

Основные этапы разработки и экспертизы  
тестовых заданий. Содержательная и  
тестологическая экспертиза БТЗ. Эмпирическая  
апробация теста. Дистракторный анализ и  
распределение опций

Доц. Баскакова И.В.

Семинар. Дни 3-4

**По итогам  
семинара  
Вы будете  
способны:**

**РО - 1** Рационально планировать процесс разработки заданий в тестовой форме

**РО - 2** Проводить содержательную экспертизу БТЗ

**РО - 3** Проводить первичную тестологическую экспертизу БТЗ

**РО- 4** Оценивать качество заданий в тестовой форме, используя стандартные тестологические параметры (уровень сложности, ID, распределение опций).

# Этапы разработки теста

1. **Определение целей тестирования**
2. **Определение ресурсных возможностей разработчиков теста**
3. **Отбор содержания учебного материала**
4. **Конструирование технологической матрицы / спецификации теста и ее экспертиза**
5. **Составление заданий в тестовой форме и их экспертиза (содержательная экспертиза)**
6. **Эмпирическая апробация заданий в тестовой форме (пре-тест)**

# Этапы разработки теста

7. Оценка качества тестовых заданий по результатам апробации (тестологическая экспертиза-1)

8. Отбраковка дефектных заданий и составление теста

9. Оснащение теста невербальными материалами (картинки, схемы, графики и др.)

10. Апробация теста

11. Тестологическая экспертиза - 2

(распределение опций, дистракторный анализ, индекс дискриминации)

12. Содержательная экспертиза -2

# Создание матрицы теста

Матрица теста может  
носить разные названия:

- технологическая матрица;
- содержательная решетка;
- сетка;
- организационно-деятельностная матрица;
- тестовая решетка;
- таблица содержания и др.



???



Для чего нужна  
спецификация/  
матрица теста?

# Для чего нужна матрица/спецификация теста:

Содержание дисциплины, которое будет отобрано для проверки.

Важность того или иного элемента содержания дисциплины.

Вся ли дисциплина охвачена предполагаемыми вопросами?

Содержание дисциплины должно полностью покрываться матрицей по всем темам / клиническим проблемам / результатам обучения по дисциплине / ОП, в зависимости от целей тестирования.

Фиксируется требование широты теста, полного учета всех разделов дисциплины, который находит выражение в матрице.



# Проблемы:

**Главной проблемой в разработке тестов, используемых для оценки обучающихся и образовательных программ, является соответствия содержания программы и уровня обучения содержанию теста.**

**Технологическая матрица может описывать только предметы, предметные области или отдельные темы разного уровня обобщения, которые должны войти в тест, и определяет соотношение заданий в тесте.**

**Важно проследить, чтобы общее количество вопросов было распределено по темам таким образом, чтобы оно отражало их важность, объем и количество учебного времени.**



# ПРОБЛЕМЫ:

Зачастую **выборка тестовых вопросов** носит субъективный характер и **нерепрезентативна** в отношении учебных целей.

**Отсутствие соответствующего программного обеспечения** не позволяет проводить оценку качества разработанных тестовых заданий и улучшать их .

## Распределение заданий между темами

Распределение заданий между темами следует увязать с целями предстоящего тестирования.

В том случае, если речь идет о **промежуточной аттестации**, следует преимущественно ориентироваться на время (объем) темы.

В случае, если тесты готовятся **для итоговой аттестации**, следует ориентироваться на важность темы для дальнейшего образования (практической деятельности).



# 1. Основные категории учебных целей (по Блуму).

## Знание (запоминание, припоминание, воспроизведение)

Эта категория означает запоминание и воспроизведение ранее изученного материала.

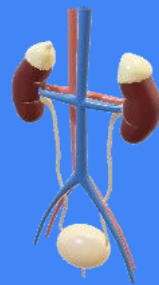
Общая черта этой категории – припоминание соответствующих сведений.

– **Знает (помнит)** –

- употребляемые термины, отдельные симптомы, классификации, теории;
  - основные понятия, конкретные факты;
  - показатели, методы, процедуры, протоколы;
  - правила, принципы и др.



## 2. Понимание



Интерпретация студентом (объяснение, краткое изложение) или же предположение о дальнейшем ходе явлений, событий (предсказание последствий, результатов).

**Такие учебные результаты превосходят простое запоминание!**

- понимает факты, правила и приемы;
- интерпретирует словесный материал (клинические ситуации);
- интерпретирует схемы, графики, диаграммы, рентгенограммы, результаты анализов:
  - преобразует словесный материал в другое выражение;
- выставляет диагноз, предположительно описывает прошедшие или будущие последствия, вытекающие из имеющихся данных и др.



## 3. Применение

- Эта категория обозначает умение использовать изученный материал в конкретных условиях и новых ситуациях.
- **Применяет** правила, законы, алгоритмы, протоколы, теории в конкретных практических ситуациях;
- **Принимает решение** о выборе тактики обследования, лечения, наблюдения, выборе лекарственного препарата и т.д.
- Демонстрирует правильное применение метода или процедуры и др. (последовательность действий).

## 4. Анализ



Эта категория обозначает умение разбирать материал на составляющие так, чтобы ясно выступала его структура.

Сюда относится вычленение частей целого, выявление взаимосвязей между ними, осознание принципов организации целого.

Учебные результаты характеризуются более высоким интеллектуальным уровнем, чем понимание и применение, поскольку требуют осознания как содержания учебного материала, так и его внутреннего строения.

- выделяет скрытые (неявные) предположения;
- видит ошибки и упущения в логике рассуждений;
- проводит различия между фактами и следствиями;
- оценивает значимость данных и др.

# 5. Синтез



Эта категория обозначает умение комбинировать элементы, чтобы получить целое, обладающее новизной (или заданными свойствами).

**Соответствующие учебные результаты предполагают деятельность творческого характера с акцентом на создание новых схем и структур.**

- пишет небольшие творческие сочинения;
- предлагает план проведения эксперимента;
- использует знания из различных областей, чтобы составить план решения той или иной проблемы и др.

# 6. Оценка



**Данная категория предполагает достижение учебных результатов по всем предшествующим категориям плюс оценочные суждения на ясно очерченных категориях.**

- оценивает логику построения материала в виде письменного текста;**
  - оценивает соответствие выводов имеющимся данным;**
  - оценивает значимость того или иного продукта деятельности, исходя из внутренних критериев;**
- оценивает значимость того или иного продукта деятельности, исходя из внешних критериев и др.**



# Когнитивные уровни

по Аванесову В.С.

- Знание названий, имен и смысла названий и имен
  - Знание определений
  - Фактуальные знания
  - Классификационные знания
  - Ассоциативные знания
- ↔
- Сравнительные сопоставительные знания
  - Знание противоположностей, противоречий
  - Причинные знания, причинно-следственных отношений, знание оснований, принципов классификации
- ↔
- Алгоритмические знания
  - Процедурные знания
  - Технологические знания
  - Структурные знания
- ↔
- Обобщенные, системные знания
  - Оценочные знания
  - Вероятностные знания
  - Процессуальные знания
  - Абстрактные знания

по Блему Б.

Запоминание

Понимание

Применение

Анализ

Синтез

Оценка



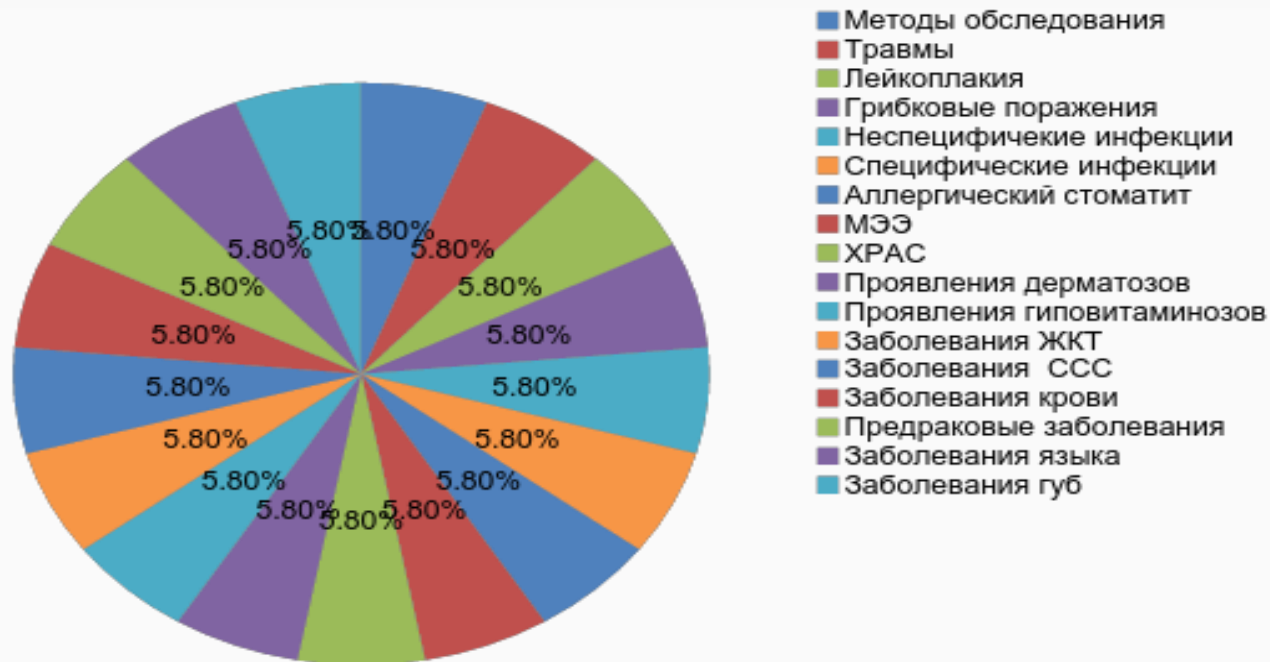
# ВЫБОР МЕТОДОВ ОЦЕНИВАНИЯ

## Методы оценки должны соответствовать результатам обучения

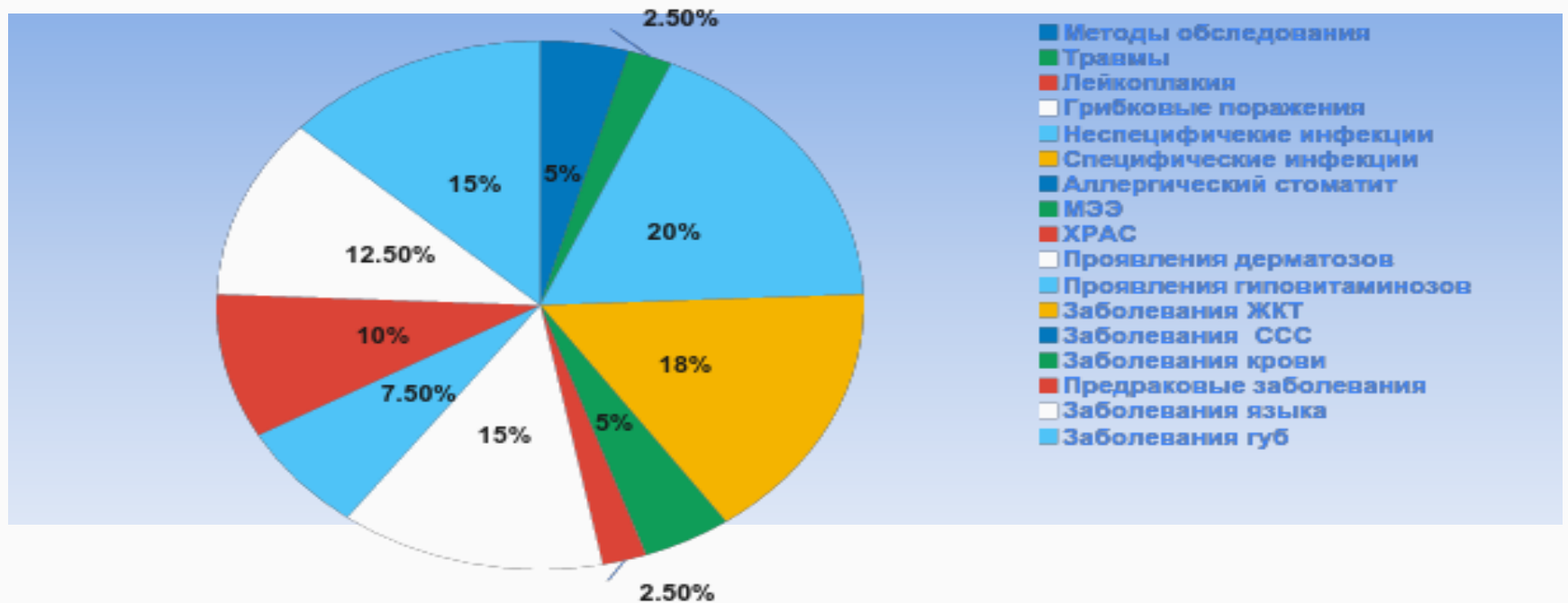


Miller GE. 1990 The assessment of clinical skills / competence / performance. Academic Medicine, 65(Supplement): p. S63-S67) *Women's health Care Competencies for Medical Students: Taking Steps to Include Sex and Gender Differences in the Curriculum*, 2005.

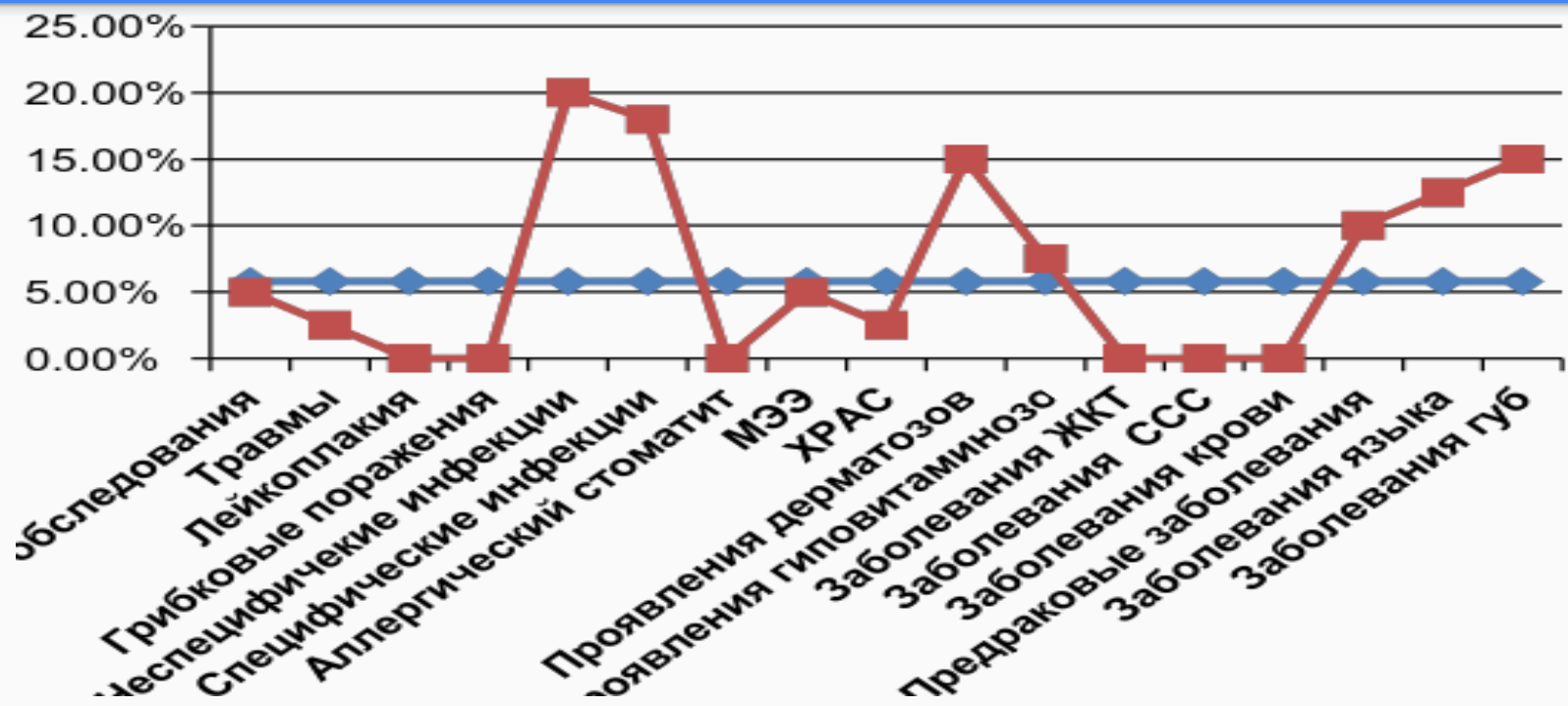
# Распределение тем внутри дисциплины «Заболевания слизистой оболочки рта и красной каймы губ» (пример идеального распределения удельного веса ТЗ по темам)



# Распределение заданий в экзаменационном тесте студента, подавшего заявление на апелляцию



# Распределение тем в реальном экзаменационном тесте студента (на примере поданного заявления на апелляцию)



# Предоставление заданий в невербальной форме

Существует множество способов представления заданий в невербальной форме или же в вербально-невербальной.

В таких формах могут быть представлены не только сами задания, но и варианты ответов.

Вариативность форм представления заданий является мощным средством сделать задания разнообразными, значительно менее утомительными и даже интересными для испытуемых.

Возможности тестирования некоторых специальных умений, например, умения работать с результатами исследований, понимать и интерпретировать их

Вопрос или задание могут быть сформулированы в виде фразы, текста, рисунка, схемы, цифры, символа, графика, формулы и пр.

# Значение невербальных материалов

Часть содержания образования представляет собой невербальные материалы.  
Использование этих материалов в соответствующих тестах должно быть обязательным.

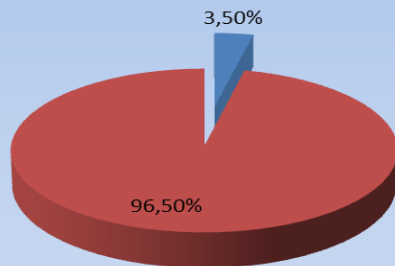
В том случае, если эти материалы не используются, есть значительная опасность деформации содержания образования.

Применение этих материалов делает тестирование менее утомительным, более разнообразным, снижает количество случайных ошибок.

Рисунки служат дополнительным мотивом при выполнении тестирования, вызывая интерес к заданию, дают возможность оценить различный уровень развития абстрактного мышления.

Наличие вербальной и невербальной форм задания одновременно снимает эту разницу в развитии.

# Различные дефекты ТМ в экзаменационных БТЗ встречались с частотой от 22 до 45%



■ БТЗ с минимальным количеством замечаний

■ БТЗ с дефектами

- повтор слов во вводном вопросе и в вариантах ответов
- словесная «мишура»
- громоздкие условия задания
- задания с коротким условием и очень длинными ответами
- вопросы типа «сверху - вниз»
- логические и грамматические подсказки
- бессистемное выражение цифровых данных
- тенденция к конвергенции
- двойные и тройные варианты ответов
- недостаточная фокусировка вопроса (условия задания)
- отсутствие логической связи между условием задания и вводным вопросом
- правильный ответ – самый длинный либо самый подробный
- негомогенные (неоднородные) либо заведомо не правдоподобные дистракторы
- использование отрицания в вопросе
- обращение к личности экзаменуемого.





# Экспертиза заданий в тестовой форме

**Рассмотрение на заседании предметного тестового Комитета (кафедра, профильный КОП и т.д.)**

**Повторная экспертиза по итогам апробации заданий в тестовой форме**

**Анализ уровня сложности и дистракторный анализ по итогам проведения экзамена - эксперты КОП и эксперты-тестологи.**

# Цели эмпрической апробации ТЕСТОВЫХ материалов

1. **Определение трудности заданий** с целью оценки его годности для испытуемых.
2. **Определение тех заданий, в которых есть существенные недостатки.**
3. **Выявление случайных ошибок** (опечаток, недостатков орфографии и пунктуации и т. д.).
4. **Определение времени, необходимого на решения задания или всего теста целиком.**
5. **Выявление недостатков** в инструкциях.

# Для чего?

- **Контроль качества ТЗ:**

Выявление дефектных компонентов  
(условие задания? вводный вопрос?  
дистракторы?)

- Определение неудачных заданий:  
**пересмотр, замена, удаление**



# По результатам эмпирической апробации часть заданий будет отбракована

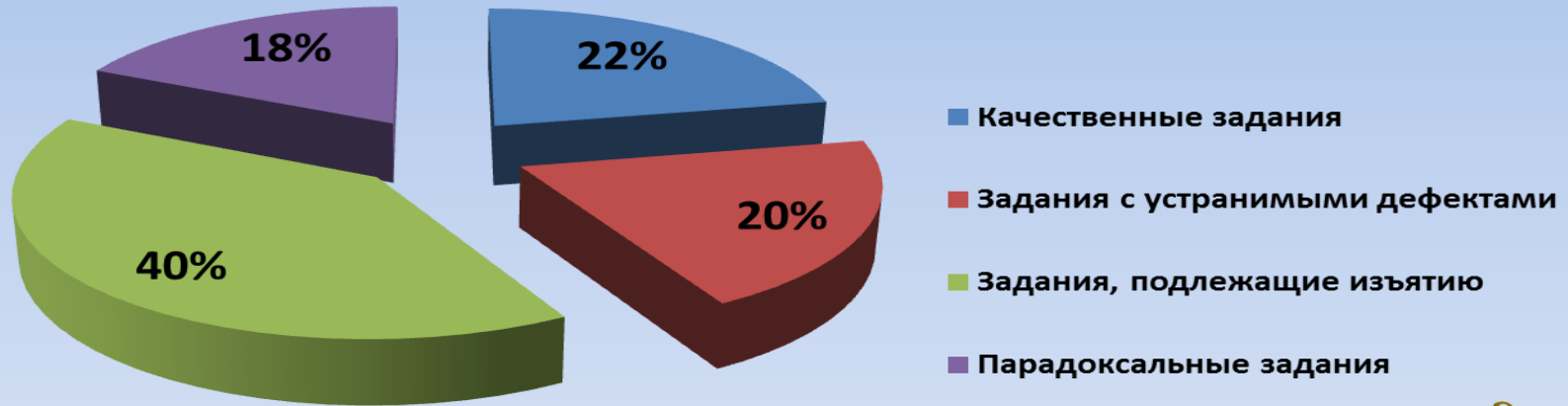
- Экспертизе подвергают от 20 до 100% от планируемого количества заданий
- Важно не только общее количество заданий, а сохранение их баланса согласно матрице/спецификации теста.

# Оценка заданий в рамках экспертизы

*В самом общем виде тестовые задания должны:*

- соответствовать содержанию учебной программы/ дисциплины / темы / раздела;
- быть составлены с учетом соответствующих правил в области методологии тестирования;
- быть ясными испытуемому;
- быть проверены на практике (апробированы).

# Результаты эмпирической апробации заданий в тестовой форме (пример)



# Уместность - содержательная экспертиза

- 1) Ясно ли из тестового вопроса, какое знание, навык или умение тестируется?
- 2) **Есть ли в вопросе какая-либо ловушка или головоломка?**
- 3) Смогут ли просто **«очень опытные в тестах»** экзаменуемые ответить правильно на вопрос? (**дефекты в виде логических либо грамматических подсказок**)
- 4) **Имеет ли вопрос достаточную степень сложности?**
- 5) **Не увеличивается ли степень трудности вопроса лишней информацией (балласт / словесная мишура)?**
- 6) **Соответствует ли тип вопроса цели теста?**
- 7) Приемлема ли степень трудности теста в-целом?
- 8) **Отвечает ли тест в-целом целям, установленным для тестовой программы?**
- 9) **Охватывает ли тест все дидактические единицы изучаемой дисциплины?**

# Применение языка – содержательная экспертиза

- 1) Сформулирован ли вопрос в соответствии с грамматическими правилами языка (казахский / русский / английский)?
- 2) Не содержит ли вопрос отрицание / двойное отрицание?
- 3) Не содержит ли вопрос ненужные / отвлекающие / непонятные / трудные слова?
- 4) Содержит ли вопрос ненужные предложения/дополнения / пояснения в скобках?
- 5) Не может ли изложение вопроса привести к затруднению понимания?



# Информация – содержательная экспертиза

- 1) Содержит ли Вопрос / условие задания достаточно информации для правильного ответа?
- 2) Достаточно ли информации в вопросе для получения ответа требуемой длины и формы? (каждая опция – в пределах одной строки)
- 3) Соблюдаются ли действующие условности при использовании символов, пунктуации и т. д.?
- 4) Проверены ли фотографии, рисунки, таблицы и т. д. на наличие ошибок?  
Могут быть спорные моменты в интерпретации невербальных символов

# Вопрос и ответ – содержательная экспертиза

1)Прочтите и ответьте на вопрос, не глядя на опции (варианты ответов)

2)Соответствует ли Ваш ответ правильному ответу (эталону)?

3)Есть ли Ваш правильный ответ среди вариантов ответа?

4)Нет ли другого, более одного правильного ответа среди дистракторов?

5)Есть ли другие правильные ответы, помимо тех, которые даны?

## Алгоритм составления задания в тестовой форме:

Сформулируйте вопрос и варианты ответов (опции).



Закройте варианты ответа листком бумаги.



Попросите опытного коллегу ответить на вопрос, основываясь только на условии задания.



Перепишите условия задания или варианты ответа (опции), если он не смог сделать этого.



# Варианты ответов (опции)

- 1) Нет ли взаимно исключающих себя вариантов?
- 2) Являются ли варианты ответа приблизительно одной длины? (для MCQ)
- 3) Являются ли варианты ответа грамматически и схематически приемлемыми продолжениями вопроса?
- 4) Может ли быть понят вариант ответа без чтения других вариантов?
- 5) Достаточно ли различимы имеющиеся варианты ответов?

# Убедитесь, что дистракторы:

- **Гомогенны по содержанию** (например, все варианты являются диагнозами; лекарственными препаратами и т.д.).
- **Менее вероятны**, чем правильный ответ.
- **Правдоподобны и привлекательны для** неинформированного мало информированного экзаменуемого.
- **Похожи** на правильный ответ **по конструкции и длине**.
- **Грамматически последовательны и логически совместимы** с условием задания.

## Вопросы

Ответы 7

## СОР и ККГ - интерны

1. Мужчина 50 лет, штурман, обратился с жалобами на чувство стянутости, шероховатости слизистой оболочки в области углов рта. Курит до 2 пачек сигарет в день в течение 20 лет. При осмотре: в области углов рта - очаги помутнения слизистой оболочки, сероватого цвета, треугольной формы. Поскабливание не дает эффекта. Проба на целостность эпителия отрицательная.

Какой из перечисленных предварительных диагнозов **НАИБОЛЕЕ** вероятен? \*

- Хронический гиперпластический кандидоз
- \*Красный плоский лишай, типичная форма
- \*Красная волчанка, типичная форма
- \*Лейкоплакия, плоская форма ✓
- \*Болезнь Боуэна

Предварительный просмотр

Темы

Сбор ответов

## Вопросы

Ответы 7

2. Мужчина 60 лет, военный пенсионер, жалуется на боль при приеме пищи. Хронические соматические заболевания отрицает. При осмотре: полость рта не санирована. Имеется некачественный мостовидный протез в области 3.5-3.8 зубов. На боковой поверхности языка слева - язва неправильной формы. Дно язвы неровное, покрыто некротическим налетом. Окружающая слизистая застойно гиперемирована. Пальпация болезненная. Какой из перечисленных предварительных диагнозов **НАИБОЛЕЕ** вероятен? \*

- Эрозивно-язвенная форма лейкоплакии
- \*Милиарно-язвенный туберкулез
- \*Декубитальная язва ✓
- \*Трофическая язва
- \*Раковая язва

3. Юноша 18 лет, студент, обратился с жалобами на сильную боль в области неба. Со слов

## Вопросы

Ответы 7

5. Мужчина 35 лет, сварщик, обратился с целью санации полости рта.. Считает себя соматически здоровым, курит с 20 лет. При осмотре: на слизистой щек и участках, прилежащих к углам рта, ограниченные пятна серо-белого цвета, не снимающиеся при поскабливании.

Какой из перечисленных лекарственных препаратов **НАИБОЛЕЕ** обоснован в плане местного патогенетического лечения? \*

- \*Ретинола ацетат ✓
- \*Аскорбиновая кислота
- \*Пиридоксин
- \*Фолиевая кислота
- \*Эргокальциферол

6. Женщина 40 лет, военнослужащая, обратилась к врачу-стоматологу с жалобами на



## Вопросы

Ответы 7

\*Системная красная волчанка



7. Мужчина 40 лет, программист, обратился к врачу-стоматологу с жалобами на дискомфорт слизистой оболочки рта. Считает себя практически здоровым, курит с 20 лет. При осмотре: на слизистой по линии смыкания зубов видны очаги ограниченных бляшек молочно-белого цвета, возвышающихся над уровнем слизистой, имеющие шероховатую поверхность и напоминающие омозолелость. Какой их перечисленных физиотерапевтических методов лечения **НАИБОЛЕЕ** целесообразен? \*

- \*Оксигенотерапия
- \*Вакуум-терапия
- \*Излучение гелий-неонового лазера
- \*Гидротерапия
- \*Криодеструкция ✓

Вопросы

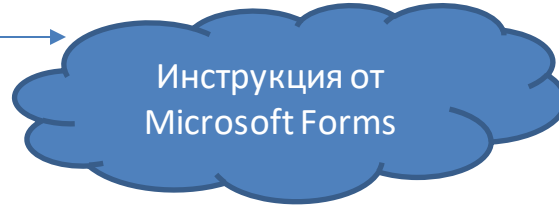
Ответы 7

8. Какие из перечисленных элементов поражения **НАИБОЛЕЕ** характерны для лейкоплакии? (возможно несколько вариантов правильных ответов) \*

Выберите несколько (2) вариантов.

- пятно ✓
- папула
- бляшка ✓
- корка
- чешуйка
- трещина

Инструкция от Microsoft Forms



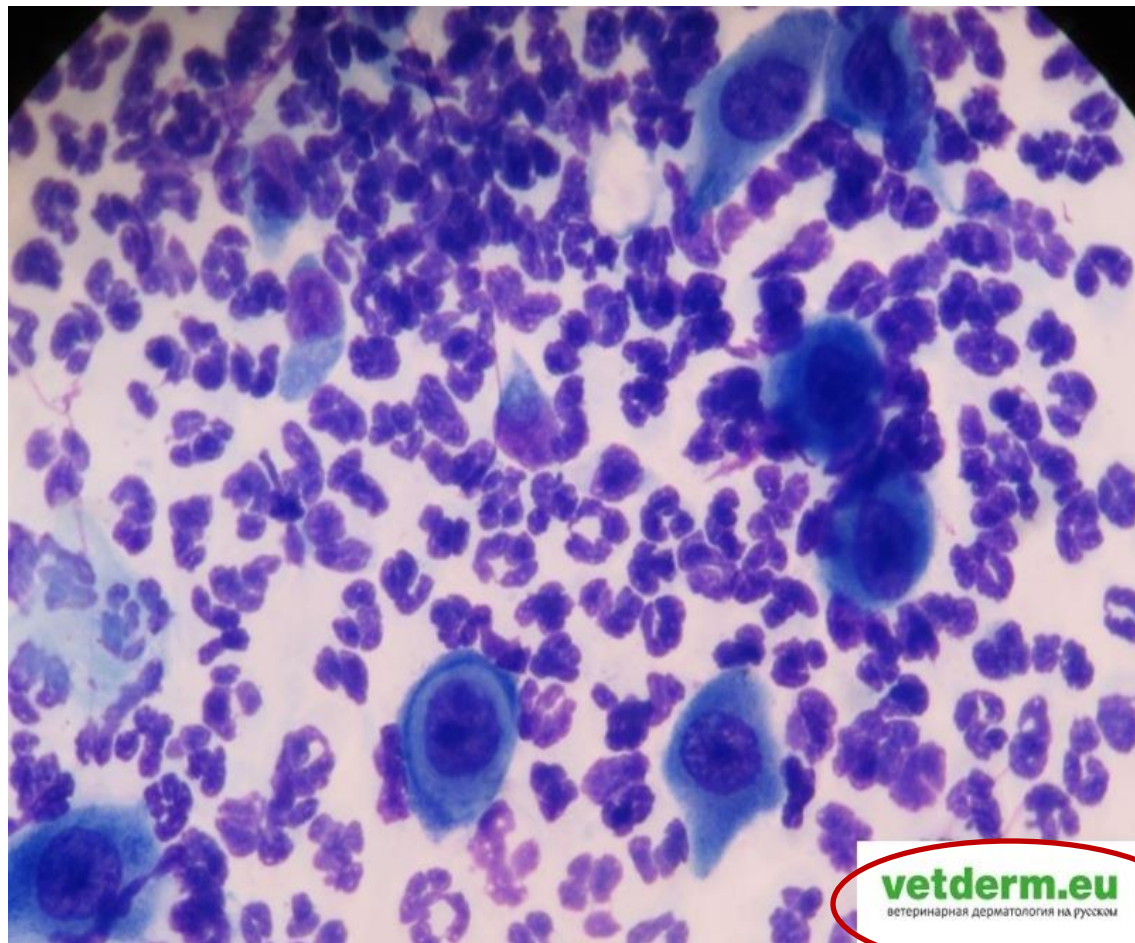
# Применение контекста – содержательная экспертиза

- 1) Является ли функциональным использование рисунков, диаграмм, рентгенограмм, результатов анализов и т. д.?
- 2) Ясны ли все рисунки, диаграммы, рентгенограммы, микро-препараты и т.д.?
- 3) Все ли надписи (пояснения) к рисункам и т. д. являются четкими, соответствующими и ясными?
- 4) Не содержит ли добавленный текст лишнюю информацию?

# Качество фотографий



# Качество микро- препаратов



# Качество таблиц (как не надо делать)

## Общий анализ крови (норма)

Показатель	Взрослые женщины	Взрослые мужчины
Гемоглобин	120—140 г/л	130—160 г/л
Эритроциты	$3,7—4,7 \times 10^{12}$	$4—5,1 \times 10^{12}$
Цветовой показатель	0,85—1,15	0,85—1,15
Ретикулоциты	0,2—1,2%	0,2—1,2%
Тромбоциты	$180—320 \times 10^9$	$180—320 \times 10^9$
СОЭ	2—15 мм/ч	1—10 мм/ч
Лейкоциты	$4—9 \times 10^9$	$4—9 \times 10^9$
Палочкоядерные	1—6%	1—6%
Сегментоядерные	47—72%	47—72%
Эозинофилы	0—5%	0—5%
Базофилы	0—1%	0—1%
Лимфоциты	18—40%	18—40%
Моноциты	2—9%	2—9%

большаясемья.рф

## Общий анализ крови кошек

Кличка животного 12.04

Ф.И.О. владельца \_\_\_\_\_

Дата « 17 » ноября

2020 г.

Показатели	Ед. изм.	Норма	Результат
Лейкоциты	$10 \times 3 \text{ мм}^3$	5,5-19,9	10
Эритроциты	$10 \times 6 \text{ мм}^3$	5-10	9,4
Гемоглобин	г/л	80-150	132
Гематокрит	%	24-45	38
Тромбоциты	$10 \times 3 \text{ мм}^3$	300-800	378
Нейтрофилы:	%	35-75	
юные	%		
сегментоядер.	%		45
палочкоядер.	%		6
Лимфоциты	%	20-55	45
Моноциты	%	1-3	7
Эозинофилы	%	2-12	3
Базофилы	%	0-1	
Тромбокрит	%	0,108-0,282	0,36
СОЭ			15
ЦП			

Вет. Врач \_\_\_\_\_

*Воспаление кровеносных сосудов*

ИД пробы:

30

Режим:

Цельн. кровь

Время анализа: 14/05/2020 10:49

Эл-ты

Результ

Ед. изм.

WBC

8.5

$10^9/L$

Lymph#

2.6

$10^9/L$

Mid#

0.6

$10^9/L$

Gran#

5.3

$10^9/L$

Lymph%

0.310

Mid%

0.072

Gran%

0.618

# Качество рентгенограмм



# Важность экспертизы

Экспертиза тестовых заданий является обязательным условием для создания тестов, **вне зависимости от уровня их применения.**

После того как **будут подготовлены вопросы, следует проверить их объективность и корректность формулировки.**

Задачу или вопрос можно считать **объективными или корректно сформулированными**, если мнения ряда экспертов о назначении вопроса, правильности формулировки и пригодности вариантов ответов **совпадают между собой.**



# Критерии оценки

В качестве экспертов тестовых заданий могут быть выступать **специалисты с большим опытом работы** (педагогической, методической, клинической).

**Первым этапом экспертизы** является обсуждение теста на кафедре. Самым простым вариантом экспертизы может быть такой, когда экспертам предлагается по трех- или пятибалльной шкале оценить задания в тестовой форме.

- 1) Соответствие целям тестирования (отвечает – частично отвечает – не отвечает);
- 2) Однозначность формулировки (однозначна – не совсем однозначна – неоднозначна);
- 3) Пригодность вариантов ответов (подходят – частично подходят – не подходят).

*Если хотя бы один из экспертов оценивает вопрос как частично не соответствующий тесту, его исключают из БТЗ.*

# Вопросы для эксперта

1. Соответствие тестового вопроса дисциплине, ОП, ГОСО и уровню обучения (бакалавриат (младший или старший курс), интернатура, резидентура ???)
2. Соответствие задания заявленной теме
3. Ориентировочное время на выполнение задания - ???
4. Сложность задания (предполагаемый удельный вес (%) испытуемых, которые могут справиться с заданием)
5. Правильность ответа (по усмотрению эксперта, задание изымается или дописывается правильный ответ).
6. Оценка приемлемости вариантов ответа (опций – ключ и дистракторы).
7. Оценка предлагаемого варианта невербальной поддержки (рисунки, фото, схемы, рентгенограммы и т.д.).

Вопросы

Ответы 10

# Тест - композиты хим. отверждения (Основы тер стом - Группа 33р - 2022 - 2023 уч. год)

Уважаемые студенты!

Вашему вниманию предлагается тест, состоящий из 28 вопросов.

Каждый вопрос теста предполагает один **НАИБОЛЕЕ** правильный ответ

Время, отведенное на тестирование - 35 минут.

Следите за таймером.

Ваш ответ будет автоматически отправлен по истечении времени тестирования, если Вы не успеете сделать это самостоятельно.

Удачи на тестировании!



# Тест - композиты хим. отверждения (Основы тер стом - Группа 33р - 2022 - 2023 уч. год)

Уважаемые студенты!

Вашему вниманию предлагается тест, состоящий из 28 вопросов.

Каждый вопрос теста предполагает один **НАИБОЛЕЕ** правильный ответ

Время, отведенное на тестирование - 35 минут.

Следите за таймером.

Ваш ответ будет автоматически отправлен по истечении времени тестирования, если Вы не успеете сделать это самостоятельно.

Удачи на тестировании!

🕒 35 мин.

Привет, Ирина, это форма со временем.

После запуска вы не сможете приостановить таймер. Не беспокойтесь, Forms предоставит вам напоминание о последней минуте перед отправкой. Ваши ответы будут отправлены автоматически, когда время истечет. Подготовьтесь перед началом, чтобы учитывать время отправки.

Инструкция  
M. Forms



Предварительный просмотр



Темы

Сбор ответов

Вопросы

Ответы 10

# Тест - композиты хим. отверждения (Основы тер стом - Группа 33р - 2022 - 2023 уч. год)



10

Ответы

23:37

Среднее время завершения

Активна

Состояние

Проверить ответы

Опубликовать оценки

Открыть в Excel

# Говорит ли среднее время завершения теста о его легкости для студентов?

Студент	Время завершения теста (мин : сек)	Балл по итогам теста
Студент 1	35:14	39
Студент 2	34:45	64
Студент 3	15:34	46
Студент 4	16:57	86
Студент 5	17:35	86
Студент 6	22:10	86
Студент 7	23:34	68
Студент 8	34:39	29
Студент 9	18:39	68
Студент 10	17:12	96

# Тестологическая экспертиза

Кроме **содержательной экспертизы** тестовых заданий, необходима **и тестологическая экспертиза**, в ходе которой должны проверяться:

- соответствие содержания задания и формы,
- **правильность пространственного и логического расположения элементов заданий**
- и другие элементы задания, которые не относятся к содержанию предмета.

# Схема тестового вопроса с одним **НАИБОЛЕЕ** правильным ответом (основа – клинический случай)

Условие задание - относительно длинное  
(максимум 7-8 строк компьютерного текста)

Вводный вопрос ( со словом **НАИБОЛЕЕ** ) – пишется с новой строки

Дистрактор 1

Дистрактор 2

Дистрактор 3

**НАИБОЛЕЕ** правильный ответ

Дистрактор 4

5 опций  
(вариантов  
ответов)



# Обратная связь по итогам экзамена

- Программное обеспечение -?
- Статистическая обработка -?
- Дистракторный анализ -?



- Уровень сложности
- Надежность
- Валидность



# Статистические параметры уровня тестового задания



- **Два основных понятия:**
- **Уровень сложности** – степень «тяжести» задания ( $P$ )
- **Дискриминационная способность** (индекс  $ID$ ) – насколько «хорош» вопрос (тестовое задание)
- **$ID$  0,35 и выше – предпочтительно**
- Отрицательное значение не допустимо
- Высокое «+» значение свидетельствует об эффективности деления испытуемых.
- Является параметром критериальной валидности.
- **Кто лучше ответил на вопрос?**
- **Сильные студенты?**
- **Слабые студенты?**
- **Все отвечают правильно?**
- **Никто не отвечает правильно?**

# Для чего делается анализ задания?

- **Контроль качества тестового вопроса MCQ: выявление дефектных компонентов вопроса (условие задания, вводный вопрос, либо дистракторы)**
- **Определение неудачных заданий:**  
**пересмотр, замена, удаление**
- **По международным стандартам, из прошлогодней БТЗ должно остаться не более 15-20% ЛУЧШИХ заданий в тестовой форме.**



# Дистракторный анализ

- Обнаружение **слишком трудных** либо **очень лёгких** тестовых заданий может свидетельствовать **о наличии дефектов:**
  - в условии задания
  - во вводимом вопросе
  - Либо необходимости замены дистракторов на более правдоподобные.

# Неудачи на тестировании

Причинами неудач обучающихся на тестировании могут стать...

Недостатки в преподавании отдельной темы

Недостатки в преподавании раздела дисциплины

Имеющиеся недоработки в учебно-методических материалах (кафедра, Университет)

Это требует детального анализа с привлечением всех участников образовательного процесса

# Уровень сложности



- Экзаменационный тест в целом должен включать в себя комплекс заданий различной сложности.
- Не вызывает сомнений тот факт, что **слишком простые** задания в тестовой форме, на которые **отвечают все**, как и **слишком сложные** задания, на которые **не может ответить ни один из испытуемых**, не обладают способностью дифференцировать студентов по уровню подготовки.
- И, в этом смысле, они **не являются качественными тестовыми заданиями.**

# Дистракторный анализ

По итогам эмпирической апробации базы тестовых заданий, важно не только получить информацию об **уровне сложности** предлагаемых испытуемым заданий в тестовой форме, но и оценить их способность **ранжировать обучающихся по уровню подготовки**.

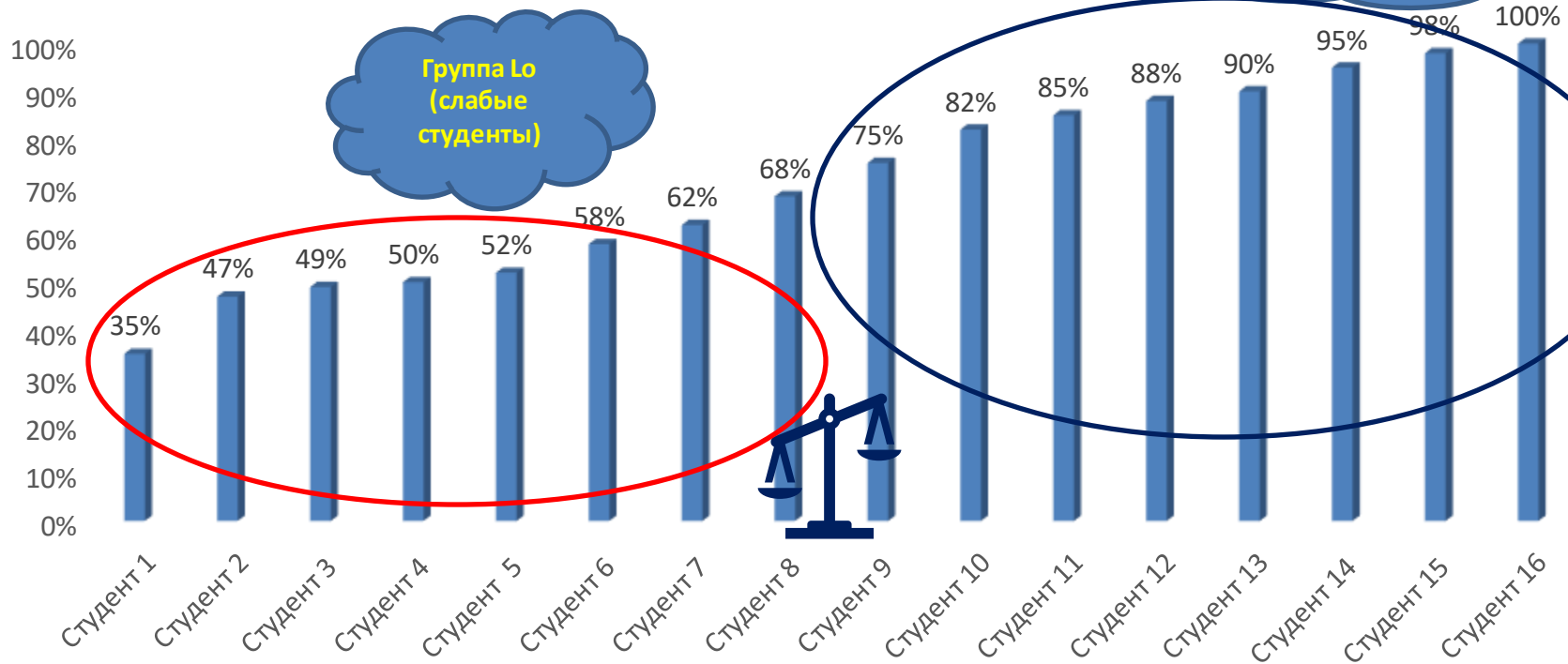
Не менее важным аспектом оценки тестового задания является получение ответа на вопрос: **«Кто правильно ответил на задание: сильные или слабые студенты?»**.

Может случиться так, что **вопрос окажется более лёгким для недостаточно подготовленных студентов**, которые более склонны к интуитивному угадыванию правильного ответа, **нежели для тех, кто привык размышлять, анализировать ситуацию и более склонен к сомнениям при выборе правильного ответа**.



# Результаты теста (пример)

Распределение студентов по итогам теста





# Пример задания

- Женщину 27 лет, страдающую ревматоидным артритом, в течение месяца беспокоят: выраженная слабость, повышение температуры тела 39,3°C. При осмотре: выраженная гиперпигментация кожи нижних конечностей, на голенях - трофические язвы. Лимфаденопатия, гепатоспленомегалия. В анализе крови: лейкоциты - 1,5 тыс. СОЭ - 60 мм/час, РФ - 1:80.
- Какое из перечисленных осложнений, **НАИБОЛЕЕ** вероятно, может развиться у пациентки в первую очередь?
- А) миокардит
- В) полисерозит
- С) токсический гепатит
- **D) инфекционные осложнения\***
- E) острая почечная недостаточность

# Распределение опций

Тестовое задание #25	Распределение опций, %				
	A	B	C	D*	E
Hi	0	0	70	30	0
Lo	38	12,5	0	38	12,5
Все	23	8	23	38	8

# Комментарии эксперта КОП (профильного специалиста)

Это задание про особую форму течения ревматоидного артрита – синдром Фелти, для которого характерна лейкопения и лимфопения.

Снижение уровня лейкоцитов закономерно становится причиной тяжёлых инфекционных осложнений.

Очевидно, **слабых студентов** информация о лихорадке, лимфаденопатии, гепатоспленомегалии натолкнула на **мысль о септическом состоянии**, и далее – об инфекционных осложнениях.

В то время, как **сильные студенты** ждали какого-либо подвоха.

Задание могло показаться им слишком «простым», и **правильный ответ они восприняли как «ловушку».**

Из отвлекающих ответов (дистракторов) был выбран миокардит и токсический гепатит (**это более логично** - гепатомегалия).

*Наверное, необходимо поменять все опции?*

# Комментарий эксперта-тестолога

На задание  
смогли  
ответить 38%  
испытуемых.

С точки зрения  
распределения опций,  
*казалось бы*, все в  
порядке: все  
дистракторы в той или  
иной степени  
востребованы  
испытуемыми.

**Однако** на тестовое  
задание правильно  
ответили только 30%  
испытуемых из группы  
Hi и 38% испытуемых из  
группы Lo.

Очевидно, что с  
заданием не все  
в порядке!!!

Отрицательное значение  
индекса дискриминации  
свидетельствует о том, что  
задание оказалось более  
лёгким для испытуемых  
из «слабой» группы».

1) Требуется  
коррекция  
формулировки  
задания.

2) Рекомендуются  
изъятие задания из  
БТЗ, и его  
переработка.

# Комментарий эксперта-тестолога

- *При разработке заданий подобного формата рекомендуется:*
- 1) термины «лимфаденопатия», «гепатоспленомегалия», являющиеся **логическими подсказками** для «слабых» студентов, не употреблять, а дать **короткое клиническое описание в тексте задания;**
- 2) **перефразировать вводный вопрос:** «Какое из перечисленных ниже осложнений является **НАИБОЛЕЕ** ожидаемым?»

## Пример задания

Женщина 28 лет отмечает утреннюю скованность суставов в течение часа, слабость, субфебрильную температуру. Больна в течение 6 лет, получала лечение. Беременность 18 недель. До беременности наблюдалась относительная ремиссия заболевания. Объективно: симметричная деформация проксимальных межфаланговых суставов II-IV пальцев обеих кистей, движения ограничены. ОАК: СОЭ 35 мм/ч.

Какой из перечисленных препаратов **НАИБОЛЕЕ** целесообразно назначить?

- А) кризано́л
- В) диклофенак
- С) метотрексат
- Д) преднизолон
- Е) сульфасалазин

## Распределение опций

Тестовое задание #38	A	B	C	D*	E	ID
H <sub>i</sub>	0	0	0	100	0	0,38
L <sub>o</sub>	0	14,5	71	14,5	0	
Все	0	8	38	54	0	

# Комментарии эксперта КОП (профильного специалиста)

- **«Сильные»** испытуемые знают, что беременным с ревматоидным артритом можно назначать **только преднизолон либо нестероидные противовоспалительные препараты**, но высокая активность заболевания предполагает **назначение именно преднизолона**.
- **«Слабые»** испытуемые, в основном, выбрали метотрексат – «золотой стандарт» для лечения ревматоидного артрита, но не учли, что **беременным его назначать нельзя** из-за тератогенного эффекта.



## Комментарий эксперта-тестолога

- Задание оказалось **достаточно сложным**, на него смогли правильно ответить **в-целом 54%** испытуемых,
- **НО с заданием все в порядке**, поскольку правильно ответили **100%** испытуемых из группы **Н<sub>1</sub>** и лишь **14,5%** испытуемых из группы **Lo**.
- Индекс дискриминации составил **0,38**.
- Рекомендуется **заменить дистракторы А и Е** на более правдоподобные, поскольку **их не выбрал никто**.

# Пример задания

- Мужчина 30 лет госпитализирован с нефротическим синдромом, повышенным уровнем креатинина, дерматитом в виде «бабочки», лимфаденитом. Выставлен диагноз системной красной волчанки, назначен преднизолон в дозе 65 мг/сут. Однако, при попытке снижения дозы препарата, развились выпотной перикардит, асцит.
- Какая лечебная тактика из перечисленных ниже **НАИБОЛЕЕ** целесообразна?
- А) Пульс-терапия метилпреднизолоном 500 мг в/в, циклофосфамидом 600 мг в/в
- В) Отмена преднизолона, назначение азатиоприна 400 мг и метотрексата 10 мг/нед
- С) Увеличение дозы преднизолона до 120 мг/сут, назначение метотрексата 10 мг/нед
- D) Пульс-терапия метилпреднизолоном в дозе 1000 мг в/в капельно, азатиоприн 200 мг
- E) Отмена преднизолона, назначение дексаметазона 20 мг в сочетании с азатиоприном 300 мг

## Распределение опций

Тестовое задание #12	A*	B	C	D	E	ID
Hi	100	0	0	0	0	0,38
Lo	14,5	0	14,5	71	0	
Все	54	0	8	38	0	

# Комментарий эксперта КОП (профильного специалиста)

- Большинство «**сильных**» и «**слабых**» испытуемых выбрали варианты лечения с пульс-терапией (**A** либо **D**).
- Но «**сильные**» студенты знают, что доза азатиоприна маловата, а сочетание метил-преднизолона и циклофосфана в приведённых дозах – наиболее стандартна для комбинированной пульс-терапии.
- Наверное, имеет смысл разработать все опции с вариантами пульс-терапии?

## Комментарий эксперта-тестолога

- Задание оказалось **достаточно сложным**, на него смогли правильно ответить 54% испытуемых.
- **Но с заданием все в порядке**, поскольку правильно ответили **100%** испытуемых **из группы Hi** и **лишь 14,5%** испытуемых **из группы Lo**.
- Индекс дискриминации составил **0,38**.
- Рекомендуется **заменить дистракторы В и Е на более правдоподобные**, поскольку их не выбрал никто.

# Пример задания

- Женщина 23 лет жалуется на повышение температуры тела до 40 С, одышку, кашель, боли в мелких суставах кистей и припухлость. Из анамнеза: самопроизвольный выкидыш месяц назад. Объективно: пониженного питания, лимфаденопатия, эритема щек и спинки носа. На коже конечностей - «сетчатое ливедо». Р-графия ОГК: выпот в полости плевры. В крови: эритроциты -2,4 млн., Нв-70 г/л, лейкоциты -2,2 тыс. СОЭ – 70 мм/час. В анализе мочи: белок-5,2 г/с, эритр-20-30 в п /зр. АНА и антитела к двуспиральной ДНК - положительные.
- Какое из перечисленных ниже обследований **НАИБОЛЕЕ** целесообразно?
  - А) бактериологический посев крови
  - В) определение ревматоидного фактора
  - **С) определение антител к кардиолипину\***
  - D) определение титра противострептококковых антител
  - E) иммуноферментный анализ на урогенитальную инфекцию

## Распределение опций

Тестовое задание #5	A	B	C*	D	E	ID
$H_i$	0	0	100	0	0	0,54
$L_o$	19	43	0	19	19	
Все	8	22	54	8	8	

# Комментарий эксперта КОП

- «Сильные» испытуемые, сопоставив выкидыш в анамнезе с волчанкой (системная красная волчанка – перечислено 8 критериев, достоверность диагноза 95%), смогли назвать наиболее частое осложнение – **вторичный антифосфолипидный синдром.**
- Может быть, заменить дистракторы А, D и E на более правдоподобные?



# Комментарий эксперта-тестолога

- С заданием смогли справиться **только студенты из группы Н<sub>i</sub> (100%)**.
- Это **сложное тестовое задание**, но **с ним все в порядке**, поскольку для всех студентов **из группы L<sub>o</sub>** дистракторы показались **достаточно привлекательными**.
- Задание имеет высокий индекс дискриминации ( 0,54 ) и **способно ранжировать испытуемых** на сильных и слабых.
- Из приведённых дистракторов **наиболее правдоподобным** показался **вариант В**, хотя варианты **А, D и Е** тоже привлекли «слабых» испытуемых с **приблизительно одинаковой долей вероятности**.