et al. Op. cit.