
РАЗДЕЛ IV

ДОПОЛНИТЕЛЬНЫЕ ТЕМЫ



Данный раздел охватывает дополнительные темы, имеющие отношение к тестированию.

ГЛАВА 8

ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ АНАЛИЗА ТЕСТОВОГО ЗАДАНИЯ



Многие школы предоставляют преподавателям итоговый анализ тестовых заданий после каждого экзамена с применением тестовых заданий множественного выбора. Этот анализ является превосходным источником информации о тестовых заданиях и полезен для оценки их качества, а также оценки точности ответов-ключей.

В качестве примера ниже приведены результаты анализа 4-х тестовых заданий, отражающие распространенную ситуацию. Тестируемые студенты были разделены на лучшую (Hi) и худшую (Lo) группы, основываясь на их результатах по всему тесту. Если у вас небольшое число экзаменуемых, включите в группу Hi 50% лучших студентов, а оставшиеся 50% включите в группу Lo. Если число экзаменуемых большое, в группу Hi можно включить 25% студентов с лучшими результатами, а в группу Lo - 25% с худшими.

Обычно, итоговый анализ отдельного тестового задания показывает процент студентов в каждой из групп, выбравших тот или иной вариант ответа. Часто в него также включаются некоторые показатели уровня сложности тестового задания (например, "р-величина" или процент студентов, выбравших правильный ответ), а также степени дискриминации (например, бисериальную или точечную). При итоговом анализе мы рекомендуем обращать больше внимания на характер ответов, а не на уровень сложности или индекс дискриминации.

В приведенных ниже примерах для каждого варианта ответа показан процент студентов, выбравших данный ответ в качестве правильного. Итоговый ряд в этой таблице содержит проценты ответов по всей группе тестируемых. Например, в тестовом задании #1 в группе Hi ответ А выбрали 1% студентов, ответ В тоже 1%, С - 94%, D - 4%, Е - 1%, F - 2%. В том же задании по группе Lo ответ А выбрали 21% экзаменуемых, ответ В - 6% и т.д. Звездочкой отмечен ответ В, который подразумевался, как правильный.

Тестовое задание #1

Группа	А	В*	С	Д	Е	F
Hi	1	1	94	4	1	2
Lo	21	6	53	15	6	3
Все	9	2	77	8	3	2

р-величина: 2 индекс дискриминации: -0,21

Интерпретация: Это типичный пример тестового задания с некорректно выбранным правильным ответом: если правильный ответ - Вариант В, то задание является очень сложным с отрицательным индексом дискриминации, а количество студентов, правильно ответивших на него, равно только 2%. Напрашивается предположение, что правильным ответом является ответ С, но в этом можно убедиться только после анализа данного тестового задания экспертом по предмету. Если правильным окажется ответ С, то р-величина становится равной 77, а индекс дискриминации 0,46. Оба эти значения являются исключительно хорошими с точки зрения статистики, и нет смысла изменять тестовое задание.

Тестовое задание #2

Группа	А	В	С*	Д	Е	F
Hi	0	1	91	4	4	3
Lo	0	1	61	29	8	6
Все	0	1	76	15	7	6

р-величина: 76 индекс дискриминации: 0,33

Интерпретация: 91% тестируемых из группы Hi и 61% группы Lo выбрали правильный ответ. Это отличные статистические данные. Следовало бы переписать ответы А и В перед тем, как снова использовать данное тестовое задание, поскольку лишь отдельные студенты выбрали их.

Тестовое задание #3

Группа	A	B	C*	D	E	F
Hi	45	1	51	2	1	2
Lo	21	15	21	23	20	2
Все	33	7	34	15	11	2

р-величина: 34

индекс дискриминации: 0,30

Интерпретация: 51% тестируемых из группы Hi и 21% группы Lo выбрали правильный ответ. Это очень сложное тестовое задание и, вероятно, С НИМ НЕ ВСЕ В ПОРЯДКЕ. Слишком много тестируемых из группы Hi выбрали вариант A; возможно, тестовое задание плохо сформулировано. Проверьте также, насколько хорошо изложен ответ A и убедитесь, не является ли он в равной степени правильным.

Тестовое задание #4

Группа	A	B	C	D	E	F
Hi	19	10	51	17	2	2
Lo	25	24	21	26	4	2
Все	23	17	34	23	3	3

р-величина: 34

индекс дискриминации: 0,30

Интерпретация: Ответы на вариант C по группам Hi и Lo разделились, как в тестовом задании #3, но с данным заданием может быть ВСЕ В ПОРЯДКЕ. В отличие от задания 3 тестируемые, не знающие правильного ответа, распределились по дистракторам. Конечно было бы желательно проанализировать варианты A, B и D на правильность и ясность формулировки.

ГЛАВА 9

КАК УСТАНОВИТЬ КРИТЕРИЙ СДАЛ/НЕ СДАЛ



ОПРЕДЕЛЕНИЯ И ОСНОВНЫЕ ПРИНЦИПЫ.

Критерии могут быть классифицированы как *относительные* и *абсолютные*. *Относительный критерий* основывается на результатах, показанных группой, сдающей тест. Экзаменуемые “сдают” или “не сдают” тест в зависимости от того, насколько хорошие результаты они показали относительно других участников тестируемой группы. Приведем примеры *относительных* критериев:

Набравшие менее среднего результата минус 1,2 стандартного отклонения считаются несдавшими тест. Несдавшими считаются 20% худших из группы.

В отличие от относительного, для определения *абсолютного критерия* не проводится сравнение результатов одного экзаменуемого относительно других, принимавших участие в тестировании. Решение “сдал” или “не сдал” в этом случае определяется только результатом экзаменуемого, независимо от результатов остальных тестируемых. При этом все экзаменуемые могут сдать тест или все могут не сдать его. Приведем пример *абсолютного* критерия:

Несдавшими считаются те, которые ответили правильно менее чем на 60% вопросов.

Если нет веских причин “отсеять” заданное количество экзаменуемых, то абсолютный критерий (основанный на результатах каждого экзаменуемого) предпочтительнее относительного (основанного на определенном проценте отсева).

Основные принципы установки критериев.

- Независимо от используемой процедуры, установка критериев требует экспертного заключения. Установка критериев всегда произвольна, но не должна быть чьей-либо прихотью.
- Если нет определенной причины отсеять заданное количество экзаменуемых (например, имеется фиксированное число вакантных мест), то критерий, основанный на уровне овладения предметом, предпочтительнее критерия, основанного на определенном проценте отсева.

- К процессу установки критериев разумно привлекать разносторонне информированных экспертов. При этом неизбежно появляются различные мнения, что уменьшает эффект ястреба/голубки.
- На определенном этапе установления критериев экспертам должны быть предоставлены данные об успеваемости экзаменуемых. Установка критериев без учета таких данных может привести к неинформативным критериям и необоснованным результатам.

Полезную информацию, как установить критерии, можно найти в:

Livingston SA, Zieky MJ. (1982) *Passing Scores: A Manual for Setting of Performance on Educational and Occupational Tests*. Princeton: Educational Testing Service.

ДВА МЕТОДА УСТАНОВКИ КРИТЕРИЕВ, ОСНОВАННЫЕ НА ЭКСПЕРТНОМ ЗАКЛЮЧЕНИИ О ТЕСТОВЫХ ЗАДАНИЯХ

Модифицированная процедура Эбеля.

- Группа экспертов обсуждает характеристики "пограничного экзаменуемого" - экзаменуемого, навыков которого как раз достаточно, чтобы сдать экзамен.
- Эксперты подразделяют тестовые задания на "Необходимые", "Важные" и "Индикаторные".
- Эксперты определяют число тестовых заданий в каждой категории, на которые должен правильно ответить по "граничный экзаменуемый".
- Критерий сдал/не сдал рассчитывается как суммарный процент возможных баллов "пограничного экзаменуемого".

Модифицированная процедура Ангоффа.

- Группа экспертов обсуждает характеристики "пограничного экзаменуемого".
- По каждому тестовому заданию эксперты оценивают возможный процент пограничных экзаменуемых, которые ответят на него правильно.
- Критерий сдал/не сдал вычисляется, как среднее значение прогнозируемых процентов по каждому тестовому заданию.

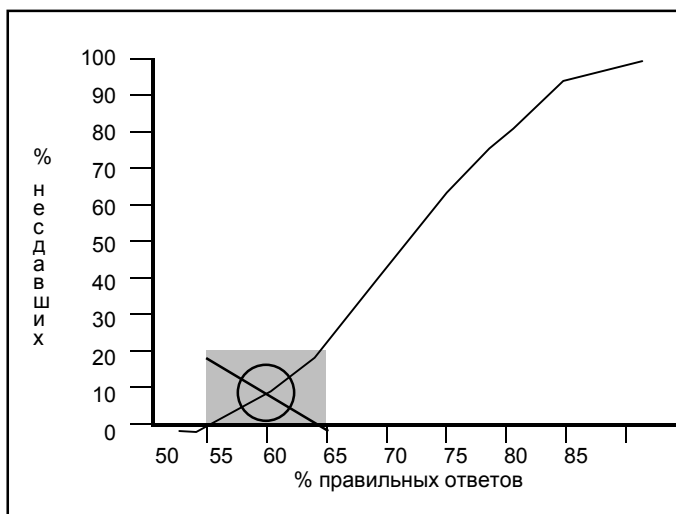
Типичные варианты процедуры Ангоффа.

- Экспертам могут предоставляться или не предоставляться правильные ответы на вопрос.
- Экспертам может предоставляться или не предоставляться информация о проценте экзаменуемых, правильно ответивших на каждое тестовое задание.
- После подготовки эксперты могут работать как в группе, так и индивидуально.

ОТНОСИТЕЛЬНЫЕ И АБСОЛЮТНЫЕ КОМПРОМИСНЫЕ КРИТЕРИИ: МЕТОД ХОФСТИ

В последнее время было разработано несколько так называемых "компромисных моделей", которые используют положительные стороны как относительных, так и абсолютных процедур установки критерия. Один из таких методов - метод Хофсти - описан ниже.

1. Эксперты просматривают копию экзамена.
2. Затем эксперты устанавливают следующие значения, которые определяют приемлемые критерии:
 - Наименьший приемлемый процент экзаменуемых, не сдавших экзамен (минимальная "норма" провалов).
 - Наибольший приемлемый процент экзаменуемых, не сдавших экзамен (максимальная "норма" провалов).
 - Наименьший результат, позволяющий сдать экзамен (минимальный проходной балл).
 - Наибольший результат, который можно потребовать от экзаменуемых для сдачи экзамена (максимальный проходной балл).
3. После тестирования строится кривая процента несдавших, как функция от проходного балла (процента правильных ответов). (На приведенном графике кривая поднимается слева направо).
4. Четыре значения, полученные в #2, формируют прямоугольник. Часто для этого используются средние значения по группе экспертов. В нашем примере соответствующий уровень отсева по оценкам экспертов находится между значениями 0 и 20% (горизонтальные пунктиры на графике), а соответствующая точка сдал/не сдал находится между значениями 50% и 60% (вертикальные пунктиры).
5. Строится диагональ от верхнего левого угла прямоугольника к нижнему правому. Точка пересечения диагонали с кривой дает численное значение критерия (т.е. в нашем примере несколько выше 55 % правильных ответов).



Полезную информацию о компромисных методах можно найти в:

de Gruijter D. Compromise models for establishing examination standards & *Journal of Educational Measurement*. 1985; 22:263-269.

ГЛАВА 10

О НЕКОТОРЫХ АСПЕКТАХ ТЕСТИРОВАНИЯ



Мы предлагаем вам свои размышления по достаточно разнородным темам. В целом высказанные идеи гипотетичны и базируются скорее на отдельных случаях, а не на доказательствах. Поэтому они отражают наши домыслы, а не являются результатом строгих научных исследований.

Многоэтапные экзамены (напр., практические экзамены, “стипельчез”, OSCEs - Objective Structured Clinical Examinations - Объективные Структурированные Клинические Экзамены).

Многоэтапные экзамены, несмотря на то, что они более сложны для подготовки и проведения, являются очень полезными в фундаментальных дисциплинах, особенно при оценке манипуляционных навыков, которые не могут быть оценены при бланковом тестировании (например, умение работать с микроскопом, выполнение лабораторных процедур). Кроме того, воспроизведение некоторых тестовых материалов (радиологические изображения, цветные иллюстрации) обходится очень дорого. В такой ситуации многоэтапный подход к проведению теста может снизить организационные расходы.

Домашние экзамены.

Домашние экзамены могут являться эффективным учебным методом для студентов, заставляя студентов читать много и глубоко вникать в важные разделы. К сожалению, студенты стремятся исписать ответами целые тома, и не всегда понятно, являются ли эти ответы их собственной работой. Такие же преимущества могут быть достигнуты путем распространения тестовых заданий заранее (в большем количестве, чем будет на экзамене), а сдача экзамена (части предварительно распространенных вопросов) будет производиться в течение фиксированного времени.

Экзамены с открытой книгой

Такие экзамены могут быть очень хороши тем, что влияют на типы вопросов, подготавливаемых преподавателями. В экзаменах с открытой книгой бессмысленно задавать вопросы об отдельных фактах, которые можно быстро найти на одной странице учебника, поэтому при разработке тестовых заданий для таких экзаменов внимание уделяется пониманию ключевых концепций и принципов в проблемных ситуациях.

Что лучше: частые короткие опросы или редкое тестирование?

Редко проводимое тестирование делает каждый экзамен важным событием; студенты могут даже прекратить посещение занятий для подготовки к тестированию, что кажется нежелательным. Кроме того, редко проводимое тестирование приводит к тому, что студенты не могут определить, изучают ли они необходимый материал, и достаточно ли глубоко он изучается. Поэтому хотя частое тестирование занимает

больше времени у преподавателей, оно снижает важность каждого отдельного экзамена и помогает студенту лучше оценить свой прогресс. В целом, частое тестирование кажется более предпочтительным, хотя студенты склонны жаловаться независимо от предлагаемого подхода.

Хранить ли тесты “в секрете” или позволять студентам знакомиться с ними ?

Поскольку тесты могут оказывать значительное направляющее воздействие на процесс изучения материала, то ознакомление студента с тестовыми заданиями безусловно помогает ему концентрировать внимание на ключевых темах, способствует достижению целей курса и освоению учебного плана (если, конечно, тестовые материалы отражают их). Однако подготовка хороших экзаменационных вопросов является очень длительной процедурой и со временем качество тестовых материалов может снизиться, если преподаватели вынуждены каждый раз разрабатывать новые тестовые материалы к началу курса. Возможно, наилучшим подходом является ознакомление студентов с примером качественных тестовых материалов с целью воздействия на процесс обучения, но при этом сохранять банк “секретных” заданий для повторного использования. Однако следует помнить, что секретность будет невысокой, поскольку обычно студенты запоминают вопросы и обмениваются друг с другом.

Использование накопительных тестов.

Накопительные тесты, которые требуют от тестируемых знания всего пройденного на данный момент учебного материала, обращают внимание студентов на взаимосвязи между темами, особенно если тестовые вопросы требуют понимания как ранее пройденного, так и нового материала. Использование тестов, которые включают только материал, усвоенный после предыдущего тестирования, способствуют тому, что студенты изучают темы как бы изолированно, теряя между ними взаимосвязь. Поскольку иногда студенты плохо справляются с серией экзаменов из-за того, что они так и не изучили базовый материал, этот подход может также побуждать студентов восполнить пробелы в знаниях.

Использование интегрированных межпредметных экзаменов.

Подобно использованию накопительных тестов, интегрированные межпредметные экзамены способствуют осмыслению студентами взаимосвязи между дисциплинами и темами. Это оказывается очень полезным для прочного усвоения материала и для успешного применения знаний фундаментальных наук в клинической практике. Вообще говоря, такие экзамены должны готовиться совместно преподавателями фундаментальных и клинических кафедр. Несмотря на затраты времени, такие объединенные усилия в результате могут дать значительно лучшие тестовые материалы, а также способствовать полезной дискуссии между преподавателями о том, какие материалы должны включаться в учебный план.

ПРИЛОЖЕНИЕ

КЛАДБИЩЕ ФОРМАТОВ ТЕСТОВЫХ ЗАДАНИЙ NBME



Первый экзамен Национального Совета был проведен в 1916 году. Это был 5-ти дневный экзамен, состоящий из письменного, устного, лабораторного и клинического тестов. После Первой Мировой войны Национальный Совет создал комиссию, которая должна была посетить Англию, Шотландию и Францию для ознакомления с методами оценки, используемыми для медицинского лицензирования. В 1922 г Национальный Совет провел новый экзамен: Часть I - трехдневный письменный (в форме эссе) экзамен по фундаментальным дисциплинам; Часть II - двухдневный письменный (в форме эссе) экзамен по основным клиническим наукам; и Часть III - однодневный устный экзамен, проходивший у постели больного. В таком виде экзамены проводились до начала 50-х годов, когда стало популярным тестирование с помощью вопросов множественного выбора. В 1951 г Национальный Совет в сотрудничестве с Образовательной Службой по Тестированию начали трехгодичное исследование по сравнению письменных (в форме эссе) тестов с тестами множественного выбора. Итогом явилось решение об изменении в 1953 г I и II Части с экзамена в виде письменного эссе на тестирование множественного выбора. С введением новых форматов Часть III была пересмотрена в начале 60-х годов и упразднена. Первый экзамен состоял главным образом из тестовых заданий А-типа, требующих вспоминания фактов, как в следующем примере:

Начиная с 40-х годов, когда был впервые проведен экзамен с вопросами множественного выбора, Национальный Совет расширил сферу применения тестовых заданий А-типа для проверки навыков рассуждения и умения решать проблемы путем введения клинических ситуаций в большинство условий заданий. В настоящее время А-тип остается наиболее широко используемым форматом в лицензионных экзаменах. Применение многих других форматов тестовых заданий, которые были разработаны на протяжении этого периода, было прекращено. Эти форматы, пронумерованные буквами в порядке разработки, описаны ниже.

Из ниже перечисленного наиболее эффективным профилактическим средством для предотвращения рецидивов ревматизма является:

- A. ацетилсалициловая кислота*
- B. парааминобензойная кислота*
- C. адренокортикотропный гормон*
- D. кортизон*
- E. сульфадиазин*

Тестовые задания, представленные в этой части взяты из Hubbard JP & Clemens WV(1961), Mulyiple-Choice Examinations in Medicine. Lea & Febiger, Philadelphia. Критика изложенных форматов взята из Hubbard JP, Measuring Medical Education. Lea & Febiger, Philadelphia, PA, 1971; Hubbard JP, Levit E.J. The National Board of Medical Examiners: The First Seventy Years. NBME, Philadelphia, PA, 1985, и из различных статей Мортон, опубликованных в Federation Bulletin в 1985 и 1986 гг.

Тестовые задания В-типа

Тестовые задания В-типа являлись заданиями соответствия, которые состояли из списка названий, обозначенных буквами, и списка пронумерованных слов или фраз. Экзаменуемый должен был выбрать название, которое больше всего подходит для каждого слова или фразы. Поскольку каждый ответ мог использоваться более одного раза или не использоваться вообще, тестовые задания В-типа не могли быть решены методом исключения. Предполагалось, что тестовые задания В-типа расширят применимость тестов множественного выбора, позволяя оценить несколько связанных предметов в одном блоке тестовых заданий. В отличие от тестовых заданий соответствия, используемых сегодня, задания В-типа не всегда имели вводный вопрос; в результате задаваемый вопрос иногда был неясен. Эти тестовые задания в целом проявили себя с хорошей стороны и их применение было прекращено недавно, когда распространился формат расширенного выбора.

Пример тестового задания В-типа

ИНСТРУКЦИЯ: Каждый блок вопросов соответствия в данном разделе содержит список из 3-5 помеченных буквами вариантов ответа (некоторые могут быть выражены цифрами), за которыми следует несколько пронумерованных тестовых заданий. К каждому пронумерованному заданию подберите **ОДИН** наиболее подходящий ответ, помеченный буквами, и закрасьте кружок, который соответствует выбранной букве в вашем бланке ответов. Каждый помеченный буквой ответ может быть выбран один раз, больше одного раза или вообще не выбран.

- A. Коарктация аорты
 - B. Открытый артериальный проток
 - C. Тетрада Фалло
 - D. Аортальное сосудистое кольцо
 - E. Атрезия трехстворчатого клапана
-
1. Улучшается вследствие системно-легочного анастомоза
 2. Наиболее распространенный тип врожденного “синего” порока сердца
 3. Хирургически корректируется путем резекции и наложения анастомоза конец-в-конец
 4. Возможная причина дисфагии у новорожденных и детей
 5. Гипертензия на верхних и гипотензия на нижних конечностях

Тестовые задания D-типа

Тестовые задания D-типа являются комплексными тестовыми заданиями соответствия, в которых каждое задание состоит из трех функциональных нарушений, отмеченных буквами, и пяти пронумерованных ситуаций. Тестируемый должен (1) выбрать функциональное нарушение или категорию, с которым связаны 4 из 5 ситуаций и (2) указать одну ситуацию, которая не относится к данной категории. Предполагалось, что такие тестовые задания дадут возможность выявить понимание различий между рядом похожих факторов, однако D-тип оказался сложным в написании, а инструкции запутанными. К тому же они не позволяли распределить знающих и незнающих экзаменуемых.

Пример тестового задания D-типа

ИНСТРУКЦИЯ: На *каждый* из следующих вопросов надо дать два ответа. В списке слева приведены три категории, обозначенные буквами. Четыре из пяти пронумерованных тестовых заданий из списка справа тем или иным образом относятся к **ОДНОЙ** из этих категорий. Дайте ответы на следующие вопросы. (1) На соответствующей строчке в бланке ответов закрасьте букву категории из списка слева, которой соответствуют четыре тестовых задания. (2) Закрасьте цифру задания из списка справа, которое **НЕ** относится к той же категории, что и остальные четыре.

- | | |
|--|-------------------------|
| A. Диагностически значимая эозинофилия | 1. Трихинеллез |
| B. Диагностически значимый плазмацитоз | 2. Множественная миелом |
| C. Диагностически значимый лимфоцитоз | 3. Синдром Леффлера |
| | 4. Болезнь Ходжкина |
| | 5. Шистосомоз |

Тестовые задания К-типа

Тестовые задания К-типа были наиболее широко используемым в Национальном Совете форматом типа верно/неверно. Они состояли из условия задания и четырех вариантов ответа, один или более из которых были правильными ответами. Считалось, что тестовые задания К-типа проверяли глубокие знания и понимание различных аспектов заболевания, процесса или метода и требовали от тестируемого знания различных фактов по данной теме. Однако, тестовые задания К-типа критиковались за то, что они слишком сложны и требовали от тестируемого постоянно помнить комбинацию ответов и соответствующую ему букву. Кроме того, возможная комбинация ответов представляла собой подсказку, что уменьшало разрешающую способность тестового задания и уменьшало надежность теста. Было сложно написать хорошие недвусмысленные тестовые задания формата верно/неверно. Поскольку задания могли включать только абсолютно верные или неверные факты, формат К-типа не мог использоваться для оценки клинического мышления, кроме как в сравнении (например, "препарат X лучше препарата Y при лечении заболевания Z"). Тестовые задания К-типа были более сложными и обладали меньшей разрешающей способностью, чем другие типы заданий. Так же они были менее эффективны, чем другие форматы вопросов множественного выбора и имели меньшую относительную достоверность на единицу времени тестирования.

Пример тестового задания К-типа

<u>Инструкции Обобщены</u>				
A	B	C	D	E
только 1,2,3	только 1,3	только 2,4	только 4	Все правильны

1. У ребенка с обострением ревматизма обычно наблюдается
 - (1) повышенная скорость оседания эритроцитов
 - (2) удлинение интервала PR
 - (3) увеличение титр антистрептолизина O
 - (4) подкожные узелки

Тестовые задания С-типа

Тестовые задания С-типа похожи на В-тип по внешнему виду, но экзаменуемый должен определить верные/неверные ответы. Тестовое задание С-типа состоит из списка обозначенных буквами названий, за которым следует перечень пронумерованных слов или фраз. Для каждого пронумерованного тестового задания экзаменуемый должен решить, правилен ли ответ А, правилен ли ответ В, оба ответа правильны (вариант С), либо все ответы неправильные (вариант D). Этот тип тестовых заданий использовался для сравнения и противопоставления двух заболеваний, признаков, симптомов, лабораторных данных и т.д. По уровню сложности тестовые задания С-типа соответствуют К-типу. Основная проблема С-типа состояла в принятии решения, в какой мере что-либо должно быть “верным”, чтобы быть выбранным. Если, например, что-либо ассоциируется с обоими ответами А и В, но больше подходит к А, тестируемый должен решить, выбирать ли ему ответ А, или оба А и В. При относительно слабой связи, экзаменуемый должен решить, достаточно ли сильна связь, или "все ответы неправильны" является подходящим ответом. Эти суждения не связаны с медицинскими знаниями, а заставляют тестируемого думать о том, что подразумевали авторы тестового задания.

Пример тестового задания С-типа

ИНСТРУКЦИЯ. Каждый блок вопросов соответствия в данном разделе состоит из списка 4-х обозначенных буквами вариантов ответов, за которым следует несколько пронумерованных тестовых заданий. Для каждого пронумерованного задания выберите ОДИН вариант ответа (обозначенный буквой), который наиболее точно соответствует заданию, и закрасьте кружок с соответствующей буквой в вашем бланке ответов. Каждый обозначенный буквой вариант ответа может быть выбран один раз, более одного раза, или не выбран ни разу.

- A. Малярия, вызванная *Plasmodium vivax*
- B. Малярия, вызванная *Plasmodium falciparum*
- C. Оба заболевания
- D. Ни одно заболевание

1. Комбинация примахина и хлорохина - средство выбора при острой атаке.
2. Клинические проявления подавляются приемом хлорохина один раз в неделю в период пребывания в эндемичном районе.
3. Инфекция полностью излечима хлорохином.
4. Инфекция предотвратима приемом хлорохина один раз в неделю.

Тестовые задания Е-типа

Тестовые задания Е-типа с множественным верными или неверными ответами основывается на анализе отношений. Тестируемые, знающие Е-тип тестовых заданий, до сих пор описывают их как "Верно, Верно и Не относится". Е-тип состоит из предложения с двумя основными частями: утверждение и причина к этому утверждению. Тестируемый получал инструкцию выбрать А, если оба варианта были верными утверждениями, а причина была верным объяснением данного утверждения; В - если оба варианта были верными утверждениями, но причина не была верным объяснением данного утверждения; С - если утверждение было верно, а причина - неверное заявление; D - если утверждение неверно, а причина верное заявление; Е - если и утверждение, и причина были неверными заявлениями. Считалось, что для правильного ответа на это задание необходимо обладать навыком рассуждения и пониманием основных принципов. Однако тестовые задания Е-типа были сложны для создания, а экзаменуемые считали их запутанными.

Пример тестового задания Е-типа

Инструкции обобщены

A	Верно	Верно	Причина является правильным объяснением
B	Верно	Верно	Причина НЕ является правильным объяснением
C	Верно	Неверно	
D	Неверно	Верно	
E	Неверно	Неверно	

Утверждение

Причина

- | | | |
|---|-----------|--|
| 1. Простой герпес обычно считается аутогенной инфекцией | ПОСКОЛЬКУ | у больных, получающих жаропонижающее лечение, часто развивается герпес |
| 2. При вскармливании младенцев коровье молоко предпочтительнее грудного | ПОСКОЛЬКУ | коровье молоко содержит больше кальция |

Тестовые задания Н-типа

Тестовые задания Н-типа были сравнительными заданиями, которые состояли из спаренных утверждений, описывающих два объекта, которые необходимо сравнить количественно. Тестируемый получал инструкцию выбрать А, если А больше чем В; В, если В больше, чем А; и С, если оба приблизительно равны.

Хотя и было общепризнано, что надо ограничивать вопросы, которые зависят от запоминания абсолютных количественных величин, однако считалось, что тестовые задания Н-типа полезны в тех случаях, когда знание количественной информации является важным. Проблема же для тестируемых состояла в определении насколько велика должна быть разница, чтобы быть значимой.

Пример тестового задания Н-типа

ИНСТРУКЦИЯ: Следующие спаренные утверждения описывают два объекта, которые надо сравнить в количественном отношении. В соответствующей строке бланка ответов отметьте нужную букву:

А если (А) больше, чем (В)

В если (В) больше, чем (А)

С если оба они равны или очень близки.

1. (А) Обычная терапевтическая доза эpineфрина
(В) Обычная терапевтическая доза эфедрина
2. (А) Ожидаемая продолжительность жизни при глиобластоме затылочной доли
(В) Ожидаемая продолжительность жизни при глиобластоме лобной доли

Тестовые задания I-типа

Тестовые задания I-типа были подобны заданиям H-типа. Они содержали пары фраз, описывающие условия или количества, которые могли изменяться относительно друг друга. Экзаменуемый должен был выбрать А, если обе фразы имели прямую связь (т.е. возрастание или уменьшение параметра в первой фразе сопровождается возрастанием или уменьшением параметра во второй); В, если фразы имеют обратную связь (т.е. возрастание или уменьшение параметра в первой приводит соответственно к уменьшению или возрастанию параметра во второй); С, если параметры в обеих фразах изменяются независимо).

Пример тестового задания I-типа

ИНСТРУКЦИЯ: Каждая из следующих парных фраз описывает условия или количества, которые могут связаны или не связаны между собой. В соответствующей строке бланка ответов отметьте соответствующую букву:

- А** если возрастание параметра в первой сопровождается возрастанием во второй, или уменьшение в первой сопровождается уменьшением во второй.
- В** если возрастание параметра в первой сопровождается уменьшением параметра во второй, или если уменьшение в первой сопровождается увеличением во второй.
- С** если изменение параметра в первой не обязательно сопровождается изменением параметра во второй.

1. (А) Объем мочи
(В) Удельный вес мочи
2. (А) Концентрация белков плазмы
(В) Коллоидное осмотическое давление плазмы

Ни H-, ни I- типы форматов тестовых заданий не получили широкого распространения. Из-за меньшего количества вариантов ответов по сравнению с другими типами тестовых заданий, вероятность угадывания правильного ответа в них была велика. Кроме того, такие задания концентрируются скорее на отдельных деталях, чем на научных положениях.

В серии своих статей в Federation Bulletin (1985-86) Мортон предположил, что различные типы тестовых заданий включались в медицинские лицензионные экзамены просто для того, чтобы внести разнообразие в эту длительную процедуру. Однако, через 25 лет после того, как Национальный Совет принял решение о переходе с письменных экзаменов в виде эссе на тестирование с помощью вопросов множественного выбора, NBME проанализировал все исследования по различным типам использованных вопросов, и все разнообразие типов тестовых заданий было ограничено типами А-, В-, С-, G-, К-, Х -, М-. Эти типы еще раз пересматривались специалистами в середине 80-х годов. Тогда они пришли к единодушному мнению, что четыре основных типа тестовых заданий обеспечивают достаточное разнообразие для того, чтобы оценить знания, необходимых для получения сертификата Национального Совета. Эти четыре базисных типа включают А-, В-, С-, К-типы тестовых заданий. G-тип (блоки тестовых заданий А-типа), N-тип (блоки заданий К-типа) и М-тип больше не рассматриваются в качестве отдельных форматов.

Еще позднее типы тестовых заданий снова подверглись пересмотру. В настоящее время лицензионные экзамены включают в себя только А- и R-типы тестовых заданий. Для улучшения экзаменов были также предприняты следующие шаги: концентрация внимания на психометрически обоснованных типах тестовых заданий; обучение авторов тестовых заданий технике их написания; больше внимания уделяется тестовым заданиям, оценивающим способность принятия решений в клинических ситуациях, а не заданиям на вспоминание; и предварительная экспертиза новых тестовых заданий.