
РАЗДЕЛ I

ОБЩИЕ СВЕДЕНИЯ О ФОРМАТЕ И СТРУКТУРЕ ТЕСТОВЫХ ВОПРОСОВ



В данном разделе рассматриваются структурные тонкости, необходимые для написания высококачественных тестовых вопросов. В следующем разделе будут рассмотрены вопросы, связанные с содержанием тестовых заданий.

ГЛАВА 1

ВВЕДЕНИЕ



КОНТРОЛЬ - ВАЖНАЯ СОСТАВНАЯ ЧАСТЬ ОБУЧЕНИЯ

Контроль является крайне важной составной частью обучения; при правильном использовании он может содействовать достижению конечных целей обучения. Нельзя переоценить важность принятия решений о том, как и когда проводить оценку знаний и умений ваших студентов.

Первоначальной целью тестирования является обсуждение того, что вы считаете самым важным. Тесты являются мощным стимулирующим фактором, и студенты будут учить именно то, что по их мнению будет оцениваться. Контроль также помогает восполнять пробелы в обучении, побуждая студентов к широкому самостоятельному чтению и активному участию в учебном процессе. Этот результат тестирования особенно важен при изучении клинических дисциплин, когда процесс обучения у студентов различен, что связано со различным спектром заболеваний у пациентов и особенностями отдельных клинических баз. Также данный результат может быть важен при изучении некоторых фундаментальных наук (например при проблемно-ориентированном обучении), когда образовательный опыт студентов отличается.

Так как тесты оказывают сильное влияние на процесс обучения студентов, необходимо разрабатывать такие тесты, которые будут способствовать достижению задач обучения. Введение контроля практических навыков заставляет студентов переходить из библиотеки в клинику, где они стремятся получить навыки врачебного осмотра; с другой стороны, введение тестов, оценивающих только запоминание изолированных фактов, заставляет студентов “зубрить” учебники. Данное руководство сфокусировано на том, как составлять высококачественные вопросы множественного выбора (MCQs), которые оценивают умение студентов интерпретировать данные и принимать решения, что, по нашему мнению, является важной составляющей клинических навыков.

Путь студентов к достаточным или даже отличным знаниям будет менее трудным, если они будут иметь постоянную информацию о своих достижениях в учебе.

ЦЕЛИ ТЕСТИРОВАНИЯ	ЧТО ДОЛЖНО БЫТЬ ПРОВЕРЕНО?
<ul style="list-style-type: none"> • Сообщить студентам, какой материал является важным • Создать у студентов мотивацию к обучению • Найти пробелы в знаниях, требующие коррекции или дополнительного изучения • Определить окончательную оценку или принять решения о переводе на следующий курс • Найти слабые стороны учебной программы/учебного курса 	<ul style="list-style-type: none"> • Содержание экзамена должно соответствовать целям курса/клинического обучения • Оценка важных тем должна иметь больший вес, чем менее важных • Время тестирования по каждой теме должно отражать относительную важность темы • Совокупность тестовых заданий должна быть репрезентативной в отношении учебных целей

ОТБОР ТЕСТОВЫХ ЗАДАНИЙ (ВЫБОРКА)

Целью любого контроля является возможность сделать выводы относительно навыков экзаменуемого - выводы, которые можно распространить с конкретных проблем (или случаев, или тестовых вопросов), включенных в экзамен, на более широкую сферу, из которой были выбраны случаи (или вопросы). Ясно, что контроль требует времени. Также ясно, что если вы уделяете больше времени одному виду деятельности, у вас останется меньше времени на остальные. Независимо от того, разрабатываете ли вы всеобъемлющий план оценки, или же решаете, что включить в конкретный тест, вы все равно столкнетесь с проблемой отбора материала. Результаты, полученные на выборке (отобранных тестовых заданиях), являются основой для оценки достижений в более широкой сфере, что и является конечной целью.

При работе с вопросами множественного выбора (MCQs), вам, во-первых, необходимо решить, что вы хотите включить в тест. Внимание, уделяемое проблеме, должно соответствовать ее относительной важности. Для проверки вы должны выбрать темы и также выбрать навыки (например, постановку диагноза, решение о следующем этапе лечения); вы не можете спрашивать сразу обо всем. Результаты, полученные на выборке (отобранных тестовых заданиях), являются основой для оценивания достижений в более широкой сфере, что и является конечной целью. Характер выборки определяет степень воспроизводимости (надежности, обобщенности) и точности (достоверности, валидности) полученной оценки истинных знаний и умений. Если выборка нерепрезентативна (например, включает только кардиологические вопросы при контроле общемедицинских знаний), результаты экзамена будут недостоверными и не дадут возможности оценить

достижения в интересующей сфере. Если выборка слишком мала, результаты экзамена могут оказаться недостаточно стабильными, чтобы отразить истинные знания и умения.

При использовании тестов с вопросами множественного выбора (MCQs), почти всегда есть единственный “оценщик” (обычно компьютер) и серии или группы вопросов; чтобы сформировать выборку, необходимо отобрать подгруппы вопросов для использования в тесте. При других методах оценки знаний (устный экзамен, основанный на клиническом случае, экзамен со стандартными пациентами, письменный экзамен), выборка является гораздо более сложным процессом, потому что при любом методе, когда нельзя провести механическую оценку, требуется отбор по второму измерению - отбор “оценщика”. На таких экзаменах вы хотите узнать результаты ответов на разные вопросы (случаи) и хотите, чтобы оценка не зависела от личности экзаменатора. Поэтому вам необходимо осуществить выборку в обоих измерениях: одну в отношении вопросов или случаев, а вторую - в отношении экспертов или оценщиков. Вы должны произвести выборку по случаям (вопросам), поскольку ответ на один случай (вопрос) не является очень хорошим прогностическим показателем в отношении ответов на другие вопросы. Вам также необходимо произвести выборку по различным экспертам, чтобы минимизировать эффект жесткости или снисходительности преподавателя или действия других факторов, например, “эффект нимба”, что может вызвать проблему противоречивости оценок у разных оценщиков. При большой выборке пики и спады в результатах ответов, а также пики и спады у различных оценщиков усредняются.

Хотя в данном руководстве больше внимания уделяется вопросам множественного выбора (MCQs), мы считаем, что необходимо также использовать и другие методы контроля. Ни один из методов не может оценить сразу все интересующие умения и навыки.

В трех приведенных ниже работах более подробно обсуждаются вопросы формата тестовых заданий:

Case SM, Swanson DB, Ripkey DR. Comparison of items in five-option and extended-matching format for assessment of diagnostic skills. *Academic Medicine*. 1994; 69 (suppl): S1-S3.

Swanson DB, Case SM. Trends in written assessment: a strangely biased perspective. In: Harden R, Hart I, Mulholland H, eds., *Approaches to Assessment of Clinical Competence*. Norwich, England: Page Brothers; 1992:38-53.

Case SM, Downing SM. Performance of various multiple-choice item types on medical specialty examinations: types A, B, C, K and X. *Proceedings of Twenty-Eight Annual Conference of Research in Medical Education (RIME)*. 1989:167-172.

ГЛАВА II

ФОРМАТЫ ТЕСТОВЫХ ЗАДАНИЙ МНОЖЕСТВЕННОГО ВЫБОРА



Хороший тестовый вопрос должен удовлетворять двум основным критериям. Во-первых, тестовый вопрос должен быть важным по содержанию. Это основополагающее условие, к которому мы обратимся далее в руководстве. Безусловно, содержание тестовых заданий является очень важным, но концентрироваться только на нем недостаточно. Тестовые задания, призванные оценивать критически важные знания, не будут выполнять своего назначения, пока они не будут иметь хорошую структуру. Необходимо избегать ошибок, от которых выигрывают опытные в тестировании студенты, а также избегать чрезмерной сложности. Это предпосылки того, что тестовые задания будут давать валидные результаты.

ВОПРОСЫ “ВЕРНО/НЕВЕРНО” В СРАВНЕНИИ С ВОПРОСАМИ С “ОДНИМ ЛУЧШИМ ОТВЕТОМ”

Все разнообразие вопросов множественного выбора можно разделить на две большие категории тестовых заданий: 1) требующие от экзаменуемого выбрать все подходящие ответы (Верно/неверно), и 2) требующие от экзаменуемого указать один ответ (один лучший ответ).

Каждая категория представлена несколькими специфическими форматами, указанными ниже:

Тестовые задания Формата “Верно/Неверно” требуют от экзаменуемых выбрать все верные варианты

- C (А/В/Оба/Ни один из вышеобозначенных)
 - K (комплексные тестовые задания)
 - X (простые тестовые задания)
- Имитационные клинические задачи, типа Проблемы Ведения Пациента (PMPs)

Тестовые задания форматов с одним лучшим ответом требуют от экзаменуемых выбрать один лучший ответ

- A (четыре или более вариантов ответа, отдельные тестовые задания или блоки)
- B (четыре или пять вариантов ответа для блока из 2-5 тестовых заданий)
- R (тестовые задания расширенного выбора в блоках по 2-20 тестовых заданий)

КАТЕГОРИЯ “ВЕРНО/НЕВЕРНО”

Категории вопросов “Верно/Неверно” и “Один Лучший Ответ” ставят перед экзаменуемым совершенно разные задачи. Тестовые задания формата “Верно/Неверно” требуют от экзаменуемого выбрать все варианты ответов, которые “верны”. В таких тестовых заданиях экзаменуемый должен решить, где проводить разграничение: в какой степени ответ должен быть “верным”, чтобы выбрать его как “верный”. Несмотря на то, что данная задача требует принятия дополнительного решения (по отношению к заданиям с одним лучшим ответом), это дополнительное решение может быть не связано с клиническими знаниями. Слишком часто экзаменуемым приходится угадывать, что подразумевал составитель теста, так как варианты ответов не являются ни абсолютно верными, ни абсолютно неверными.

Следующий пример демонстрирует приемлемое с точки зрения структурного построения тестовое задание¹ формата “верно/неверно”. Условие задания ясно, и варианты ответов либо абсолютно верны, либо абсолютно неверны без двусмысленности.

Что из нижеперечисленного наследуется по X- сцепленному рецессивному типу?

1. Гемофилия А (классич. гемофилия)
2. Муковисцидоз
3. Мышечная дистрофия Дюшенна
4. Болезнь Тея-Сакса

Варианты ответов графически можно представить следующим образом:

2	1
4	3

*Абсолютно
неверные ответы*

*Абсолютно
верные ответы*

К верным утверждениям о муковисцидозе (МВ) относятся:

1. Частота встречаемости МВ 1:2000.
2. Больные МВ обычно умирают в подростковом возрасте.
3. Мужчины с МВ стерильны.
4. МВ наследуется по аутосомно-рецессивному типу.

Данное задание формата “верно/неверно” имеет ошибку. Варианты ответов № 1, 2 и 3 не могут быть расценены как абсолютно верные или неверные даже группой экспертов. Оценивая первый вариант ответа, заметим, что частота встречаемости муковисцидоза не составляет точно 1:2000, экспертам потребовалась бы дополнительная информация: Проводится ли данная оценка в США? Включает ли она все этнические группы? Изменение формулировки на “около 1:2000” не поможет, поскольку данные не уточнены. Подобные проблемы возникнут при обдумывании вариантов ответов №2 и 3, тогда как

¹ По традиции, варианты ответов в заданиях формата “верно/неверно” обозначаются цифрами, а в заданиях с одним лучшим ответом - буквами.

Написанное в шутку, это тестовое задание формата “верно/неверно” иллюстрирует общую проблему - некорректно составленное условие задания. В зависимости от вашего мнения, варианты ответов № 1, 2 и 3 могут быть верными; с другой стороны, №1, 2 и 3 варианты ответов могут быть неверными, тогда как №4 верен.

В данном примере туманная формулировка вариантов ответа является подсказкой для опытного в тестировании студента. Использование глагола “может” в вариантах ответа № 1, 2 и 3 подсказывает, что они верны. Вариант ответа №4 угадать труднее, поскольку надо решить, что означает “обычно”? Исследования показали, что определение частоты явления столь туманным образом не дает однозначного ответа. Эксперты не смогли бы решить, является ли вариант ответа №4 верным или неверным.

Ошибки в этом тестовом задании менее заметны. Трудность заключается в том, что студент должен сделать предположение о серьезности заболевания, возрасте больного и предшествующем лечении. Разные предположения приведут к разным ответам, даже среди экспертов.

Обратите внимание, что во всех приведенных примерах тестовых заданий с дефектами, условие тестового задания неясно, варианты ответов содержат туманные термины, или варианты ответов верны лишь частично. В каждом случае группа экспертов испытывала бы затруднения при выборе правильного ответа.

Путь к сердцу мужчины лежит через:

1. *аорту*
2. *легочные артерии*
3. *легочные вены*
4. *желудок*

При клинической оценке хронических болей

1. *личное отношение врача к боли может повлиять на его медицинское заключение*
2. *отрицательные эмоции могут преобразиться в жалобы на физическую боль*
3. *боль может иметь символическое значение*
4. *выражение лица и поза больного обычно являются ключом к оценке силы болевых ощущений.*

Дефект межжелудочковой перегородки у детей сопровождается:

1. *систолическим шумом*
2. *легочной гипертензией*
3. *тетрадой Фалло*
4. *цианозом*

Поскольку экзаменуемые должны выбрать все “верные” варианты ответа, тестовые задания формата “верно/неверно” должны удовлетворять следующим требованиям:

- Условие задания должно быть ясным и недвусмысленным. Необходимо избегать использования неточных выражений типа “сопровождается”, “полезно для”, “важно для” и слов, являющихся подсказкой - “может” или “возможно”. Неопределенные термины типа “обычно” или “часто” не должны применяться.
- Варианты ответов должны быть либо абсолютно верными, либо абсолютно неверными; никакие оттенки недопустимы. Исключайте использование фраз и слов, перечисленных выше.

КАТЕГОРИЯ ВОПРОСОВ С ОДНИМ ЛУЧШИМ ОТВЕТОМ

В отличие от вопросов формата “верно/неверно”, вопросы с одним лучшим ответом (А-тип) точно определяют число вариантов ответа, которое нужно выбрать. Вопросы с одним лучшим ответом (А-тип) являются наиболее широко и часто используемым форматом вопросов множественного выбора. Они состоят из условия задания (клинической ситуации), вводного вопроса и серии из 5 вариантов ответа (обычно одного верного и четырех дистракторов - отвлекающих ответов). В следующем примере дано описание больного и экзаменуемого просят поставить наиболее вероятный диагноз.

Условие задания:

32-х летний мужчина в течение 4-х дней отмечает нарастающую слабость в конечностях. Был ранее здоров, но перенес респираторную инфекцию 10 дней назад. Температура 37,8°C, АД 130/80 мм.рт.ст, пульс 94/мин, ЧДД 42/мин. Дыхание поверхностное. У больного наблюдается симметричная слабость мышц обеих половин лица и проксимальных и дистальных мышц конечностей. Чувствительность сохранена. Глубокие сухожильные рефлексы не вызываются; отмечаются сгибательные подошвенные рефлексы.

Вводный вопрос:

Какой из нижеперечисленных диагнозов является наиболее вероятным?

Варианты ответа:

- A. Острый диссеминированный энцефаломиелит
- B. Синдром Гийена-Барре
- C. Миастения гравис
- D. Полиомиелит
- E. Полимиозит

Заметьте, что неверные варианты ответа не являются абсолютно неверными. Варианты ответа можно представить следующим образом:

D	C	A	E	B
<i>наименее верный</i>			<i>наиболее верный</i>	

Несмотря на то, что неверные варианты ответа не являются абсолютно неверными, они все же менее правильны, чем эталонный ответ. Экзаменуемый должен выбрать “наиболее вероятный диагноз”. Все эксперты согласятся, что в данном случае наиболее вероятным диагнозом является ответ В. Также они согласятся, что остальные диагнозы тоже вероятны в какой-то степени, но менее, чем ответ В. Когда варианты ответа представляют собой однородную совокупность (в данном случае от “Наиболее Вероятного” до “Наименее Вероятного” диагноза), эти варианты не должны быть абсолютно неверными.

Данное тестовое задание имеет дефект. После чтения условия задания студент имеет совершенно неясное представление о сути вопроса. Чтобы найти “наилучший” ответ экзаменуемый должен решить, является ли вариант ответа А (“Чаще болеют женщины”) более или менее верным, чем вариант ответа В (“Редко встречается острая боль в суставе”). Это сравнение яблок и апельсинов. Чтобы разграничить относительную правильность вариантов ответа они должны быть однородны (гомогенны), в противном случае все варианты ответа должны быть стопроцентно верны или неверны.

Что из ниже перечисленного верно в отношении псевдоподагры?

- A. Чаще встречается у женщин
- B. Редко сопровождается острой болью в суставе
- C. Может сопровождаться хондрокальцинозом
- D. В большинстве случаев наследственное заболевание
- E. Хорошо поддается лечению аллопуринолом

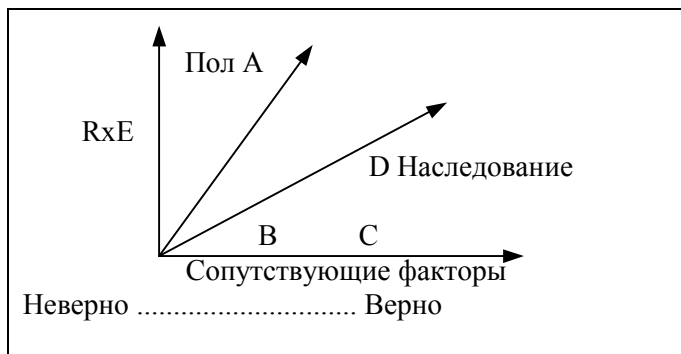


Диаграмма данных вариантов ответа может выглядеть подобным образом. Варианты ответа гетерогенны и затрагивают различные факты; они не могут быть расположены от наименее к наиболее верному в одном измерении. Хотя данный вопрос призван оценить знания различных сторон заболевания, присущие тестовому заданию дефекты не позволяют это сделать. Вопрос не ясен сам по себе; на него невозможно ответить, не видя вариантов ответа.

В отличие от вариантов ответа в тестовом задании о псевдоподагре, варианты ответа в тестовом задании о синдроме Гийена-Барре гомогенны (все они являются диагнозами); знающие экзаменуемые могут расположить варианты ответа в одном измерении.

Хорошо построенные вопросы с одним наилучшим ответом удовлетворяют правилу “закрытых вариантов ответа” - возможности ответить на вопрос, не видя вариантов ответа. Такие вопросы могут использоваться для письменного экзамена.

ИТОГ ОБСУЖДЕНИЯ ФОРМАТОВ ТЕСТОВЫХ ЗАДАНИЙ

Мы рекомендуем вам не использовать тестовые задания формата “Верно/Неверно”. Хотя многие авторы ТЗ считают, что тестовые задания данного формата писать легче, чем задания с одним лучшим ответом, мы обнаружили, что с ними больше проблем. При написании вопроса автор имеет в виду что-то определенное, а тщательный разбор написанных тестовых заданий показывает, что в них есть некоторые недостатки, незаметные автору. Часто различия между “верным” и “неверным” неясны, и не были редкостью ситуации, когда при последующем пересмотре менялся эталон ответа. В результате, при обсуждении тестовых заданий формата “Верно/Неверно” на встречах тестовых комитетов их приходится переписывать и исключать гораздо чаще, чем тестовые задания, написанные в других форматах. В некоторых случаях двусмысленность можно устранить, в других это сделать нельзя.

И, наконец, есть еще одна причина, гораздо более серьезная, чем вышеперечисленные. Мы обнаружили, что, исключая двусмысленность, мы вынуждены оценивать только запоминание отдельных фактов, а это то, чего мы активно избегаем. Мы нашли, что применение знаний, интеграция, синтез и умение принимать решение лучше оцениваются с помощью вопросов с одним лучшим ответом. NBME полностью прекратил использование тестовых заданий формата “Верно/Неверно” в своих экзаменах.

Мы также рекомендуем вам не использовать А-тип тестовых заданий, содержащих отрицание. Наиболее проблемными являются тестовые задания, содержащие формулировки типа “Что из следующего верно, **КРОМЕ**“ и “Какое из следующих утверждений **НЕ** верно?” Такие вопросы имеют те же проблемы, что и формат “Верно/Неверно”: если варианты ответа невозможно расположить в одном измерении, экзаменуемые не могут определить ни “наименее”, ни “наиболее” правильный ответ. С другой стороны, в некоторых экзаменах мы иногда используем четко сфокусированные отрицательные тестовые задания А-типа с короткими (одно слово) вариантами ответа. Главным образом они выступают в качестве достаточно плохой замены тех тестовых заданий, которые требуют от экзаменуемого выбрать несколько ответов. Оптимальный формат тестовых заданий, который может использоваться для этой цели, - формат N-типа, при котором экзаменуемые должны выбрать “N” ответов, - обсуждается далее в этом руководстве.

В Приложении даны примеры форматов тестовых заданий, которые сейчас не используются в экзаменах NBME.

ГЛАВА 3

ТЕХНИЧЕСКИЕ ДЕФЕКТЫ ТЕСТОВЫХ ЗАДАНИЙ



Данный раздел описывает два типа технических дефектов тестовых заданий: связанных с опытом студентов в тестировании и безотносительной сложностью. Наличие дефектов первого типа облегчает некоторым студентам выбор правильных ответов. При этом студенты основываются только на своем опыте в тестировании. Данные дефекты обычно наблюдаются в тестовых заданиях, которые не сфокусированы и не удовлетворяют правилу “закрытых вариантов ответа”. При наличии дефекта в виде безотносительной сложности, вопрос становится трудным вне связи с оцениваемым параметром.

Целью данного раздела является представить типичные дефекты и помочь преподавателям исключить данные дефекты из своих вопросов, чтобы обеспечить равные возможности как для умудренных, так и для менее опытных в тестировании студентов. Вероятность правильного ответа на вопрос должна зависеть от уровня знаний экзаменуемого по оцениваемой теме и не должна зависеть от опыта студентов в тестировании.

ДЕФЕКТЫ, СВЯЗАННЫЕ С ОПЫТОМ В ТЕСТИРОВАНИИ

Грамматические подсказки: Один или более дистракторов грамматически не соответствуют условию тестового задания.

Так как авторы тестового задания стараются уделить большее внимание правильному ответу, чем дистракторам (отвлекающим ответам), грамматические несоответствия чаще встречаются в последних. В данном примере опытные в тестировании студенты сразу исключают А и С варианты, так как они грамматически не совпадают с условием задания. Далее таким студентам останется выбрать только между вариантами ответа В, D и Е.

60-летний алкоголик доставлен полицией в приемный покой в состоянии эпилептического статуса. После того, как вы убедились, что его дыхательные пути свободны, первым шагом в дальнейшем ведении пациента будет внутривенное введение:

- A. исследование спинномозговой жидкости*
- B. глюкозы с витамином В₁ (тиамин)*
- C. компьютерная томография головы*
- D. фенитоина*
- E. диазепама*

Логические подсказки: часть вариантов ответа исчерпывает все возможные варианты

В данном тестовом задании варианты ответа А, В и С включают все возможности. Опытный в тестировании студент знает, что ответы А, В или С должны быть правильными, тогда как неопытный теряет время на обдумывание ответов D и E. Часто авторы тестовых заданий добавляют ответы D и E только для доведения числа вариантов до 5. В таких ситуациях авторы тестовых заданий могут не обратить внимания на содержание ответов D и E; иногда они могут быть частично верными и смущать студентов, так как не могут быть расположены в одном измерении с ответами А, В и С. Данные дефекты часто встречаются в тестовых заданиях, содержащих варианты ответа: “Повышается”, “Понижается” и “Остается на прежнем уровне”.

Преступность

- A. одинаково распространена среди всех социальных слоев*
- B. наиболее широко представлена среди бедного населения*
- C. наиболее широко представлена среди людей среднего достатка и богатых*
- D. главным образом указывает на психо-сексуальную дизадаптацию*
- E. достигает уровня терпимости общества*

Абсолютные термины: использование таких терминов, как “всегда” или “никогда” в вариантах ответа.

В данном тестовом задании варианты ответа А, В и Е содержат термины, которые менее категоричны, чем содержащиеся в вариантах С и D. Опытный в тестировании студент исключит варианты С и D из возможных, так как они менее вероятны из-за абсолютности утверждений. Заметьте, что данный дефект не появился бы, если бы условие задания было сфокусировано, а варианты ответа короткими; дефект возникает тогда, когда глагол включается в варианты ответа, а не во вводный вопрос.

У больных с прогрессирующей деменцией типа Альцгеймера дефект памяти

- A. можно лечить фосфатидилхолином (лецитином)*
- B. может быть результатом раннего паркинсонизма*
- C. никогда не встречается у пациентов с нейрофиброматозом на аутопсии*
- D. никогда не бывает глубоким*
- E. возможно вовлекает холинэргическую систему*

Длинный правильный ответ: правильный ответ длиннее, более конкретен или более полон, чем другие варианты ответа.

В данном тестовом задании вариант ответа С длиннее, чем все остальные варианты ответа; также это единственный двойной ответ. Авторы тестовых заданий уделяют большее внимание правильному ответу, чем дистракторам. Поскольку вы преподаватель, то вы пишете длинные правильные ответы, содержащие дополнительную информацию, комментарии, предупреждения и т.д. Иногда это бывает слишком выражено: правильный ответ длиной в абзац, а дистракторы - только одно слово.

“Вторичное приобретение“ (получение определенных преимуществ и выгод от болезни)-

- A. это синоним симуляции
- B. это частая проблема при обсессивном синдроме
- C. это осложнение множества заболеваний, ведущее к удлинению болезни.
- D. не встречается при органическом поражении мозга

Повторение слов: слово или фраза из условия задания включены в правильный ответ.

В данном тестовом задании используется слово “нереальный” в условии задания и слово “дереализация” в правильном ответе. Иногда слово повторяется только метафорически: условие задания упоминает костные боли, а правильный ответе начинается со слова osteo.

58-летний мужчина, злоупотребляющий алкоголем и ранее госпитализировавшийся в психиатрическую больницу, растерян и возбужден. Он воспринимает окружающий мир, как нереальный. Этот симптом называется

- A. дереализация
- B. деперсонализация
- C. дезориентация
- D. частичная потеря памяти
- E. сигнальное возбуждение

Тенденция к конвергенции: правильный ответ включает в себя наибольшее количество общих элементов из других вариантов ответа.

Дефект этого тестового задания менее заметен, чем в предыдущих примерах, но он часто встречается и заслуживает внимания. Дефект может быть нескольких видов. Основополагающим является то, что правильным ответом служит вариант, имеющей наибольшее сходство с остальными вариантами ответа; для правильного ответа не типично быть исключением. Например, в вариантах ответа, представленных цифрами, правильным ответом скорее всего будет среднее значение, а не крайние величины. В двойных вариантах правильным ответом скорее всего будет вариант, имеющий наибольшее сходство с дистракторами. Например, среди вариантов ответа “Карандаш и ручка”, “Карандаш и фломастер”, “Карандаш и кисточка”, “Ручка и маркер”, правильным ответом скорее всего будет “Карандаш и ручка” (что выявляется путем простого подсчета: “Карандаш” встречается 3 раза, “Ручка” - дважды, а другие варианты лишь по одному разу). Хотя это может показаться забавно, но данные дефекты возникают из-за того, что авторы тестовых заданий начинают с правильного ответа, а вариации с правильным ответом являются дистракторами. Таким образом правильный ответ и содержит наибольшее число общих элементов среди всех вариантов ответа; неправильные ответы чаще являются исключе-

ниями, так как авторам тестовых заданий бывает трудно создать правдоподобные дистракторы. В приведенном примере опытный в тестировании студент исключит “анионную форму” как маловероятную; этот студент также исключит “снаружи нервной мембраны”. Таким образом студенту останется выбрать только между вариантами В и Е. Поскольку три из пяти вариантов ответа упоминают заряд, такой опытный студент выберет вариант В.

Местные анестетики наиболее эффективны в

- A. анионной форме, действуя изнутри нервной мембраны*
- B. катионной форме, действуя изнутри нервной мембраны*
- C. катионной форме, действуя снаружи нервной мембраны*
- D. незаряженной форме, действуя изнутри нервной мембраны*
незаряженной форме, действуя снаружи нервной мембраны

ДЕФЕКТЫ, СВЯЗАННЫЕ С ЧРЕЗМЕРНОЙ СЛОЖНОСТЬЮ

Варианты ответов являются длинными, сложными или двойными.

Данное тестовое задание иллюстрирует собой типичный дефект. Условие задания содержит посторонние сведения, и, что еще более важно, варианты ответов очень длинные и сложны. Выбор из данных вариантов ответа требует долгого и внимательного чтения в связи с большим количеством подробностей в каждом варианте. Это может привести к тому, что вместо оценки знаний мы будем оценивать скорость

чтения. Заметьте, что данный дефект относится только к вариантам ответа. Существует много хорошо составленных тестовых вопросов, которые имеют длинное условие задания. Решение о размере условия задания должно приниматься в соответствии с целью тестового задания. Если его целью является оценка способности студента интерпретировать и синтезировать информацию, например, для постановки наиболее вероятного диагноза, тогда целесообразно, чтобы условие задания включало подробное описание пациента.

Комиссии по оценке работы коллег в некоторых организациях здравоохранения (НМО - Health Maintenance Organization) могут предпринять действия в отношении документов, подтверждающих образование и квалификацию врача, в защиту интересов участников НМО. Существует требование убедиться в том, что в ходе таких действий к врачу применяются надлежащие законные процедуры. Что из следующего обязательно включается в надлежащие законные процедуры?

A. Соответствующее уведомление, суд, уполномоченный принять решение, возможность лично оспорить показания свидетелей против себя; и возможность представить доказательства в свою защиту.

B. Уведомление, беспристрастность участников, возможность услышать и лично оспорить показания свидетелей против себя.

C. Разумное и своевременное уведомление, беспристрастная комиссия, уполномоченная принять решение, возможность услышать свидетельства против себя и лично оспорить показания свидетелей и возможность представить доказательства в свою защиту.

Цифровые данные выражены бессистемно

Когда используются цифровые варианты ответа, они должны быть расположены в порядке возрастания и быть выражены в едином формате (т.е. в виде отдельного числа или в виде диапазона чисел). Путаница возникает, когда форматы цифровых данных смешиваются, и когда варианты ответа располагаются в нелогичном порядке или бессистемно.

В данном примере варианты ответа А, В и С выражены в виде диапазона, тогда как варианты D и E - точных процентные данные. Все варианты ответа должны быть либо в виде диапазона, либо в виде точных процентных данных; смешивать эти форматы не рекомендуется. Кроме того, вариант ответа С включает в себя D и E, что почти несомненно исключает варианты D и E из числа правильных ответов.

После повторно перенесенного сальпингита какова вероятность бесплодия у женщины?

- A. Меньше 20%*
- B. От 20 до 30%*
- C. Более 50%*
- D. 90%*
- E. 75%.*

Использование в вариантах ответа неопределенных терминов, характеризующих частоту (напр., редко, обычно)

Исследование показало, что неопределенные термины, характеризующие частоту явления, даже экспертами не всегда понимаются одинаково. Ниже в руководстве приведен более подробный анализ данного исследования.

Выраженное ожирение в раннем подростковом возрасте

- A. обычно лечение диетой является эффективным*
- B. часто связано с эндокринными нарушениями*
- C. в 75% случаев исчезает спонтанно*
- D. имеет плохой прогноз*
- E. обычно поддается фармакотерапии и интенсивной психотерапии*

Стилистическая неоднородность вариантов ответа; нелогичный порядок вариантов ответа.

Данное тестовое задание иллюстрирует распространенный дефект. При этом варианты ответа длинные, а стилистика делает тестовое задание трудным и требующим времени для принятия решения о наиболее правильном ответе. Обычно данный дефект можно устранить при внимательном редактировании. В данном конкретном тестовом задании вводный вопрос можно изменить следующим образом: “По какой из следующих причин невозможно сделать выводы из представленных данных?”. Варианты ответа можно отредактировать следующим образом: А. “Не было наблюдения за невакцинированными детьми”; В. “Количество наблюдений слишком мало”; С. “Исследование включает только мальчиков”, а для D может быть написан новый вариант ответа.

При исследовании эффективности вакцины двести 2-летних мальчиков получили вакцину против определенного заболевания, а затем, в течение 5 лет за ними проводилось наблюдение для выявления частоты данного заболевания. 85% детей этой группы не имели признаков заболевания. Какое из следующих утверждений о данных результатах верно?

- A. Заключение нельзя сделать, так как не было наблюдения за невакцинированными детьми*
- B. Число случаев (30 за 5 лет) слишком мало для достоверного статистического заключения*
- C. Заключение нельзя сделать, так как в исследовании были включены только мальчики*
- D. Эффективность вакцины (%) вычислена как 85-15/100*

Использование фразы “Ничего из вышеперечисленного” в качестве варианта ответа.

Фраза “Ничего из вышеперечисленного” является проблемной для тестовых заданий, где требуется принять решение, а варианты ответа не являются абсолютно верными или абсолютно неверными. Если предполагаемый ответ заключен в одном из перечисленных вариантов ответа, то знающий студент сталкивается с дилеммой, поскольку он должен выбрать между очень подробным хорошим ответом и вашим собственным правильным ответом. Студент сам может придумать вариант ответа, который будет более правильным, чем предложенный вами. Использование фразы “ничего из вышеперечисленного” полностью изменяет тестовое задание на формат Верно/неверно; каждый вариант ответа должен оцениваться как более или менее верный по сравнению с множеством неуказанных ответов.

Диагноз большой яичниковой кисты наиболее вероятен в случае

- A. притупления по передней брюшной стенке, латерального тимпанита
- B. ослабления перистальтики
- C. флюктуации
- D. перемещающегося притупления
- E. ничего из вышеперечисленного

Условия задания излишне сложны или запутаны

Иногда авторы тестовых заданий могут превратить достаточно легкий вопрос в нечто настолько запутанное, что лишь очень отважный студент решится прочитать его. Данное тестовое задание является таким примером.

Расположите родителей следующих детей с синдромом Дауна в порядке от наибольшего к наименьшему риску повторного рождения ребенка с данным заболеванием. Условно примите, что возраст матери во всех случаях 22 года, а следующая беременность наступит в течение 5 лет. Кариотипы дочерей следующие:

- I. 46, XX, -14, +T (14q21q) по отцу
- II. 46, XX, -14, +T (14q21q) впервые
- III. 46, XX, -14, +T (14q21q) по матери
- IV. 46, XX, -21, +T (14q21q) по отцу
- V. 47, XX, -21, +T (21q21q) (кариотип родителей неизвестен)

- A. III, IV, I, V, II
- B. IV, III, V, I, II
- C. III, I, IV, V, II
- D. IV, III, I, V, II
- E. III, IV, I, II, V

ОБОБЩЕНИЕ ТЕХНИЧЕСКИХ ДЕФЕКТОВ ТЕСТОВЫХ ЗАДАНИЙ

Дефекты, относящиеся к опытности в тестировании

- **Грамматические подсказки** - один или более дистракторов грамматически не соответствуют условию задания
- **Логические подсказки** - часть вариантов ответа исчерпывает все возможные варианты
- **Абсолютные термины** - использование терминов “всегда” или “никогда” в некоторых вариантах ответа
- **Длинный правильный ответ** - правильный ответ длиннее, более конкретен или более полон, чем другие варианты ответа.
- **Повторение слов** - слово или фраза из условия задания повторяется в правильном варианте ответа
- **Тенденция к конвергенции** - правильный ответ имеет наибольшее сходство с элементами других вариантов ответа

Дефекты, относящиеся к чрезмерной сложности

- Варианты ответа являются длинными, сложными или двойными
- Цифровые данные выражены бессистемно
- Использование неопределенных терминов в вариантах ответа (напр., “редко”, “обычно”)
- Стилистическая неоднородность вариантов ответа
- Нелогичный порядок вариантов ответа
- Использование фразы “Ничего из вышеперечисленного” в качестве варианта ответа
- Условия задания слишком сложны или запутаны
- Ответ на одно тестовое задание вытекает из ответа на предыдущее тестовое задание

Общие указания по составлению тестовых заданий

- Убедитесь, что на тестовое задание можно ответить, не глядя на варианты ответа, ИЛИ что варианты ответа на 100% верны или неверны.
- Включите в условие наибольшую часть информации из тестового задания; условие задания должно быть длинным, а варианты ответа короткими.
- Избегайте избыточной информации.
- Избегайте “запутанных” или слишком сложных тестовых заданий.
- Составляйте грамматически правильные и логически вытекающие из условия задания варианты ответа; размещайте их в логическом или алфавитном порядке. Пишите правдоподобные дистракторы, имеющие приблизительно ту же длину, что и правильный ответ.
- Избегайте использования в вариантах ответа категоричных утверждений, типа *всегда*, *никогда* и *все*; также избегайте неопределенных формулировок, типа *обычно* или *часто*.

- Избегайте тестовых заданий с отрицаниями (со словами **кроме** или **не** во вводном вопросе). Если использование отрицания в условии задания необходимо, варианты ответа должны быть короткими, предпочтительно в одно слово.

И самое главное: Сосредоточьте внимание на важных моментах; не теряйте время на проверку знания тривиальных фактов.

ИСПОЛЬЗОВАНИЕ НЕОПРЕДЕЛЕННЫХ ТЕРМИНОВ В ЭКЗАМЕНАЦИОННЫХ ВОПРОСАХ

Хотя мы используем неопределенные термины в повседневной речи и при письме, они могут вызывать затруднения при использовании в экзаменационных тестовых заданиях. Во время исследования, проведенного в Национальном Совете Медицинских Экзаменаторов (NBME), 60 членов восьми тестовых комитетов, составляющие вопросы для экзаменов по различным медицинским специальностям, проанализировали список терминов, используемых в вопросах множественного выбора для выражения частоты явления, и попытались определить отрезок времени, который отражается каждым из этих терминов.

Результаты, приведенные ниже, показали, что эти термины не имеют устойчивого определения, которое было бы общепринято, даже среди самих авторов тестовых заданий. Средняя величина плюс-минус одно стандартное отклонение превышала 50% для более чем половины фраз. Например, в среднем, авторы тестовых заданий подразумевали под термином “*часто*” 65% встречаемости; две трети полагали, что это значение находится между 45-85%; а в действительности разброс мнений составил 15-95%. Особо следует заметить, что величины для термина **часто** перекрывались с величинами, характеризующими термин **редко**.

Влияние данных фактов варьирует в зависимости от формата тестового задания. Неопределенные формулировки вызывают более серьезные проблемы в тестовых заданиях формата верно/неверно (К-, С- и Х-типы), чем в тестовых заданиях с одним лучшим ответом. Например, неопределенные термины являются основной проблемой в следующем тестовом задании формата верно/неверно:

К правильным утверждениям о псевдоподагре относятся:

1. *Часто встречается у женщин*
2. *Часто связана с острой болью*
3. *Обычно передается по наследству*
4. *Уровень кальция в сыворотке часто повышен.*

В тестовом задании формата верно/неверно экзаменуемый должен решить, является ли каждый вариант ответа верным или неверным. Когда варианты ответа не являются абсолютно верными или неверными, экзаменуемые полагаются на свою собственную интерпретацию двусмысленных терминов или пытаются угадать, что же подразумевал под данным термином автор тестового задания. С другой стороны, ответы экзаменуемых могут отражать их личный стиль ответов (тенденцию предпочитать ответы “

верно” или “неверно”, если правильный ответ неизвестен). Данный стиль ответов может сильнее влиять на правильность ответов, чем знание экзаменуемым предмета. Это может быть одной из причин, почему тестовые задания формата верно/неверно имеют плохие результаты.

Перефразирование вариантов ответа с указанием конкретных цифр не улучшает ситуацию. Например, в приведенном примере изменение варианта ответа №1 следующим образом: “частота встречаемости среди женщин составляет 1:2000” не будет приемлемой модификацией. Частота встречаемости не составляет точно 1:2000, и поскольку диапазон не указан, экзаменуемые будут определять свой собственный диапазон: уже или шире, в зависимости от стиля ответов экзаменуемого. В тестовых заданиях формата верно/неверно приемлемым решением в случае с цифровыми вариантами ответа является или проведение сравнения (например, частота встречаемости выше, чем у остеоартроза), или указание диапазона (например, частота встречаемости находится между 1:1000 и 1:2000).

Проблемы, характерные для тестовых заданий формата верно/неверно, обычно не возникают в хорошо сформулированных тестовых заданиях с одним лучшим ответом (т.е., имеющих ясный вопрос и однородные варианты ответа). Например, следующий вопрос содержит неопределенный термин в условии тестового задания, но так как задание состоит в выборе одного наилучшего ответа, вопрос не является двусмысленным.

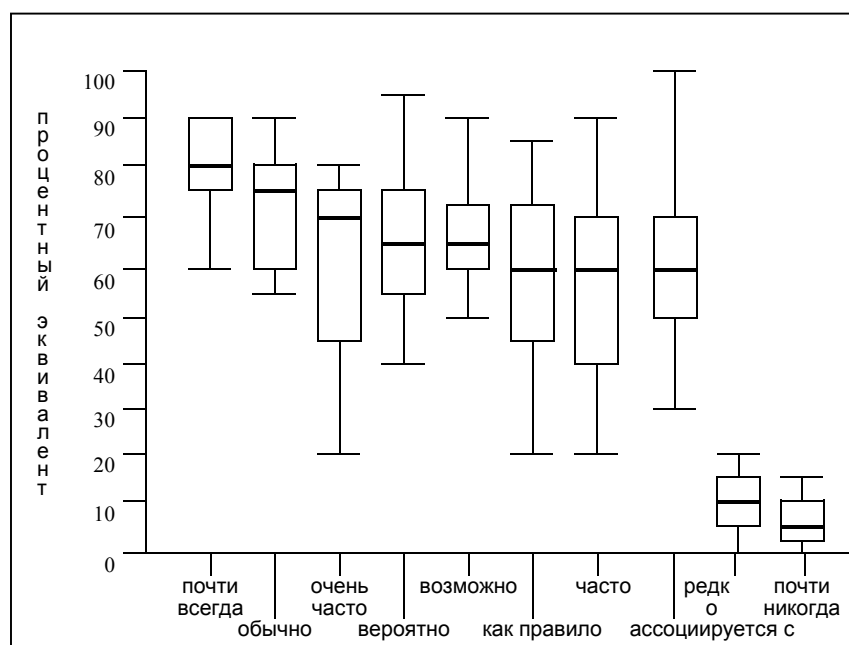
Какой из перечисленных лабораторных показателей обычно повышается у больных псевдоподагрой?

Однако при использовании неопределенных формулировок в вариантах ответа проблемы возникают и в тестовых заданиях с одним лучшим ответом, как в приведенном примере.

Единственный путь еще больше запутать данное тестовое задание - это добавить пятый вариант ответа- “ничего из вышеперечисленного”.

Больные псевдоподагрой испытывают боль

- A. Очень часто*
- B. Обычно*
- C. Часто*
- D. Довольно часто*



Данная диаграмма показывает распределение ответов по терминам, характеризующим частоту. Эти данные основаны на ответах 60 членов восьми тестовых комитетов. Горизонтальная линия в каждой клетке - средняя величина; клетки включают диапазон 50% ответов. Вертикальные линии продлеваются до наибольшей и наименьшей из указанных величин. Например, средним ответом для термина “часто” является 70%; половина опрошенных полагала, что “часто” находится между 45% и 75%; а диапазон всех ответов варьировал от 20% до 80%.

По материалам: Case SM. The use of imprecise terms in examination questions: how frequent is frequently? *Academic Medicine*. 1994;69 (suppl):S4-S6.