

УДК 378

АВТОМАТИЗАЦИЯ РАСЧЕТА ОСНОВНЫХ ХАРАКТЕРИСТИК ТЕСТА**Р. Р. Абдулвелеева****Р. И. Абдулвелеев****Н. В. Воронина**

*Кандидат педагогических наук,
доцент,
ведущий программист,
студентка,
Орский гуманитарно-технологический
институт, Оренбургский государственный
университет (филиал),
г. Орск, Оренбургская область, Россия*

**AUTOMATION OF CALCULATION
OF THE MAIN CHARACTERISTICS OF THE TEST****P. R. Abdulveleeva****R. I. Abdulveleev****N. V. Voronina**

*Candidate of pedagogical sciences,
assistant professor,
Lead programmer,
student,
Orsk Humanitarian Institute of Technology,
Orenburg State University (branch),
Orsk, Orenburg region, Russia*

Summary. In this paper we describe the use of the computer program intended for the analysis of result of test. The program is created in the environment of Excel with use of Visual Basic for Application, allows automation of calculating the main mathematic-statistical characteristics of test.

Keywords: testing; approbation of test; validity; reliability; automation of processing of results of test.

Контроль знаний учащихся можно осуществлять различными способами. Одним из современных средств оценивания результатов обучения становится тестирование. Процедура диагностики достижений учащихся с помощью тестов включает в себя процесс разработки диагностических материалов или теста, организацию процесса тестирования, обработку и интерпретацию результатов тестирования, анализ полученных результатов и выявление качества диагностических заданий.

Эффективная реализация диагностики в учебном процессе школы возможна при использовании информационных технологий. Большую роль при этом играет наличие у будущих учителей знаний и умений подбора, применения инструментария и проведения процедуры диагностики знаний учащихся. Методическая готовность учителя включает в себя умение осуществлять компьютерную диагностику знаний, которая позволяет оптимизировать процесс оперативного опреде-

ления результатов обучения, осуществлять оперативную обратную связь, корректировать имеющиеся недостатки в процессе подготовки [1].

Не всегда у учителя для контроля знаний учащихся имеется нужный тест или комплекс диагностических заданий. Зачастую педагогу необходимо самостоятельно разработать тестовые задания, а значит необходимо обосновать их качество, воспользовавшись современной теорией педагогических измерений.

Основными характеристиками, которые измеряются при апробации и обработке результатов тестирования являются: величина трудности заданий теста, дискриминативности, корреляции заданий, дисперсии, надежности, валидности, корреляции теста [4]. Расчет этих характеристик состоит в обработке большого количества числовых данных, в использовании громоздких формул и в проведении множества промежуточных вычислений. Иногда после проведения тестирования требуется

коррекция тестовых заданий, а это значит, что для повторной апробации теста все расчеты нужно повторить. Если расчеты делать вручную с помощью калькулятора, то это весьма затруднительно и требует много времени. Поэтому возникает потребность автоматизации расчета основных статистических характеристик теста.

В целях автоматизации расчета характеристик теста воспользуемся компьютерной программой Excel. Табличный процессор Excel входит в пакет Microsoft Office и предназначен для работы с табличными данными. В программу встроено большое количество формул, которые позволяют производить самые сложные расчеты. Автоматическое проведение вычислений при изменении исходных данных в таблице упрощает расчеты при повторных измерениях. Удобное построение диаграмм, гистограмм, графиков упрощает проведение анализа и интерпретации полученных результатов. Использование языка программирования Visual Basic Applications (VBA) значительно упрощает процесс оценки качества теста, так как позволяет создать интерфейс, функционирующий в диалоговом режиме, который не требует от пользователя глубоких знаний программирования, вычислительной техники и математики.

Проводить тестирование можно как над одним или несколькими испытуемыми, так и с несколькими группами людей. В педагогической практике чаще всего применяют групповые тесты, так как они являются более удобными для школьной практики в системе классно-урочной формы обучения. Использование индивидуальных тестов происходит тогда, когда исследуемое нельзя измерить в условиях проверки группы людей. В наши дни групповое тестирование проводится чаще, чем индивидуальное, однако уровень развития компьютерной техники, программного обеспечения, сетевых и интернет технологий позволяют эффективно организовать индивидуальное тестирование.

Процесс информатизации процесса обучения позволил повысить возможности тестирования. Современные тестовые компьютерные программы и тестирующие системы имеют расши-

ренные возможности в определении режима тестирования, режима работы программы. Они позволяют задать жесткий режим тестирования: запретить возвращаться к предыдущим заданиям, чтобы исправить ответ, ограничить время тестирования, ограничить количество попыток тестирования, осуществить случайный выбор определенного количества заданий по теме или разделу из обширной базы заданий, запретить выход из программы до окончания тестирования, отключить доступ к рабочему столу и панели задач.

Качество разработанного теста определяется значениями существенных показателей. Основными показателями качества теста являются коэффициенты надежности и валидности тестовых заданий.

В большинстве случаев надежность определяется как коэффициент. Источники неудовлетворительной надежности тестов могут быть распределены на три группы связанные с качеством теста, с процедурой его проведения и оценивания и связанные с испытуемыми.

Качество тестовых заданий может значительно повысить надежность тестов. Этому способствует увеличение длины теста, проверка объективности критериев оценки, снижение вероятности угадывания правильных ответов, ужесточение инструкций для учащихся, повышение качества инструктажа и подготовки ведущих. При определении надежности теста следует обратить внимание на критериально-ориентированные тесты. Такие тесты могут целиком состоять из заданий, с которыми справляются все учащиеся и это позволяет педагогу констатировать тот факт, что тему или раздел усвоили все учащиеся, или не справляется никто из учащихся. В этих случаях значение показателя надежности будет иметь неудовлетворительные значения.

Еще одной важнейшей характеристикой теста является его валидность. Валидность особенно важна для критериально-ориентированных тестов, так как определение надежности для этих тестов затруднено. Валидность и надежность — связанные понятия. Валидность определяет, насколько тест отражает то, что он должен оценивать.

Автоматизировать процесс апробации теста и оценить его качество можно с помощью компьютерной программы [3]. Разработанный педагогом тест предлагается испытуемым. Форма предъявления заданий испытуемым может быть различной: в бумажном варианте или в виде тестовой компьютерной программы. Результаты выполнения тестовых заданий заносятся в матрицу результатов теста созданную в Excel.

Существует большое количество различных тестовых программ и оболочек для создания тестов. Для предъявления заданий испытуемым использовался пакет программ AD Tester предназначенный для проведения тестирования. В AD Tester в режиме конструктора вносятся разработанные тестовые задания, а затем в режиме тестирования испытуемые выполняют задания теста. В этой программе есть полезная функция – «экспорт результатов выполнения заданий теста в Excel». Использо-

вание этой функции упрощает процесс переноса результатов выполнения тестов испытуемыми в матрицу результатов (рис. 2). Это производится простым копированием ячеек. Далее внесенные значения используются для обработки и расчета различных параметров теста. Эти параметры и промежуточные характеристики рассчитываются методом ввода формул в соответствующие ячейки таблицы листа Excel. Для расчета различных параметров теста использовались формулы, содержащиеся в методике разработанной и предложенной М. Б. Челышковой [4]. Полученные промежуточные результаты позволяют рассчитать валидность и надежность теста, которые дают возможность выявить необходимость внесения корректив в содержание тестовых заданий (заменить, усложнить или облегчить).

Главное окно программы автоматизации расчета характеристик теста показано на рис. 1.

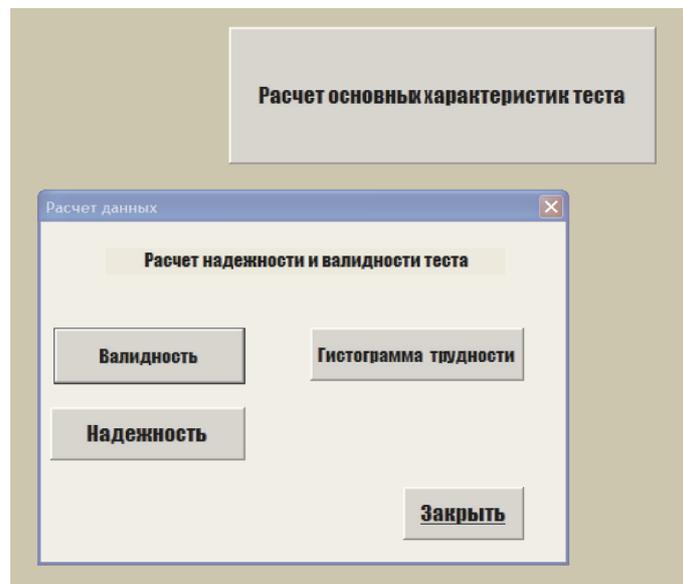


Рис. 1. Главное окно программы

Окно программы содержит кнопку расчета основных характеристик теста. Нажатие кнопки вызывает форму для расчета надежности и валидности теста. На этой же форме расположена кнопка для просмотра гистограммы трудности заданий теста.

Методика расчета основных характеристик теста состоит из нескольких этапов. Одним из этапов апробации разработанного теста является расчет валидности и надежности теста, для этого сначала формируют результирующую матрицу ответов испытуемых

В ПОМОЩЬ УЧИТЕЛЮ

в табличном процессоре Excel (рис. 2), в которой строки соответствуют испытуемым, столбцы – заданиям теста, а в ячейки таблицы экспортируются ответы испытуемых, задаваемые в некоторой шкале. В разработанной компьютерной программе используется дихотомическая шкала (0 – не выполнено, 1 – выполнено).

В матрицу результатов вносят следующие данные: фамилия, имя испытуемых, количество заданий, результат ответа каждого испытуемого на каждое

задание. Количество правильных ответов по каждому вопросу рассчитывают суммированием единиц по столбцам, количество верных ответов испытуемым по всем вопросам рассчитывают суммированием единиц по строкам. Для подсчета общего количества правильных ответов вначале суммируют количество верных ответов каждого испытуемого (сумма единиц по строке), и затем суммируют количество верных ответов всех испытуемых. В последнем столбце, получаем результат суммы.

| Повторная обработка результатов тестирования в экспериментальной группе | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|-----------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| # | ФИО | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 1 | Студент 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| 2 | Студент 2 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| 3 | Студент 3 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 4 | Студент 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| 5 | Студент 5 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| 6 | Студент 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 7 | Студент 7 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 |
| 8 | Студент 8 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | Студент 9 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| число правильных ответов | | 9 | 6 | 8 | 7 | 8 | 3 | 7 | 6 | 8 | 7 | 7 | 8 | 7 | 5 | 6 | 7 | 8 | 5 | 6 | 4 | 5 | 4 | 4 | 6 | 5 | 4 | |

| Упорядоченная матрица данных тестирования | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|-----------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| # | ФИО | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 8 | Студент 4 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | Студент 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| 6 | Студент 2 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | |
| 5 | Студент 7 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 3 | Студент 5 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| 2 | Студент 3 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 7 | Студент 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 1 | Студент 9 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |

Рис. 2. Матрица результатов

Далее вносят данные во вторую матрицу результатов испытуемых – упорядоченную по строкам результатов правильных ответов от общего количества заданий. Результаты каждого испытуемого упорядочены по возрастанию, то есть распределены от наименьшего результата к наибольшему (рис. 3).

Затем производится расчет основных математико-статистических значений: расчет индивидуальных баллов теста, дисперсии, среднее выборочное значение, надежность теста, валидность.

Валидность позволяет определить способность тестовых заданий выявить то, ради чего проводилось тестирование.

| | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | |
|---|------------------|------|------|------|------|------|------|-------|-------|--|
| 2 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | всего | | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 36 | |
| | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 18 | 324 | |
| | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 23 | 529 | |
| | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 29 | 841 | |
| | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 29 | 841 | |
| | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 34 | 1156 | |
| | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 34 | 1156 | |
| | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 35 | 1225 | |
| | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 38 | 1444 | |
| | 6 | 6 | 6 | 5 | 6 | 6 | 6 | 246 | 60516 | |
| | 0,78 | 0,67 | 0,67 | 0,67 | 0,56 | 0,67 | 0,67 | 0,67 | | |
| | 0,22 | 0,33 | 0,33 | 0,33 | 0,44 | 0,33 | 0,33 | 0,33 | | |
| | 0,82 | 0,54 | 0,11 | 0,70 | 0,37 | 0,32 | 0,19 | 0,57 | | |
| | 0,78 | 0,67 | 0,67 | 0,67 | 0,56 | 0,67 | 0,67 | 0,67 | 27,33 | |
| | надежности теста | | | | | | | | 39 | |

Рис. 3. Упорядоченная по результатам матрица

Чтобы обеспечить удобство пользователю в эксплуатации программы при получении результатов расчета валидности и надежности теста были использованы возможности языка программирования VBA. С помощью кодов программы была создана кнопочная форма. Основ-

ная кнопка формы «Расчет основных характеристик теста». При нажатии на эту кнопку появляется окно формы, в котором созданы объекты – кнопки «валидность», «надежность» «выход» (рис. 4).

Результат нажатия на кнопку «валидность», показан на рис. 5.

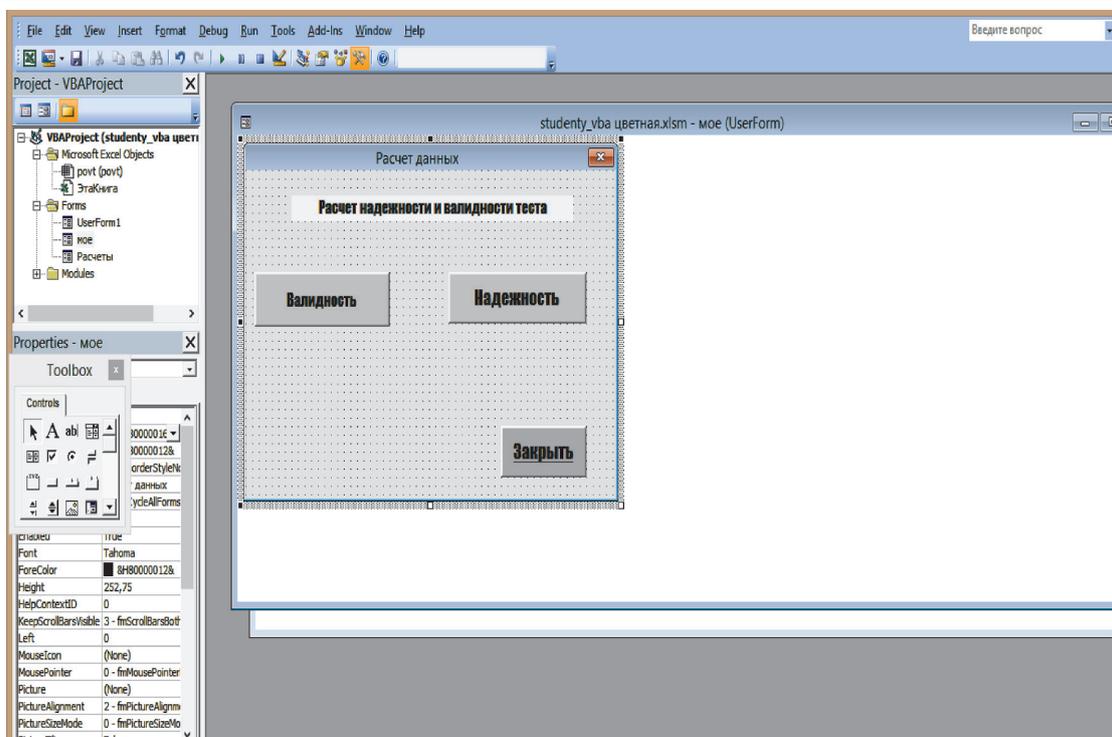


Рис. 4. Форма окна расчета характеристик теста

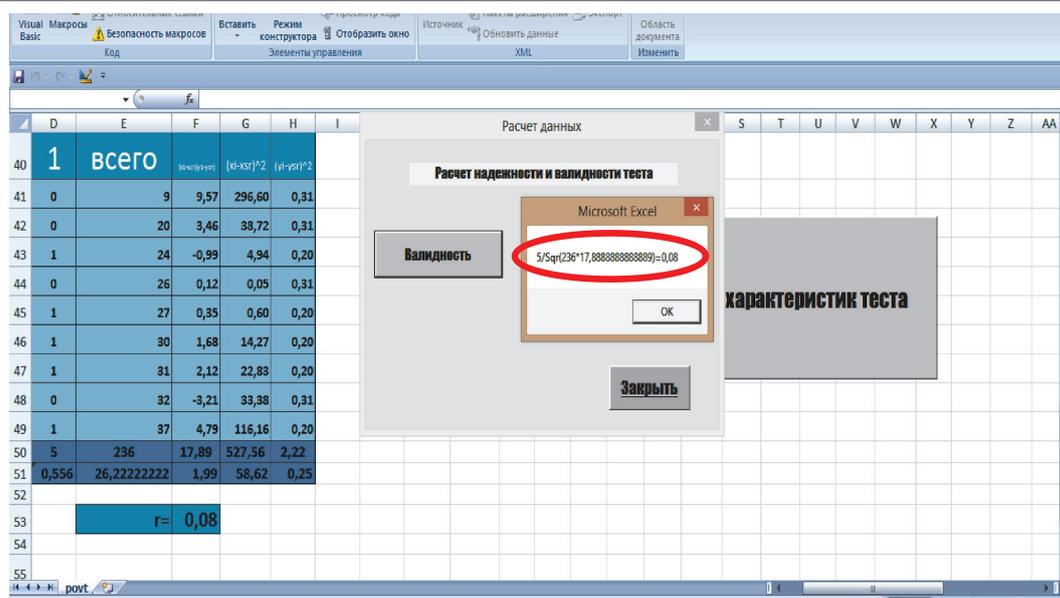


Рис. 5. Результат расчета валидности

Разработанная компьютерная программа, предназначена для автоматизации расчета основных характеристик теста, вычисления валидности и надежности теста. В случае неудовлетворительных результатов, разработчик теста производит коррекцию содержания тестовых заданий, количества заданий и осуществляет повторную апробацию теста.

Таким образом, компьютерная программа позволяет, во-первых, автоматизировать и повысить качество экспертизы, во-вторых, повысить эффективность процесса апробации разработанного теста и обеспечить достоверное определение результатов тестирования и, в-третьих, свести к минимуму недостатки тестового контроля знаний.

Создание таких программ позволит обеспечить представление педагога о теории педагогических измерений и ее применении в оценке качества разработанного теста, повысить профессиональную культуру и компетентность педагога в области современных средств оценивания результатов обучения, стимулировать будущих педагогов к изучению возможностей информационных технологий с целью применения в исследовательской деятельности, рефлексии своей профессиональной деятельности [2].

Библиографический список

1. Абдулвелеева Р. Р. Технология моделирования компьютерной диагностики методической готовности студентов педагогического вуза // Информатизация образования и науки. – 2014. – № 3 (23). – С. 42–54
2. Абдулвелеева Р. Р. Диагностика уровня развития рефлексии учебно-профессиональной деятельности студента // Сборник научных трудов SWorld. – № 1. – Т. 19. – Одесса, 2013. – С. 68–72.
3. Компьютерная программа оценки качества теста / Р. Р. Абдулвелеева, Н. В. Воронина // Мир науки и инноваций. – Вып. № 1. Т. 6. – Иваново : Научный мир, 2015. – С. 68–71
4. Чельшкова М. Б. Теория и практика конструирования педагогических тестов: учебное пособие. – М. : Логос, 2002. – 432 с.

Bibliograficheskij spisok

1. Abdulveleeva R. R. Tehnologija modelirovanija kompjuternoj diagnostiki metodicheskoj gotovnosti studentov pedagogicheskogo vuza // Informatizacija obrazovanija i nauki. – 2014. – № 3 (23). – S. 42–54
2. Abdulveleeva R. R. Diagnostika urovnja razvitija refleksii uchebno-professionalnoj dejatel'nosti studenta // Sbornik nauchnyh trudov SWorld. – № 1. T. 19. – Odessa, 2013. – S. 68–72.
3. Kompjuternaja programma ocenki kachestva testa / R. R. Abdulveleeva, N. V. Voronina // Mir nauki i innovacij. – Vyp. № 1. T. 6. – Ivanovo : Nauchnyj mir, 2015. – S. 68–71
4. Chel'shkova M. B. Teorija i praktika konstruirovanija pedagogicheskix testov: Uchebnoe posobie. – M. : Logos, 2002. – 432 s.

© Абдулвелеева Р. Р., Абдулвелеев Р. И., Воронина Н. В., 2015