

- международная финансовая и техническая помощь;
- займы международных и зарубежных кредитно-финансовых учреждений;
- средства отечественных и иностранных инвесторов, в т. ч. средства национальных и транснациональных финансово-промышленных групп;
- капитализация части прибыли и амортизационные отчисления субъектов хозяйственной деятельности.

К сожалению сегодня можно говорить только о тенденциях, существующих в экономике и в обществе в целом, об осознании необходимости ресурсосбережения, чем о реальных экономических результатах. Меры по активации источников энергосбережения должны приниматься как на государственном уровне, так и на уровне региональных и районных властей.

АНАЛИЗ ВЛИЯНИЯ СЛУЧАЙНЫХ ОТВЕТОВ НА РЕЗУЛЬТАТЫ ПРОВЕДЕНИЯ ТЕСТИРОВАНИЯ СТУДЕНТОВ

А.И. Шемаров, к.т.н., доцент

Белорусский национальный технический университет (г. Минск),

Е.Г. Гриневич, ст. преподаватель,

Белорусский государственный аграрный технический университет (г. Минск)

Не вызывает сомнений, что при проведении различных вариантов тестирования, даже хорошо подготовленные студенты иногда угадывают правильный вариант ответа. Как показывает практика, особенно для дисциплин гуманитарного и естественнонаучного циклов, вопросы с множественным выбором при формировании тестов составляют подавляющее большинство. Поэтому целью данной работы является изучение того, какое влияние оказывает угадывание правильного ответа тестируемым на его результирующую оценку.

Эксперимент проводился на базе системы дистанционного обучения МООДУС в Белорусском государственном аграрном техническом университете (БГАТУ). Для исследования использовались результаты тестирования студентов дневной формы обучения факультета предпринимательства и управления во время зимней экзаменационной сессии 2008/09 учебного года по дисциплине «Операционные системы».

База вопросов для контроля знаний по этой дисциплине включала 185 вопросов (основная база вопросов) по проверке знаний учебного материала (теоретическая, практическая составляющая курса и материал управляемой самостоятельной работы студентов) и 10 вопросов-«пустышек» (дополнительная база вопросов), имеющих псевдонаучный смысл, не имеющих правильного ответа или смысла априори. Обе базы вопросов включали в себя различные их типы: в закрытой форме (множественный выбор); на соответствие и «верно» и «неверно».

Каждое тестовое задание состояло из 50 вопросов, выбранных случайным образом: 48 вопросов — из основной базы и 2 вопроса из дополнительной. Вопросы и варианты ответов не только выбирались случайно, но и перемешивались при формировании тестовой последовательности. Время прохождения теста было ограничено 32 минутами. Такой период времени прохождения теста определяется опубликованными данными, по которым при ответе на вопрос, не требующий использования расчетов и справочной литературы, должно отводиться от 30 до 45 секунд. Две дополнительные минуты даются на возможные задержки при загрузке теста на сетевой терминал и отправке выполненного теста на сервер для обработки системой.

Ответы на вопросы из основной базы оценивались, исходя из одинаковой максимальной оценки равной единице за верный ответ. За неправильный ответ и за полное отсутствие ответа на вопрос оценка равна нулю. Частично правильный ответ позволяет получить оценку в интервале значений от нуля до единицы. Ответы на вопросы из дополнительной базы не оценивались и не вносили вклад в формирование общей оценки за тест. До проведения экзаменационных испытаний тестируемые были предупреждены о том, что ряд заданий, возможно, не будет иметь правильного ответа на поставленный вопрос.

По завершении теста, тестируемый получал оценку, нормированную по десятибалльной шкале. Оценка пропорциональна общей доле правильных ответов по всем вопросам теста. После окончания тестирования студенты могли сравнивать собственные ответы с базовыми, а

также проанализировать общую оценку за тест и оценку за каждый вопрос, что является общепринятой практикой при проведении экзаменационных испытаний в университете. Всего в зимнюю сессию 2008/09 учебного года было протестировано 149 студентов. Результаты указанного тестирования являются исходными данными для проводимых в статье исследований.

Основными предпосылками для проведения эксперимента стали следующие допущения.

При недостаточном усвоении материала по дисциплине тестируемые на базе общих знаний могут путем удаления заведомо неверных ответов сократить количество правильных ответов, из которых осуществляется выбор. Таким образом, существенно повышается вероятность правильного ответа. Если количество правильных ответов сокращается до двух или даже одного, то исходя из вероятностных оценок, тестируемый будет получать положительную оценку, выбирая ответы наугад, не имея минимального уровня знаний по дисциплине.

Цель работы оценить, насколько часто студенты прибегают к угадыванию при выборе правильного ответа и выработка рекомендаций по устранению этой ситуации при тестировании в СДО Моодус.

Примеры вопросов-«пустышек» представлены ниже:

Вопрос **E01** — «Какая часть ядра ОС В-31 была включена в ядро ОС Windows XP?». ОС В-31 в природе не существует, но прослеживается связь с маркировкой вооружения, используемого в армии США и, следовательно, находящаяся «на слуху» благодаря средствам массовой информации. В качестве ответов на вопрос были предложены три варианта ответа:

- 1). 10%. Утвердительно ответил 1 тестируемый из 29 (3%).
- 2). 25%. Утвердительно ответило 19 тестируемых из 29 (66%).
- 3). 50%. Утвердительно ответило 5 тестируемых из 29 (17%).

Не дали ответ на вопрос 4 тестируемых студента из 29 (14%), что в первом приближении можно считать правильным ответом, если вопрос вообще рассматривался тестируемым;

Вопрос **E03** — «Эффективность ОС измеряется в – ». Не существует общеупотребительных методик сравнения различных ОС, не говоря уже об единицах измерения, но возможно в вводной лекции были упомянуты фамилии «отцов-основателей» вычислительных машин и кибернетики. В качестве ответов на вопрос были предложены три ответа:

- 1). В Винерах. Утвердительно ответило 3 тестируемых из 31 (10%).
- 2). В Тьюрингах. Утвердительно ответило 4 тестируемых из 31 (13%).
- 3). В Гилбертах. Утвердительно ответило 20 тестируемых из 31 (65%) — возможно, слышали о смесях задач, на которых проверяется производительность вычислительных систем, комплексов и сетей.

Не дали ответ на вопрос 4 тестируемых студента из 31 (12%);

Вопрос **E07** — «Плотность загрузки операционной системы определяется в – ». Здесь вообще представлена благоглупость, но в таких похожих на общеупотребительные величины типа $кг/мм^2$ и т.д. В качестве ответов на вопрос были предложены три ответа:

- 1). В байтах на $мм^2$ кристалла процессора. Утвердительно ответили 9 тестируемых из 28 (32%).
- 2). В пикселях на байт. Утвердительно ответило 11 тестируемых из 28 (39%).
- 3). В командах на $мм^2$ кристалла процессора. Утвердительно ответило 3 тестируемых из 28 (11%).

Не дали ответ на вопрос 5 тестируемых студента из 28 (18%).

Анализ ответов на контрольные вопросы показал, что студенты, даже хорошо успевающие, пытаются угадать правильный ответ. Ответ на бессмысленный вопрос в среднем давали около 77% тестируемых студентов. Эти действия усугубляются тем, что перед тестированием они были предупреждены о возможном введении вопросов, не имеющих правильных ответов. Это подтверждает наши предпосылки о возможных проблемах качественного проведения тестирования с целью выявления реального уровня знаний испытуемых.

В таблице представлены результаты тестирования, отсортированные по группам. Всего выделено 15 групп. Основным признаком сортировки является классификация по результатам ответа на «вопросы-пустышки». В первую группу вошли тестируемые студенты, ответившие на оба «вопроса-пустышки», во вторую группу — ответившие на один вопрос, и в третью — не отвечавшие на один из «вопросов-пустышек». Каждая основная группа поделена на пять подгрупп по результатам ответов на основные вопросы. В первую подгруппу вошли студенты, ответившие на все основные вопросы, во вторую, третью, четвертую и пятую подгруппы — студенты, не ответившие на один или два основных вопроса, три или пять, шесть или десять, более десяти вопросов соответственно. Для каждой подгруппы определяется число студентов попавших в нее, их средняя оценка по результатам тестирования и среднеквадратическое отклонение оценки по подгруппе. Также представляется эти же оценки по каждой группе и для всей вы-

борки в целом. Вычисления проводились по каноническим методикам.

Таблица — Распределение результатов тестирования по группам и подгруппам тестируемых студентов

Группа	Параметр	Подгруппы по количеству ответов на основные вопросы					По всей группе
		48	47..46	45..43	42..38	37..25	
Ответившие на два «вопроса-пустышки»	Число студентов	64	22	6	6	2	100
	Средняя оценка	6,1	5,8	5,3	5,9	4,3	5,9
	Отклонение оценки (ско)	1,14	1,07	0,68	0,50	0,26	1,11
Ответившие на один «вопрос-пустышку»	Число студентов	11	16	5	2	4	38
	Средняя оценка	5,9	6,1	5,9	6,0	4,8	5,9
	Отклонение оценки (ско)	1,36	0,98	0,36	0,79	0,72	1,09
Не отвечавшие на «вопросы-пустышки»	Число студентов	3	3	0	2	1	9
	Средняя оценка	5,7	6,3	0,0	5,7	4,6	5,8
	Отклонение оценки (ско)	1,4	0,6	0,0	0,4	0,0	1,1
По всем группам	Число студентов	78	41	11	10	7	147
	Средняя оценка	6,06	5,96	5,57	5,87	4,64	5,92
	Отклонение оценки (ско)	1,19	1,03	0,64	0,57	0,61	1,10

Анализ распределения оценок по группам показывает, что статистические характеристики полученных средних оценок по группам практически не отличаются, а это означает, что существует очень неприятная возможность существенного нивелирования оценок действительно работающих в семестре студентов и тех, кто возможно вообще не готовится к экзаменам, и использует тактику угадывания ответов на вопросы.

Чтобы убедиться в этом, попробуем найти работы студентов, которые даже не маскируют свою тактику получения положительной оценки по экзамену без какой-либо подготовки. Критерием выявления таких студентов может служить аномально низкое время прохождения теста. По выборке всех ответов, этому критерию соответствуют работы двух студентов. При общей тенденции использовать все отведенное для тестирования время, данные работы были выполнены за 6 минут 30 секунд оценка 5,31 и 7 минут 59 секунд оценка 5,23 соответственно. При этом были даны ответы на все пятьдесят вопросов. Среднее время ответа на вопрос соответственно 7,8 и 9,6 секунды. Следующие, по затраченному на ответы времени, две работы, скорее всего также соответствуют выбранному критерию. Работы были выполнены за 19 минут 49 секунд оценка 5,371 и 22 минуты 58 секунд оценка 5,07 соответственно. При этом также были даны ответы на все пятьдесят вопросов, включая и бессмысленные «вопросы-пустышки».

Таким образом, существующая система проведения тестирования существенно нивелирует знания студентов, так как позволяет получать положительные оценки без сколько то ни было существенной подготовки к экзамену. Необходимо менять систему оценивания знаний. Необходимо внедрять алгоритмы, активно выявляющие попытки угадывания правильных ответов. В качестве промежуточной меры можно использовать в каждом вопросе добавочный ответ — «нет правильного ответа», который должен быть истинным для определенного количества вопросов, и позволит снизить вероятность угадывания правильного ответа.