

Карпинский Виктор Болеславович

**СОВРЕМЕННЫЕ МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ РЕЗУЛЬТАТОВ  
ПЕДАГОГИЧЕСКОГО ТЕСТИРОВАНИЯ**

Адрес статьи: [www.gramota.net/materials/1/2012/2/32.html](http://www.gramota.net/materials/1/2012/2/32.html)

Статья опубликована в авторской редакции и отражает точку зрения автора(ов) по рассматриваемому вопросу.

Источник

**Альманах современной науки и образования**

Тамбов: Грамота, 2012. № 2 (57). С. 82-84. ISSN 1993-5552.

Адрес журнала: [www.gramota.net/editions/1.html](http://www.gramota.net/editions/1.html)

Содержание данного номера журнала: [www.gramota.net/materials/1/2012/2/](http://www.gramota.net/materials/1/2012/2/)

**© Издательство "Грамота"**

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: [www.gramota.net](http://www.gramota.net)

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: [almanac@gramota.net](mailto:almanac@gramota.net)

11. Калюжный А. А. Теоретико-методологические предпосылки исследования имиджа как социально-психологического явления // Проблемы социальной психологии личности. 2005. № 4. С. 52-60.
12. Ковальчук А. С. Основы имиджологии и делового общения. 3-е изд. Ростов н/Д: Феникс, 2004. 256 с.
13. Кузин Ф. А. Имидж бизнесмена. М.: Ось-89, 1996.
14. Лосев А. Ф. Форма. Стиль. Выражение. М.: Мысль, 1995.
15. Панасюк А. Ю. Вам нужен имиджмейкер? Или о том как создавать свой имидж. М.: Юр-райт, 2008. 235 с.
16. Перелыгина Е. Б. Психология имиджа. М.: Логос, 2004. 211 с.
17. Почепцов Г. Г. Имиджмейкер. Киев: Наукова думка, 2005. 415 с.
18. Сергиенко С. К. Центр оценки: групповая деятельность, управленческая технология, исследовательская лаборатория. М.: Нута, 2004. 345 с.
19. Сидоренко Е. В. Методы математической обработки в психологии. СПб.: ООО «Речь», 2007. 350 с.
20. Спивак В. А. Организационная культура. СПб.: Издат. дом «Нева», 2004. 224 с.
21. Феофанов О. США: реклама и общество. М., 1974.
22. Шепель В. М. Имиджология: секреты личного обаяния. М.: Фильнь, 2007. 325 с.

УДК 519.2:37

*Виктор Болеславович Карпинский*

*Новгородский государственный университет им. Ярослава Мудрого*

#### СОВРЕМЕННЫЕ МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ РЕЗУЛЬТАТОВ ПЕДАГОГИЧЕСКОГО ТЕСТИРОВАНИЯ<sup>©</sup>

«Наука начинается с тех пор, как начинают измерять», - говорил Дмитрий Иванович Менделеев. В педагогике, как и во многих приложениях естественных, социальных, экономических и технических наук, зачастую необходимо измерять нечто, недоступное непосредственному наблюдению. В педагогике это, например, уровень подготовленности обучающихся в рамках тех или иных учебных дисциплин. Такие, недоступные непосредственному наблюдению, величины называют латентными. И измерены, точнее, оценены, они могут быть по совокупности каких-то косвенных данных. Обычно измеряемый объект тестируют, то есть подвергают серии специально выбранных испытаний, по отклику объекта на которые и можно сделать заключение о значении его латентной характеристики.

Одним из актуальных примеров этого является педагогическое тестирование: уровень подготовленности учащихся оценивается по результатам выполнения ими тестовых заданий.

Любое измерение должно давать объективный, надёжный и достоверный результат. Особенно важно это для измерения, основанного на косвенной информации, в том числе для педагогического тестирования.

Одной из актуальных задач, решаемых с использованием современных математических методов, является оценивание достоверности результатов тестирования. И такие методы последнее время активно разрабатываются, как в западной, так и в отечественной науке. Об этом и, в частности, о методах теории моделирования и параметризации тестов, основанных на математических моделях Георга Раша, можно прочитать в [1].

Математические модели Раша имеют вероятностную природу. Так, определяется вероятность того, что испытуемый, обладающий определённым уровнем подготовленности, при выполнении задания известной трудности получит некоторый балл. Уровень подготовленности каждого испытуемого, равно как и уровни трудности всех заданий теста, являются параметрами модели Раша и теми латентными характеристиками, которые подлежат измерению. Доступны для наблюдения баллы, фактически полученные каждым испытуемым за выполнение каждого задания. Эти баллы образуют матрицу ответов. Каждая строка матрицы ответов характеризует одного испытуемого и называется его профилем. Сумма баллов по строке называется первичным баллом испытуемого и является, в совокупности с другими аналогичными первичными баллами, достаточной статистикой для латентной характеристики - уровня подготовленности испытуемого.

Однако первичный балл испытуемого лежит на порядковой шкале. То есть испытуемых по результатам тестирования можно ранжировать, определить, что один испытуемый имеет более высокий уровень подготовленности, чем другой. Но нельзя ответить на вопрос: насколько более высокий, так как численно одинаковая разность первичных баллов в разных частях шкалы имеет объективно неодинаковое значение. Кроме того, результаты измерения, полученные с использованием разных тестов, нельзя непосредственно сравнивать друг с другом, так как они лежат на разных шкалах. Немного улучшает дело пересчёт первичных баллов в процентные, но и процентные баллы не образуют метрической шкалы, необходимой для объективности измерений. Тогда как шкалы, по которым измеряются физические величины (расстояния в метрах, интервалы времени в секундах и т.п.) - метрические. Хотелось бы и уровень подготовленности обучающихся измерять по результатам тестирования тоже баллами на метрической шкале.

То, что тестовые баллы, полученные указанными выше математическими методами, основанными на моделях Раша, лежат на метрической шкале, доказано [Там же, с. 39]. Что есть одно из важнейших преимуществ этих методов.

Другим важным преимуществом математических моделей Раша является объективность выполняемых с их помощью измерений. Это свойство моделей Раша позволило разработать специальные математические методы анализа результатов тестирования [2]. В частности, по согласию или несогласию данных тестирования с моделью оказывается возможным сделать заключение о достоверности или недостоверности результатов измерения уровня подготовленности испытуемого. Причинами недостоверности измерения могут быть нарушения процедуры тестирования, ошибки разработчиков теста и так далее. В любом случае педагоги, применяющие тестирование (а также иные лица, принимающие решения на основании тестовых измерений), должны быть уверены в достоверности полученных оценок. А при выявлении случаев недостоверности необходимо принять меры по совершенствованию тестовых материалов и (или) процедуры тестирования, либо устранению иных причин этой недостоверности. Математические методы, описываемые в настоящей статье, помогают в этом.

Поскольку модель вероятностная, то для оценки достоверности измерения используются методы математической статистики, и, в частности, методы проверки статистических гипотез. Стандартный подход предполагает сравнение значений некоторых статистик, вычисленных на эмпирических данных, с соответствующими критическими значениями. При проверке гипотезы, что эмпирические данные (профили испытуемых) согласуются с моделью Раша, используются так называемые статистики согласия. Их эффективность для обнаружения недостоверных измерений была подвергнута специальному исследованию [3]. Были выявлены недостатки упомянутых стандартных методов, определены причины их недостаточной эффективности и разработаны способы совершенствования этих методов применительно к указанной задаче.

В частности, было показано и теоретически, и эмпирически в сериях вычислительных экспериментов, что распределение статистик согласия существенно отличается от предполагаемых для них теоретических распределений. А, следовательно, неэффективно использовать при проверке гипотезы согласия предлагаемые стандартным подходом критические значения. То есть можно было предположить, что корень проблемы не в самих статистиках согласия, а в выборе критических значений для них. Соответственно, надлежало исследовать зависимость эффективности выявления искажённых профилей от выбора критического значения. Была определена зависимость от критического значения для частоты совершения ошибки первого рода (профиль испытуемого искажён, но критерием это не выявлено) и для частоты совершения ошибки второго рода (профиль не содержит искажений, но критерием выбран как искажённый).

Как обычно, в данном случае невозможно оптимизировать обе величины одновременно. Поэтому был использован известный из математической теории принятия решений подход: оптимизация функции суммарных потерь. В соответствии с ним разработан алгоритм получения критического значения для каждой статистики согласия посредством оптимизации функции потерь на специально сконструированных модельных данных, аналогичных реальным данным, подлежащим анализу. Интересной особенностью этого метода является отсутствие необходимости в знании теоретического характера распределения, что позволяет, - и это проверено, - использовать и иные критерии, не связанные со статистиками согласия, в том числе непараметрические.

Попутно исследование эффективности статистик согласия показало, что они эффективны в разных ситуациях по-разному. Одни показывают высокую эффективность при малом количестве искажений, но теряют её при увеличении этого количества. Другие, наоборот, менее эффективны в целом, но более устойчивы к увеличению числа искажённых профилей.

Поэтому было предложено сформировать на основе наиболее эффективных статистических критериев сводный критерий, являющийся их линейной композицией, коэффициенты которой выбираются так, чтобы статистика оказалась нормирована к единице. Тогда её семантика очевидна - это степень уверенности в том, что результаты тестирования данного испытуемого содержат искажения и недостоверны. Априорные соображения позволяли предполагать, что такой сводный критерий окажется эффективнее каждого из входящих в композицию статистического критерия отдельно. Посредством такого же, как и для всех остальных статистик, исследования его эффективности на всех модельных данных было показано, что это действительно так.

На основе метода оптимизации критических значений и алгоритма формирования сводного критерия была создана технология обработки данных тестирования, позволяющая оценивать достоверность его результатов по каждому испытуемому. Иначе говоря, разработано математическое и программное обеспечение оценки достоверности результатов тестирования [4].

Указанное программное обеспечение прошло всестороннюю проверку как в ходе серий вычислительных экспериментов, так и при обработке реальных данных тестирования, включая не только результаты педагогического тестирования, но и данные аналогичной структуры, но относящиеся к иным прикладным сферам. Результаты исследования прошли апробацию в Федеральном центре тестирования Министерства образования и науки Российской Федерации, были доложены на российских и международных симпозиумах и конференциях и опубликованы в научных журналах.

Целью публикации настоящей статьи являются трансляция наработанного опыта и приглашение коллег, заинтересованных в развитии и использовании методов обработки результатов тестирования, к сотрудничеству. Связаться с автором можно по электронной почте [V\\_B\\_Karpinsky@mail.ru](mailto:V_B_Karpinsky@mail.ru). Автор использует описанные выше методы в своей педагогической деятельности. Они оказываются полезны при разработке и

апробации тестовых материалов как текущего, так и промежуточного тестирования по разным учебным дисциплинам, а также тестирования остаточных знаний. Если по результатам апробации выявляются случаи недостоверного измерения уровня подготовленности некоторых испытуемых, то чаще всего причина в несовершенстве тестовых материалов (реже - процедуры тестирования), причём используемые математические методы помогают разобраться, в чём именно состоит это несовершенство, и как его исправить.

Перед нами пример того, как достижения современной науки могут найти конкретное применение в нашей системе образования. А именно, в частности, для совершенствования педагогического тестирования, вызывающего, как известно, много нареканий, но являющегося, на данный момент неизбежно, одним из важнейших средств педагогической диагностики.

#### Список литературы

1. Карданова Е. Ю. Моделирование и параметризация тестов: основы теории и приложения. М.: ФГУ «Федеральный центр тестирования», 2008. 296 с.
2. Карданова Е. Ю., Карпинский В. Б. Специальные методы анализа результатов тестирования, основанные на свойстве объективности моделей Раша // Информационные технологии. 2008. № 4. С. 72-80.
3. Карпинский В. Б. Исследование эффективности общих статистик согласия для обнаружения искажений при массовом тестировании // Вопросы тестирования в образовании. 2006. № 1 (17). С. 7-14.
4. Карпинский В. Б. Математическое и программное обеспечение оценки достоверности результатов массового тестирования: автореф. дисс. ... канд. техн. наук. Великий Новгород, 2009. 20 с.

УДК 376.42

*Людмила Сергеевна Круглова, Александр Павлович Коротыш  
Специальная (коррекционная) школа-интернат № 6 VIII вида, г. Астрахань*

### ТРУДОВОЕ ВОСПИТАНИЕ УЧАЩИХСЯ С ИНТЕЛЛЕКТУАЛЬНОЙ НЕДОСТАТОЧНОСТЬЮ<sup>©</sup>

Трудовое обучение и трудовое воспитание в специальной (коррекционной) школе VIII вида представляет собой единую систему. Главная задача трудового обучения состоит в том, чтобы максимально облегчить детям их адаптацию в обществе после окончания школы. Особенно актуально это в последнее время, когда медики и официальная статистика констатируют факты увеличения числа рождаемости детей с отклонениями в развитии, ограниченными возможностями, детей-инвалидов. Естественно, что такие дети испытывают значительные трудности в своей жизни, в приобретении необходимых трудовых навыков, выборе и освоении своего пути в жизни.

Учащиеся специальной (коррекционной) школы с самого начала пребывания в школе нуждаются в постоянном и последовательном обогащении своего мировосприятия, мироощущения, социального опыта и, что особенно актуально, - в поэтапном приобщении к осознанной трудовой деятельности.

Общеизвестно, что уроки труда в коррекционной школе выступают не только как предмет, но и как одно из важнейших и неотъемлемых средств коррекции умственных и физических недостатков детей.

Используя все познавательные возможности детей, нужно развивать у них жизненно необходимые навыки, чтобы, став взрослыми, они могли самостоятельно себя обслуживать, выполнять в быту и на производстве простую работу, жить по возможности в семье и в трудовом коллективе.

Достижение поставленных целей обеспечивается решением следующих основных задач в коррекционной школе:

- 1) развитие всех психических функций и познавательной деятельности детей в процессе обучения и коррекция их недостатков; основное внимание в этой работе должно быть направлено на умственное развитие;
- 2) воспитание детей с ОВЗ (ограниченными возможностями здоровья), формирование у них правильного поведения; основное внимание в этом разделе работы направлено на нравственное воспитание;
- 3) трудовое обучение и подготовка к посильным видам труда, физическое воспитание, самообслуживание;
- 4) бытовая ориентировка и социальная адаптация - как итог всей работы.

Большое значение в определении трудовых возможностей детей с ОВЗ и профиля доступного им вида труда имеет их физическое развитие, развитие общей и ручной моторики. Правильно организованное трудовое обучение благоприятно влияет на развитие учащихся, содействуя коррекции их недостатков.

В работах Б. И. Пинского, Г. И. Дульнева, Е. А. Ковалева, С. Л. Мирского, исследовавших вопросы трудового воспитания детей с недоразвитием интеллекта, отмечается, что педагогически правильно организованное трудовое обучение и воспитание является эффективным средством коррекции недостатков развития умственной деятельности учащихся. Наибольшая сложность трудового обучения в специальной (коррекционной) школе VIII вида состоит в формировании обобщенных умений, способности выполнять трудовые задания не только в определенной ситуации, но и при изменении условий.