

Теория

ДИСТРАКТОРНЫЙ АНАЛИЗ

Вадим Аванесов

testolog@mail.ru

Рассматриваются вопросы оценки качества правдоподобных, но неправильных ответов к тестовым заданиям. Это дистракторы, от которых зависит качество отдельного задания и теста в целом. Хороший дистрактор часто выглядит подходящим ответом для тех, кто не знает правильного ответа на задание теста. Чем больше некачественных дистракторов в заданиях, тем хуже тест.

Рассмотрены два метода оценки качества дистракторов — процентный анализ, используемый в статистической (классической) теории педагогических и психологических измерений и метод, используемый в математической теории измерений (IRT).

Ключевые слова: тестовое задание, тестирование, выбор правильного ответа, дистрактор, вероятность угадывания.

«Мир постоянно изменяется и подает нам сигналы.

Наша задача — вовремя их улавливать, понимать и адекватно реагировать на них, то есть измениться.

Причём — и это важно — преобразования должны охватывать как саму реальность, так и её восприятие».

Люк Де Бранддер

Введение

В данной статье затрагивается самая консервативная проблема тестирования. Это проблема изменения формы заданий. Вот уже бо-

лее ста лет в практике применяются задания с выбором одного правильного ответа. Нет силы, которая побудила бы переход к заданиям других форм. Между тем, эти задания уже изрядно устарели. В них довольно высока вероятность угадывания правильного ответа. Эти задания не случайно являются главным предметом критики со стороны тех педагогов, которые противостоят тестовым методам.

Определения

Дистрактором (от англ. to distract — *отвлекать*) называется неправильный, но правдоподобный ответ, в заданиях с выбором одного или нескольких правильных ответов. Таким образом, дистрактор используется в процессе тестирования для отвлечения внимания от правильного ответа тех испытуемых, которые либо совсем не знают правильный ответ, либо, имея весьма туманные представления о предмете тестирования, пытаются угадать его во время тестирования.

Сделать неправильные ответы правдоподобными для незнющих студентов — одна из важных задач разработчика теста.

Психологический смысл применения дистрактора заключается в создании интерферентной среды тестирования, развивающей рефлексивное

мышление. В такой среде один ответ кажется привлекательнее другого: приходится искать аргументы в пользу того или иного ответа. В итоге появляется привычка аргументировать выбор ответа. Известно, что аргументированные знания всегда ценнее и прочнее знаний неаргументированных.

Дистракторным анализом называется теоретико-эмпирическое исследование правильности и эффективности подбора всех ответов, прилагаемых к каждому заданию теста. Без такого анализа ни тестов, ни педагогических измерений не бывает. Теоретическая часть анализа касается вопросов содержания и формы заданий с ответами. Эмпирическая часть касается организации пробного тестирования с целью определения свойств заданий и свойств ответов.

Невольно напрашивается вопрос: а в российском ЕГЭ проводят дистракторный анализ? Мы не знаем. В отчётах ФИПИ нет никакой информации по данному вопросу. Если такой анализ не проводится, и результаты не публикуются, то это и есть самая главная причина негодности той продукции ФИПИ, в которой применяются задания с выбором одного правильного ответа. Это одна из причин, почему КИМы ЕГЭ не являются методом педагогических измерений¹.

1
Аванесов В.С.
Являются ли КИМы
ЕГЭ методом педагогических измерений?
(Вторая редакция).
Первая редакции — 5
апреля 2009 г. Вторая
редакция — 25 мая
2009 г. <http://viperson.ru/wind.php?ID=563869&soch=1>

Пример задания с выбором одного правильного ответа

Рассмотрим пример задания с выбором одного правильного ответа из пяти ответов, предлагаемых испытуемому. Задание начинается с инструкции.

Нажимайте на клавишу с номером правильного ответа:

1. КАРЛ МАРКС РОДИЛСЯ В

- 1) Берлине
- 2) Карлмаркшгадте
- 3) Трире
- 4) Мюнхене
- 5) Лондоне

Из инструкции становится понятным, что в предлагаемом задании есть только один правильный ответ. Для тех, кто знает, это город Трир. Остальные четыре ответа — неправильные, но, по убеждению разработчика, достаточно правдоподобные (привлекательные) для тех испытуемых, кто не знает места рождения рассматриваемой персоны.

Убеждения каждого разработчика могут основываться на разных основаниях. Для тех, кто не знает, но должен был знать место рождения, *Берлин* может быть приемлемым дистрактором, поскольку это столица страны. Многие полагают, что известные люди нередко

родом из столиц. *Карлмаркштадт* привлекает своим названием, данным, по их мнению, возможно в честь рассматриваемой личности. *Мюнхен* — это город, в котором случилось многое, так что нетрудно допустить, что это и есть тот город, в котором родился К.Маркс. И наконец, основанием для выбора *Лондона* в качестве правильного ответа на задание может стать ассоциация с Энгельсом, который жил в Англии.

Надо подчеркнуть, что все убеждения такого рода абсолютно субъективны. А потому они могут подтверждаться, а могут и не подтверждаться в практике тестирования. Отсюда берёт начало понимание необходимости предварительной, до начала тестирования, проверки приемлемости каждого дистрактора в данной или похожей группе испытуемых.

Алогизмы

Самый частый случай ошибочно сделанных дистракторов — это использование алогизмов.

Посмотрим несколько примеров:

2. ВОССТАНИЕ СПАРТАКА БЫЛО

- а) в 73 г. до н.э.
- б) в 73 г. н.э.

- в) в 1973 г.
г) пока ещё не было²

3. ПРИЕХАВ В СТОЛИЦУ

- 1) всех поражает шум и суета;
- 2) у вас не будет времени сходить в музей;
- 3) идите сразу на Красную площадь;
- 4) у меня было прекрасное настроение³.

4. ШЁЛК ХРАНИТСЯ В СПИРТЕ, КОТОРЫЙ МЕНЯЕТСЯ КАЖДЫЕ

- 1) 2 дня
- 2) 5 дней
- 3) 10 дней
- 4) 20 дней

Это задание лучше сформулировать так:

5. СПИРТ, В КОТОРОМ ХРАНИТСЯ ШЁЛК, МЕНЯЕТСЯ КАЖДЫЕ

- 1) 2 дня
- 2) 5 дней
- 3) 10 дней
- 4) 20 дней

Пример задания с выбором нескольких правильных ответов

Сейчас почти везде начинают применяться задания с выбором нескольких правильных ответов⁴. Оно тоже начинается с инструкции:

Вашему вниманию предлагаются задания, в которых могут быть один, два, три и большее число правильных ответов. Нажимайте на клавиши с номерами всех правильных ответов:

6. БУКВА «О» ПИШЕТСЯ В СЛОВАХ

- 1) упл...тнить сроки
- 2) обн...жить пороки
- 3) р...птать на судьбу
- 4) прил...скать собаку
- 5) отк...заться от услуг
- 6) опр...вдать поступок
- 7) угр...жать расправой
- 8) в...пиющий производ
- 9) раск...лить сковороду
- 10) усл...жнить обстановку

Безошибочный выбор всех правильных ответов в каждом задании с выбором нескольких правильных ответов обычно оценивается баллом 2; одна любая допущенная ошибка снижает балл до 1; две допущенные ошибки обнуляют ответ испытуемого.

Задания с выбором нескольких правильных ответов труднее заданий с выбором одного правильного ответа. Они проверяют знания полнее, глубже и точнее. Они технологичнее и эффективнее заданий с выбором одного правильного ответа. Это самые перспективные задания для образовательных технологий XXI века.

Теория

2

2

Источник: Российская газета, 23 марта 2001 г.

3

Источник: «Педагогическая диагностика». № 2, 2004. С. 88. Задание написано так, как это было в источнике.

4

Аванесов В.С. Форма тестовых заданий. М.: Центр тестирования. 2006. 156 с.

Необходимость эмпирического обоснования качества дистракторов

Качество дистрактора формируется совокупностью свойств, повышающих у испытуемых вероятность его выбора в процессе тестирования. Это может быть высокая правдоподобность ответа, привлекающая оригинальность, краткость, кажущаяся обоснованность. Хорошие дистракторы обычно создают опытные педагоги, знающие типичные ошибки студентов и учащихся.

Если предполагаемый дистрактор не выбирается ни одним испытуемым, то это может означать, что он не обладает свойствами дистрактора. Это случай несоответствия названия элемента его сути, а потому он теряет право на название дистрактора. Его надо либо переформулировать и опробовать заново, либо заменить другим дистрактором. При отсутствии возможности замены количество ответов просто сокращается.

Часто спрашивают, должно ли быть в каждом задании одинаковое число ответов? Ответ такой: нет, не должно быть. В каждом задании общее число ответов и число правильных ответов желательно иметь различным.

Каждое тестовое задание имеет заранее проверенное число дистракторов. Задание с непроверенными дистракторами в педагогической теории подпадает под название задания в тестовой форме.

Эффективность дистрактора

Дистрактор можно считать эффективным, если его выбирает заметная часть испытуемых, имеющих низкую подготовленность. Это может означать, что он хорошо выполняет свою миссию дифференциации испытуемых по уровню подготовленности.

Если дистрактор оказывается привлекательным для испытуемых с высоким уровнем подготовленности, то это нежелательное свойство дистрактора. Этот случай может означать необходимость переформулирования как самого задания, так и ответа.

В идеале каждый дистрактор имеет свой контингент. Одни дистракторы привлекают самых слабых испытуемых, другие — испытуемых более высокого уровня подготовленности. Если случится так, что на какой-то дистрактор среагируют испытуемые с очень высоким уровнем подготовленности, то такой дистрактор надо признать неадекватным.

Другой идеал для задания — чтобы каждый его дистрактор притягивал к себе примерно одинаковое число испытуемых соответствующего уровня подготовленности. Но так бывает редко.

Самый низкий порог выбора для дистрактора обычно устанавливается в 5 процентов. Если дистрактор выбирается меньшим числом испытуемых, то он, очевидно, малопривлекательный, и его лучше заменить на более привлекательный для испытуемых дистрактор.

Если какой-то дистрактор притягивает к себе очень большой процент испытуемых — 80% и более, то и такой дистрактор лучше заменить. Остальные дистракторы при этом фактически не работают. Здесь срабатывает правило — когда слишком хорошо, то это плохо.

Актуальность дистракторного анализа

Ежегодно по всему миру миллионы школьников и студентов подвергаются тестированию, добровольно или принудительно, по множеству различных тестов. В России качественного тестового процесса нет и пока что не предвидится. Вместо него есть ряд тестоподобных форм, которые применяются при проведении т.н.

Единого государственного экзамена и при аттестации студентов высших учебных заведений.

Одна из причин отсутствия в России качественных тестов — игнорирование необходимости проводить дистракторный анализ ответов для каждого задания проектируемого теста. Другая причина — отсутствие публикаций по результатам дистракторного анализа.

Необходимость проведения дистракторного анализа вытекает из логики организации качественного процесса педагогического тестирования. Не бывает теста без качественных тестовых заданий. Не бывает и хороших тестовых заданий без качественных и эффективных дистракторов. Поэтому без проведения дистракторного анализа и публикации информации о качестве каждого использованного задания и ответа к ним утверждения о «качестве тестов или измерений» носят, как писали классики, спекулятивный, т.е. предположительный характер.

Методы проведения дистракторного анализа

В практике тестирования используют два метода проведения дистракторного анализа.

Первый метод можно условно назвать частотным или процентным.

Второй метод основан на IRT.

Применение IRT для проведения дистракторного анализа

Среди многочисленных применений IRT для решения множества прикладных образовательных задач в данной статье рассматривается вариант использования IRT для проведения т.н. дистракторного анализа качества заданий в тестовой форме. Напомним, что дистракторами называют неправильные, но правдоподобные ответы в заданиях с выбором одного или нескольких правильных ответов.

Посмотрим на пример задания И. Парчева с выбором одного правильного ответа из трёх, предлагаемых на выбор⁵.

7. ПЛОЩАДЬ КРУГА С РАДИУСОМ 3 см РАВНА

- 1) 9,00 см²
- 2) 18,85 см²
- 3) 28,27 см²

Поскольку каждый испытуемый может выбрать только один ответ из трёх предлагаемых взаимоисключающих случаев, то применима теорема сложения вероятностей: сумма вероятностей выбора любого одного из трёх ответов равна единице.

На рис. 1. из работы И. Парчева приведены три графика, отражающих идею вероятностной функции выбора каждого из трёх ответов. Эти кривые удобно получать посредством программы RUMM 2020.

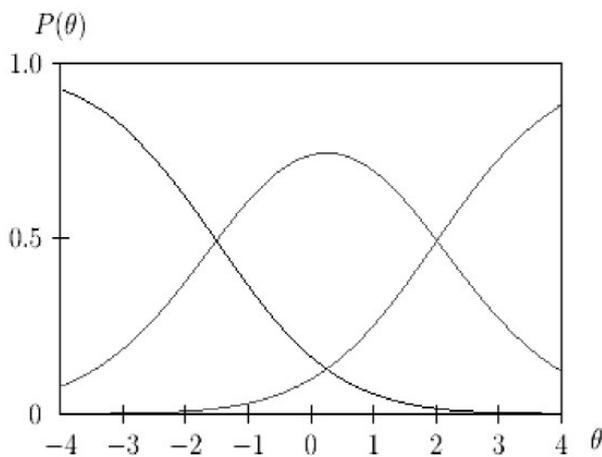


Рис. 1. A partial credit model

Первый и второй ответы неправильные, а третий — правильный. Интересно отметить, что первый и второй дистракторы имеют ответы неодинаковой содержательной ценности.

Первый график указывает на динамику изменения вероятности выбора испытуемыми первого, самого слабого дистрактора. Трудно объяснить, почему подготовленные испытуемые могли бы выбрать первый ответ в качестве правильного, в то время как выбор второго ответа можно объяснить интерференцией — испытуемые путают формулы определения площади (третий, правильный ответ) и длины окружности (второй ответ).

Иначе говоря, выбор первого ответа в этом задании свидетельствует о полном незнании данного фрагмента геометрических знаний, в то время как выбор второго дистрактора — о знаниях хотя бы одной из двух формул, применённых, однако, неправильно. И наконец, выбор только третьего ответа может свидетельствовать о знании и умении применить формулу для определения площади круга.

Соответственно, первый график, изображённый на рис. 1, показывает, что по мере роста уровня подготовленности испытуемых вероятность выбора первого ответа заметно уменьшается. Этот математический факт вполне согласуется с пе-

дагогической логикой. Чем выше уровень подготовленности испытуемых, тем меньше следует ожидать выбора неправильного ответа. Этот ответ в качестве правильного могут выбирать только незнающие испытуемые.

Второй график, изображённый на рис.1, показывает на иной характер изменения вероятности выбора испытуемыми второго, интерферентного, т.е. похожего на правильный ответ, дистрактора. Из графика этой функции видно, как меняется вероятность выбора второго ответа в зависимости от уровня подготовленности испытуемых. Максимум вероятности выбора такого ответа приходится на испытуемых среднего уровня подготовленности.

И только график третьей функции, изображённый на рис. 1, чётко указывает на педагогически осмысленное поведение испытуемых с хорошей подготовкой: чем выше уровень подготовленности испытуемых, тем выше у них вероятность правильного ответа.

Заключение

Качество и эффективность дистракторов можно улучшить вместе с качеством каждого задания в тестовой форме. Для этого рекомендуется:

ПЕД
измерения

1) представлять задания в логической форме высказывания, а не вопроса;

2) устранять в заданиях глаголы в повелительном наклонении типа «оцените», «назовите», «решите» и т.п.;

3) прибавить число ответов, там, где можно, с целью уменьшить вероятность угадывания. Число ответов и правильных ответов в разных заданиях может быть разным. Они

не должны быть одинаковыми во всех заданиях;

4) использовать в тесте для проверки знаний, в основном, задания с выбором нескольких правильных ответов. Такие задания практически исключают догадку;

5) в ответах вместо номера и точки лучше писать номер и круглую скобку, что даёт возможность писать текст ответов со строчных букв.