

О ПРОДОЛЖИТЕЛЬНОСТИ ТЕСТИРОВАНИЯ

Теория

Содержание

Олег Деменчёнок

Восточно-Сибирский институт МВД России

AskSystem@yandex.ru

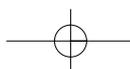
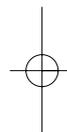
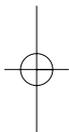
Рассмотрены подходы к определению продолжительности тестирования. Показано, что время тестирования — не единственный фактор, значимо влияющий на дисперсию результатов. Сделан вывод о проблематичности практического нахождения времени тестирования по максимуму дисперсии тестовых результатов ввиду невозможности выявить в «чистом виде» влияние продолжительности тестирования. Рассмотрено влияние свойств нервной системы обучаемого на время тестирования. Предложен способ коррекции продолжительности тестирования для обучаемых с инертным типом личности.

Ключевые слова: тест, продолжительность тестирования, дисперсия результатов тестирования, сила нервной системы, лабильность.

Во многих вузах отмечается повышение интереса к компьютерному тестированию. Это связано с интенсификацией процесса обучения, что требует внедрения новых информационных технологий и адекватного мониторинга подготовленности обучаемых.

Достоинства тестирования достаточно известны: при правильном применении тесты позволяют достичь высокой точности и обоснованности результатов контроля, исключают влияние субъективных факторов, технологичны, стимулируют студентов к обучению. Значительно меньше изучены недостатки автоматизированного контроля, к числу которых следует отнести:

- необходимость формализации заданий. Легко формализуются задания по проверке знания фактов, типовые задачи (они встречаются чаще всего). Задачи повышенной сложности хотя и поддаются формализации, но используются реже ввиду неоднозначности решения и больших затрат времени при тестировании. Нестандартные, творческие задачи почти не поддаются формализации, поскольку их решение нельзя оценивать формализовано.



Поэтому необходимость формализации заданий несколько ограничивает возможности и область применения тестового метода;

- сложность учёта индивидуальных особенностей учащихся. На результат тестирования влияет способность испытуемого к саморегуляции в стрессовых ситуациях. Лучшими могут оказаться менее подготовленные студенты с высокой устойчивостью в стрессовых ситуациях, а уровень психической активности во многом определяет скорость выполнения теста. При прочих равных условиях холерик закончит тестирование раньше флегматика или меланхолика. Указанные факторы способны внести искажения в результат контроля знаний;

- сложность принятия обобщающих решений. Если, например, по результатам тестирования нужно выдать полные и обоснованные рекомендации по устранению выявленных пробелов в знаниях, то опытный педагог сделает это лучше любой системы компьютерного тестирования.

Как правило, тестирование ограничено по времени, что облегчает планирование, создаёт определённую психологическую нагрузку на обучаемых, а также затрудняет получение несанкционированной помощи — от банального списыва-

ния и подсказок однокурсников до использования возможностей компьютерной техники. Так, В.С. Аванесов полагает, что время проверки знаний не должно превышать 40 минут, поскольку в противном случае теряется одно из самых существенных свойств теста — кратковременность процесса контроля знаний¹.

Обычно ограничивается общее время выполнения теста. Известны также рекомендации ограничивать время выполнения каждого тестового задания². Испытуемому, не уложившемуся в заданное время, либо ставится неудовлетворительная оценка, либо засчитываются неправильные ответы.

Выставление неудовлетворительных оценок оправдано в ситуациях, когда скорость решения является профессионально значимым фактором, например, при проверке знаний о порядке действий в экстренных ситуациях. В некоторых компьютерных программах предусмотрено дифференцирование тестового балла в зависимости от длительности ответа³.

По истечении основного времени тестирования может быть предоставлено дополнительное время, однако тестовый балл в этом случае снижается либо на фиксированную величину, либо пропорционально перерасходу времени.

1

Аванесов В.С.
Item Response Theory:
Основные понятия и по-
ложения. // ПИ. 2007.
№ 2. С. 3–28.

2

*Башмаков А.И.,
Башмаков И.А.*
Разработка компьютер-
ных учебников и обуча-
ющих систем. М.:
ФИЛИНЪ, 2003. 616 с.

3

Войтов А.Г.
Учебное тестирование
для гуманитарных и эко-
номических специаль-
ностей. М.: Дашков и К,
2006; 402 с.

Например, базовое время тестирования — 30 минут, при необходимости дополнительно выделяется ещё 5 минут, но тестовый балл снижается на 15%.

Подходы к определению продолжительности тестирования

Сразу отметим, что нет общепризнанной методики определения продолжительности тестирования. Существует несколько основных подходов к решению этой задачи. Так, время тестирования может определяться на основе:

1) времени выполнения теста преподавателем (испытуемым выделяют в 3–5 раз больше);

2) усреднённых эмпирических данных выполнения отдельного задания из расчёта 0,5–2 минуты на одно задание. Среднее время ответа испытуемого на тестовое задание не должно превышать 5 минут⁴;

3) данных пробного тестирования.

Данные пробного тестирования могут использоваться по-разному. Некоторые методики предусматривают очень жёсткое ограничение времени — продолжительность тестирования принимается такой, чтобы только 5% испытуемых могли пройти тест в полном объёме. Известнейший тест «на время» — американский

TOEFL. Ограничения во времени выполнения там вполне обоснованы⁵. Известны скоростные тесты для водителей. Но то, что хорошо для отбора пилотов и водителей, не всегда подходит для контроля знаний.

Если скорость выполнения теста не имеет первостепенного значения — а это справедливо почти для всех учебных дисциплин — то жесткие временные рамки становятся препятствием для объективного измерения уровня подготовленности. В этом случае, как правило, принимается время, достаточное для полного прохождения теста абсолютным большинством испытуемых.

Кроме того, существуют методы, основанные на анализе дисперсии результатов тестирования, полученных при разных временных ограничениях. Оптимальной признаётся продолжительность тестирования, при которой дисперсия (вариация значений) результатов тестирования начинает увеличиваться⁶ или достигает максимума⁷.

Поиск максимума дисперсии широко используется в технических системах автоматического регулирования. Например, автоматическая фокусировка в фотоаппаратах основана на определении контраста изображения. Неверная наводка объектива на резкость приводит к «размытости» изобра-

Теория

4

*Васильев В.И.,
Киришок А.А.,
Тягунова Т.Н.*

Культура компьютерного тестирования. Требования к программно-дидактическим тестовым материалам и технологиям компьютерного тестирования. М.: МГУП, 2005. 29 с.

5

Морев И.А.

Образовательные информационные технологии. Часть 2. Педагогические измерения: Учебное пособие. Владивосток: Изд-во Дальневост. ун-та, 2004. 174 с.

6

Майоров А.Н.

Теория и практика создания тестов для системы образования (Как выбирать, создавать и использовать тесты для целей образования). М.: Интеллект-центр, 2001. 296 с.

7

Ким В.

Выбор оптимального времени тестирования. //ПИ. 2008. № 2, С. 43–50.

ПЕД
измерения

жения и, следовательно, к низкой контрастности. Электроника камеры изменяет наводку объектива на резкость до тех пор, пока не будет найден максимум контраста. Максимальный контраст изображения (т.е. максимальная дисперсия яркости точек) соответствует наилучшей резкости (чёткости линий, различимости отдельных деталей и т.д.).

К сожалению, такой подход представляется не совсем оправданным для практического применения в тестировании. Для подбора оптимальной продолжительности тестирования нужно сравнивать дисперсии (т.е. показатели вариации баллов) нескольких групп испытуемых при разном времени тестирования. Но время тестирования — не единственный фактор, значимо влияющий на дисперсию результатов.

Группы испытуемых могут — и, скорее всего, будут — разными по качественному составу студентов. Следовательно, изменение дисперсии может быть обусловлено как изменением времени тестирования, так и изменением качественного состава испытуемых.

Могут оказывать влияние и другие факторы. Например, тестирование в одной группе проводится утром, во второй — в перерыве между занятиями, а в третьей — после напряженного учебного дня. Как в таких

условиях выявить в «чистом виде» влияние продолжительности тестирования — неизвестно.

Поскольку нельзя однозначно назвать причину, по которой вариация результатов тестирования в одной группе больше, чем в другой, то простое сравнение дисперсий не даёт достаточных оснований для определения оптимальной продолжительности тестирования. Конечно, можно несколько раз протестировать одну и ту же группу. Но это также не устраняет проблему обеспечения сопоставимости дисперсий: во второй (третьей и т.д.) раз решать одни и те же задания гораздо легче.

Влияние свойств нервной системы испытуемого

Жёсткая регламентация продолжительности тестирования ставит в заведомо невыгодные условия испытуемых с инертным типом личности. Таким студентам объективно требуется больше времени на решение теста. Причиной того, что конкретный испытуемый не уложился в заданное время, может быть как незнание ответов, так и низкий уровень психической активности (т.е. студент может справиться с тестом при менее жёстких временных ограничениях). Если в последнем слу-

чае время решения не имеет принципиального значения, то результаты контроля знаний нельзя считать достоверными. Они не соответствуют фактическому уровню подготовленности студента, оценка получается искусственно заниженной.

Интерес представляет оценка взаимосвязи успешности выполнения тестовых заданий и времени решения при различных свойствах нервной системы обучаемого. Знание этих зависимостей даёт возможность правильно распределить временные ресурсы учебной дисциплины или занятия, добиваясь при этом максимальной достоверности контроля знаний.

К числу индивидуально-типологических свойств нервной системы студента, влияющих на продолжительность тестирования, следует отнести её силу и лабильность^{8,9}. Сила нервной системы характеризуется выносливостью, работоспособностью обучаемого и уровнем абсолютной чувствительности, а лабильность — скоростью возникновения и прекращения нервных процессов. Индивидуально-типологические свойства нервной системы определяют индивидуальный стиль деятельности студента.

Исследованиями установлено, что эти процессы генетически обусловлены. Следова-

тельно, указанные свойства стабильны, а сделанные по ним выводы будут справедливы для данного обучаемого в течение продолжительного промежутка времени.

Время контроля в значительной мере определяется лабильностью; влияние силы нервной системы незначительно, хотя и значимо¹⁰. Результаты решения тестовых заданий почти не зависят от силы и лабильности нервной системы обучаемого. Эти данные подтверждаются при практическом использовании автоматизированного контроля знаний.

Действительно, результат тестирования в основном определяется знанием учебного материала, а время прохождения теста — индивидуальными свойствами нервной системы обучаемого. То есть время прохождения теста конкретным студентом мало изменяется; взаимосвязь этого времени с уровнем подготовки к занятию, а также с результатом тестирования практически отсутствует.

Для достижения одинаковых по эффективности результатов контроля обучаемому с инертным типом личности требуется больше времени на решение тестовых заданий, чем для получения того же результата обучаемому с лабильным типом личности. Возможен обратный вывод о том, что если при одинаковом уровне знаний

Теория

Знание

8

*Свиридов А.П.,
Вербицкий А.Л.* Психолого-педагогические особенности контекстного обучения. М.: Знание, 1987. 109 с.

9

*Заркин В.Г.,
Фадеева Н.В.,
Свиридов А.Б.* Прочность усвоения знаний и психофизиологические особенности учащихся. /Новые исследования в психологии, 1983, № 2. С. 41–45.

10

Там же: *Свиридов А.П.,
Вербицкий А.Л.* Психолого-педагогические особенности контекстного обучения. М.: Знание, 1987. 109 с.

ПЕД
измерения

один обучаемый закончил выполнение контрольного задания раньше, чем другой, то можно предполагать, что индивидуально-типологические свойства нервных систем обучаемых различны: второй обучаемый, по-видимому, обладает более инертным и слабым типом нервной системы.

При стандартном подходе к организации тестирования ограничение по времени устанавливается равным для всех. При этом получается, что не укладываются в заданное время одни и те же обучаемые. Далеко не всегда эти обучаемые — отстающие. Среди них есть и те, кто усвоил учебный материал, но не в состоянии из-за свойств нервной системы выполнить тест в отведенное время. Глядя на таких студентов, почти физически ощущаешь, с каким напряжением «проворачиваются шестерёнки» в их мозгах.

Снижать тестовый балл за перерасход лимита времени в таких случаях нецелесообразно, так как это нарушает объективность педагогического измерения. Кроме того, необъективные результаты контроля знаний создают у обучаемого отрицательную стимуляцию обучения, что затрудняет дальнейшую работу преподавателя с таким студентом.

На взгляд автора, продолжительность тестирования для обучаемых, систематически не

укладывающихся в отводимое время, в дальнейшем должна быть увеличена. Практика показывает, что количество таких студентов не превышает 5% (т.е. 1–2 человека на группу). Обычно для обучаемых с инертным типом личности бывает достаточным увеличение времени на 50% или адекватное снижение числа тестовых заданий при неизменном времени (т.е. на 33%). Такая простейшая и весьма доступная для преподавателя корректировка позволяет правильно распределить временные ресурсы занятия (время занятия, отводимое на контроль знаний, практически не изменяется), добываясь при этом высокой достоверности контроля знаний всех обучаемых.

При систематическом использовании автоматизированных систем контроля знаний несложно организовать подобную коррекцию продолжительности тестирования в зависимости от индивидуальных психологических характеристик обучаемых программными средствами.

Таким образом, учёт индивидуальных свойств нервной системы обучаемого может служить необходимым условием достоверности результатов контроля знаний и «разумной» интенсификации процесса обучения с помощью автоматизированных систем.