

бавлять тайм-коды на протяжении всего видео. Текстовые материалы необходимо подкреплять наглядными примерами, инфографикой, диаграммами, таблицами.

Подводя итог, следует отметить, что для достижения максимального восприятия визуально представленной информации необходимо:

- 1) руководствоваться принципами представления информации на экране;
- 2) обеспечить курсы, расположенные на платформе дистанционного обучения, однотипными элементами, упрощающими навигацию и позволяющими задержаться на курсе и быстрее включиться в процесс обучения;
- 3) привести текст в максимально читабельный вид;
- 4) использовать цветовые сочетания, позитивно влияющие на настрой к обучению;
- 5) наполнить курс наглядными примерами, инфографикой, диаграммами, таблицами и различными мультимедийными средствами обучения.

### Список литературы

1. *Дорожкин Е. М.* Психолого-педагогические проблемы использования электронного обучения / Е. М. Дорожкин, М. Д. Щербин // Научный диалог. 2016. № 5 (53). С. 199–213.
2. *Живой журнал* [Электронный ресурс]. Режим доступа: <http://arsenische.livejournal.com/51864.html>.
3. *Зрение человека* [Электронный ресурс]. Режим доступа: [https://ru.wikipedia.org/wiki/Зрение\\_человека](https://ru.wikipedia.org/wiki/Зрение_человека).
4. *Колмакова Л. А.* Совершенствование учебно-познавательной деятельности учащихся профессиональной образовательной организации на основе технологии визуализации учебной информации / Л. А. Колмакова // Образование и наука. 2015. № 6. С. 50–62.
5. *Кувандыкова Х. Б.* Модульное обучение – преимущества и недостатки / Х. Б. Кувандыкова // Молодой ученый. 2015. № 4. С. 579–581.
6. *Размещение информации на экране* [Электронный ресурс]. Режим доступа: <http://vgs1949.ru/?p=865>.
7. *Чувства из прошлого влияют на настоящее* [Электронный ресурс]. Режим доступа: <http://dervish-city.ru/article-feelings-from-the-past-influence-the-present.html>.

УДК 371.263:001.891.573

Н. С. Толстова, В. А. Сорокин, Н. В. Сердитов

N. S. Tolstova, V. A. Sorokin, N. V. Serditov

*ФГАОУ ВО «Российский государственный профессионально-педагогический университет», Екатеринбург*  
*Russian state vocational pedagogical university, Ekaterinburg*  
*natalya.tolstova@rsvpu.ru, vladimir.sorokin@rsvpu.ru*

### АВТОМАТИЗАЦИЯ ОЦЕНКИ КАЧЕСТВА ТЕСТА

### AUTOMATION OF TEST QUALITY ASSESSMENT

**Аннотация.** Рассматриваются характеристики теста, позволяющие судить о его качестве, и математическая модель, автоматизирующая расчет этих характеристик.

**Abstract.** The article discusses the characteristics of the test, which allows to judge its quality and a mathematical model that automates the calculation of these characteristics.

**Ключевые слова:** оценка качества, надежность теста, корреляция заданий друг с другом.

**Keywords:** quality assessment, reliability of the test, correlation of tasks with each other.

Одно из преимуществ, которым обладают тесты относительно других видов измерений, заключается в том, что они имеют основания для сравнения. На основе результатов прохождения теста возможно определить различные характеристики, которые позволяют судить о качестве теста, в частности, оценить его надежность. К таким характеристикам относятся:

- доля верных и неверных ответов на каждое задание, определяющая трудность задания;
- дисперсия баллов, характеризующая дифференцирующую способность задания;
- корреляция заданий друг с другом;
- средний коэффициент корреляции;
- точечный бисериальный коэффициент корреляции.

Надежность, определяющая воспроизводимость результатов тестирования и их точность, является одной из важнейших характеристик теста. Допустим, у нас есть гипотетическая группа испытуемых, которые немедленно забывают содержание теста по его завершении. Тогда, в случае надежного теста, повторяя тестирование многократно, мы должны получать одни и те же результаты. Для малонадежного теста результаты будут меняться каждый раз. Тест представляет собой систему заданий. Надежность теста зависит от качества заданий [2, 4].

В основном для определения надежности реальных тестов используют коэффициент корреляции Пирсона для результатов разных сеансов тестирования. Существует несколько способов вычисления надежности теста:

- 1) тестирование с помощью двух параллельных тестов;
- 2) повторное тестирование с помощью одного и того же теста через определенный промежуток времени;
- 3) расщепление результатов одного сеанса тестирования;
- 4) использование среднего коэффициента корреляции.

Для определения надежности теста на основе одного прохождения группой испытуемых удобно использовать метод, основанный на использовании среднего коэффициента корреляции всех заданий между собой:

$$r_t = \frac{M\bar{R}}{1 + (M-1)\bar{R}}, \quad \bar{R} = \frac{\bar{r}_{xy1} + \dots + \bar{r}_{xyM}}{M},$$

где  $r_t$  – коэффициент надежности теста;

$M$  – количество заданий в тесте;

$\bar{R}$  – средняя корреляция заданий друг с другом;

$\bar{r}_{xy}$  – среднее значение коэффициента корреляции для каждого задания.

Мы можем говорить о достаточной надежности теста, если его коэффициент надежности не меньше 0,7.

Корреляция между величинами означает, что они связаны. Определение корреляции как между заданиями, так и заданий с тестом в целом позволяет оценить систем-

ные качества теста. Благодаря такому анализу можно избавить тест от заданий, нарушающих его системные свойства.

Для вычисления корреляции заданий друг с другом используют формулу коэффициента корреляции Пирсона для дихотомических (бинарных) данных:

$$r_{xy} = \frac{p_{xy} - p_x p_y}{\sqrt{p_x q_x p_y q_y}},$$

где  $p_x$  – доля верных ответов на задание  $x$ ;  
 $q_x$  – доля неверных ответов на задание  $x$ ;  
 $p_{xy}$  – доля верных ответов одновременно на задания  $x$  и  $y$ .

Корреляция заданий друг с другом не должна быть слишком высокой ( $r_{xy} \leq 0,3$ ), иначе задания начинают дублировать друг друга. Если корреляция между двумя заданиями близка к единице, то одно из них лишнее.

Отрицательная корреляция заданий друг с другом заданиями нежелательна. Если задание отрицательно коррелирует с большим количеством других заданий, то это означает, что исход ответов на него противоположен результатам по другим заданиям. По всей вероятности, у такого задания либо имеются грубые ошибки в содержании или оформлении, либо проверяются знания из другой предметной области.

Точечный бисериальный коэффициент корреляции представляет собой коэффициент корреляции каждого задания с индивидуальным тестовым баллом испытуемого:

$$r_{pb} = \frac{\bar{X}_1 - \bar{X}_0}{s_x} \sqrt{\frac{n_1 n_0}{n(n-1)}},$$

где  $\bar{X}_1$  – средний индивидуальный балл испытуемых, справившихся с данным заданием;  
 $\bar{X}_0$  – средний индивидуальный балл испытуемых, не справившихся с данным заданием;  
 $n_0$  – число испытуемых, не выполнивших данное задание;  
 $n_1$  – число испытуемых, выполнивших данное задание;  
 $n$  – общее количество испытуемых;  
 $s_x$  – стандартное отклонение для индивидуальных баллов всех испытуемых.

Коэффициент  $r_{pb}$  характеризует валидность отдельных заданий. Необходимо стремиться к тому, чтобы данное значение было достаточно высоким. В. С. Аванесов рекомендует  $r_{pb}$  не ниже 0,5 [1].

Стандартное отклонение для индивидуальных баллов испытуемых рассчитывается по формуле

$$s_x = \sqrt{p_x q_x},$$

где  $p_x$  – доля верных ответов на задание  $x$ ;  
 $q_x$  – доля неверных ответов на задание  $x$ .

Тесты должны хорошо дифференцировать испытуемых, т. е. индивидуальные тестовые баллы должны в достаточной степени отличаться друг от друга. Для проверки данной характеристики теста используется величина дисперсии тестовых баллов:

$$s_x^2 = p_x q_x.$$

Малая величина дисперсии говорит о том, что тест плохо разделяет испытуемых по уровню знаний, не позволяет с приемлемой точностью ранжировать их. Слишком большая дисперсия указывает на сильную неоднородность группы испытуемых, на возможные нарушения процедуры тестирования, не достаточно ясные формулировки заданий и т. п. В случае оптимальной величины дисперсии распределение тестовых баллов близко к нормальному.

По мнению М. Б. Челышковой, дисперсию можно считать оптимальной, а распределение баллов близком к нормальному, если среднее арифметическое индивидуальных тестовых баллов примерно равно утроенному стандартному отклонению [3]:

$$\bar{X} \approx 3s_x.$$

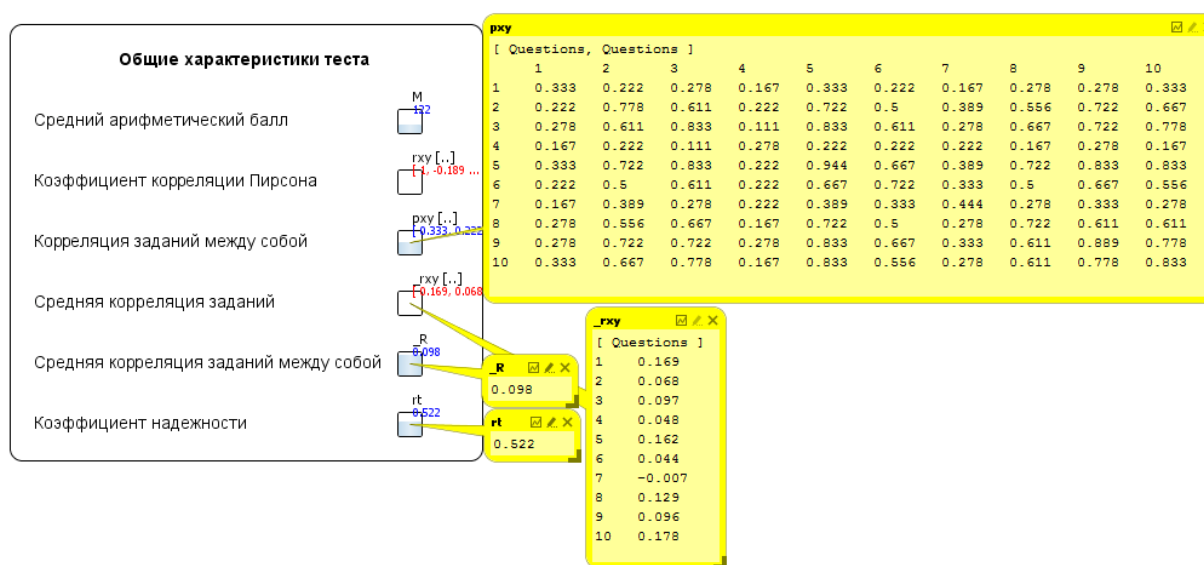
Для автоматизации расчета характеристик теста на основе двоичной матрицы результатов тестирования была создана модель в системе AnyLogic для автоматизации расчета характеристик теста на основе двоичной матрицы результатов тестирования. В качестве начальных данных заносятся количество вопросов в тесте, количество испытуемых прошедших тест и результаты ответов на вопросы в двоичной форме (1 – верно, 0 – неверно).

Все параметры модели разделены на четыре группы:

1. Начальные характеристики, заданные пользователем.
2. Характеристики для каждого задания, представленные в виде массивов с вычитанными числовыми значениями.
3. Характеристики для каждого студента, вычисленные на основе количества верных ответов.
4. Общие характеристики теста, содержащие матрицы значений корреляции и коэффициент надежности.

При запуске расчета подсчитывается:

- количество верных и неверных ответов по каждому заданию;
- доля верных и неверных ответов по каждому заданию;
- количество верных ответов, данных каждым испытуемым.



Результаты вычислений характеристик теста

Далее на основе этих данных по формулам, представленным в статье, происходит расчет следующих характеристик:

- дисперсия баллов для каждого задания;
- стандартное отклонение для каждого задания;
- коэффициент усвоения для каждого студента;
- точечный бисериальный коэффициент корреляции;
- корреляция заданий между собой;
- средний коэффициент корреляции;
- средняя корреляция заданий между собой;
- коэффициент надежности теста.

Все полученные данные хранятся в переменных и массивах, и их можно легко просмотреть. Это позволяет с помощью модели быстро оценить характеристики теста и сделать выводы о его качестве и качестве отдельных вопросов. Для примера внесем в модель данные о прохождении теста из 10 вопросов группой из 18 студентов. После произведения расчетов модель выдает характеристики теста, изображенные на рисунке. Коэффициент надежности в данном случае равен 0,522, что говорит о недостаточной надежности теста. Это может быть связано с малым количеством вопросов в тесте или с их низким качеством.

### Список литературы

1. *Аванесов В. С.* Применение тестовых форм в Rasch Measurement / В. С. Аванесов // Педагогические измерения. 2008. № 4. С. 3–20.
2. *Ким В. С.* Тестирование учебных достижений: монография / В. С. Ким. Уссурийск: Изд-во Уссурийского гос. пед. ин-та, 2007. 214 с.
3. *Челышкова М. Б.* Теория и практика конструирования педагогических тестов: учебное пособие / М. Б. Челышкова. Москва: Логос, 2002. 432 с.
4. *Хеннер Е. К.* Оценка прочности знаний на основе сопоставления результатов различных видов тестирования / Е. К. Хеннер, Т. С. Ознобихина // Образование и наука. 2012. № 1. С. 17–25.

УДК [378.016:004.92]:004.896

Т. А. Унсович, Е. В. Завьялова

T. A. Unsovich, E. V. Zavyalova

*ФГАОУ ВО «Российский государственный  
профессионально-педагогический университет», Екатеринбург  
Russian state vocational pedagogical university, Ekaterinburg  
tauns1@mail.ru*

### ВЫБОР АВТОМАТИЗИРОВАННЫХ СИСТЕМ ПРОЕКТИРОВАНИЯ ДЛЯ ГРАФИЧЕСКОЙ ПОДГОТОВКИ БАКАЛАВРОВ THE CHOICE OF AUTOMATED DESIGN SYSTEMS FOR THE GRAPHIC PREPARATION OF BACHELORS

**Аннотация.** Рассматриваются критерии выбора автоматизированных систем проектирования в преподавании графических дисциплин.

**Abstract.** The article discusses the selection criteria for automated systems design in the teaching of graphic disciplines.