

ОЦЕНКА КАЧЕСТВА ТЕСТОВЫХ ЗАДАНИЙ MCQ ПО ПОКАЗАТЕЛЯМ  
СЛОЖНОСТИ И ДИСКРИМИНАТИВНОСТИ: ПОПЕРЕЧНОЕ  
ИССЛЕДОВАНИЕГ.К. КОШМАГАНБЕТОВА, Л.М. ЖАМАЛИЕВА, А.Р. КАШКИНБАЕВА,  
Г.С. ДИЛЬМАГАМБЕТОВА

Западно-Казахстанский медицинский университет имени Марата Оспанова, Актөбе, Казахстан

Кошмаганбетова Г.К. – <https://orcid.org/0000-0001-5895-346X>Жамалиева Л.М. – <https://orcid.org/0000-0003-3625-3651>Кашкинбаева А.Р. – <https://orcid.org/0000-0001-7173-0770>Дильмагамбетова Г.С. – <https://orcid.org/0000-0002-8123-3476>

## Citation/

библиографиялық сілтеме/  
библиографическая ссылка:

Koshmaganbetova GK, Zhamaliyeva LM, Kashkinbayeva AP, Dilmagambetova GS. Assessment of the quality of MCQ test items in terms of complexity and discrimination: a cross-sectional study. West Kazakhstan Medical Journal. 2020;62(3):146-151.

Кошмаганбетова Г.К., Жамалиева Л.М., Кашкинбаева А.Р., Дильмагамбетова Г.С. MCQ тест тапсырмаларының сапасын күрделілік және дискриминативтілік көрсеткіштері бойынша бағалау: көлденең зерттеу. West Kazakhstan Medical Journal. 2020;62(3):146-151.

Кошмаганбетова Г.К., Жамалиева Л.М., Кашкинбаева А.Р., Дильмагамбетова Г.С. Оценка качества тестовых заданий MCQ по показателям сложности и дискриминативности: поперечное исследование. West Kazakhstan Medical Journal. 2020;62(3):146-151.

**Assessment of the quality of MCQ test items in terms of complexity and discrimination: a cross-sectional study**

G.K. Koshmaganbetova, L.M. Zhamaliyeva, A.P. Kashkinbayeva, G.S. Dilmagambetova

West Kazakhstan Marat Ospanov Medical University, Aktobe, Kazakhstan

Multiple Choice Questions (MCQs) are the most widely used test format today. The effectiveness of MCQ is an effective assessment tool depends solely on their quality, which is best assessed by analyzing difficulty and discrimination indices.

**Purpose:** to assess the quality of MCQ test items in the “Research Management” discipline at the residency in terms of difficulty and discrimination index and cognitive levels of MCQ.**Methods.** The study was conducted among 215 residents of the first year of study in the first and second semesters of the 2018-2019 academic year at Aktobe Medical University. 73 MCQ test cases were analyzed for the difficulty and discrimination indices by the StartExam program.**Results.** The average difficulty index was 0.27 [95% CI 0.20; 0.34], the average discriminative score was 0.32 [95% CI 0.19; 0.40]. More than half of the test items had a low discrimination index. 42% of the test items showed a high discrimination index. 40% of the test tasks had optimal difficulty with a good share of correct and incorrect answers, almost half of the test tasks were too easy, it is recommended to change them. While 13.7% of the tests were too complicated also with the recommendation of a review. The discrimination index in our study did not correlate well with the difficulty index ( $r = 0.1$ ).**Conclusions:** test tasks with average optimal difficulty and a high discrimination index of the corresponding cognitive level should be included in the future to improve the quality of the test.**Keywords:** *difficulty index, discrimination index, item analysis, multiple-choice questions.***MCQ тест тапсырмаларының сапасын күрделілік және дискриминативтілік көрсеткіштері бойынша бағалау: көлденең зерттеу**Г. К. Кошмаганбетова, Л. М. Жамалиева, А. Р. Кашкинбаева,  
Г. С. Дильмагамбетова

Марат Оспанов атындағы Батыс Қазақстан медицина университеті, Ақтөбе, Қазақстан

Бүгінгі таңда тест форматының ең көп қолданылатын түрі бірнеше таңдау сұрақтары бар (MCQs) болып табылады. MCQ тиімді бағалау құралы ретіндегі тиімділігі тек олардың сапасына байланысты, ол күрделілік пен дискриминация көрсеткіштерін талдау арқылы жақсы бағаланады.

**Мақсаты:** Резидентурадағы «Зерттеулерді басқару» пәні бойынша MCQ тестілерінің сапасын күрделілік, дискриминация және MCQ когнитивтік деңгейіКошмаганбетова Г.К.  
e-mail: [gulbakit66@mail.ru](mailto:gulbakit66@mail.ru)Received/  
Келіп түсті/  
Поступила:  
05.06.2020.Accepted/  
Басылымға қабылданды/  
Принята к публикации:  
07.09.2020.ISSN 2707-6180 (Print)  
© 2020 The Authors  
Published by West Kazakhstan Marat Ospanov  
Medical University

тұрғысынан бағалау.

**Әдістері.** Зерттеу Ақтөбе медицина университетінде 2018-2019 оқу жылының бірінші және екінші семестрлерінде бірінші курстың 215 резиденттері арасында жүргізілді. 73 MCQ тест StartExam бағдарламалық жасақтамасының көмегімен күрделілік және дискриминация көрсеткіштері бойынша талданды.

**Нәтижелер.** Орташа күрделілік - 0,27 [95% CI 0,20; 0,34], орташа дискриминациялық балл - 0,32 [95% CI 0,19; 0,40] болды. Тест тапсырмаларының жартысынан көбі төмен дискриминацияға ие болды. Тест тапсырмаларының 42%-ы жоғары дискриминацияны көрсетті. Тест тапсырмаларының 40%-ы оңтайлы күрделілікке ие болды, дұрыс және дұрыс емес жауаптардың үлесі жақсы болды, тест тапсырмаларының жартысына жуығы өте оңай болды, сонымен бірге, тестілеудің 13,7% -ы өте күрделі болды. Біздің зерттеуіміздегі дискриминация индексі күрделілік индексімен жақсы сәйкес келмеді ( $r = 0.1$ ).

**Қорытынды.** Тесттің сапасын жақсарту үшін болашақта орташа оңтайлы күрделілігі мен жоғары дискриминациясы бар тест тапсырмалары тиісті когнитивтік деңгейге енгізілуі керек.

**Негізгі сөздер:** күрделілік индексі, дискриминация индексі, тақырыпты талдау, бірнеше таңдау сұрақтары.

#### Оценка качества тестовых заданий MCQ по показателям сложности и дискриминативности: поперечное исследование.

Г.К. Кошмаганбетова, Л.М. Жамалиева, А.Р. Кашкинбаева,  
Г.С. Дильмагамбетова

Западно-Казахстанский медицинский университет имени Марата Оспанова,  
Ақтөбе, Казахстан

Вопросы с множественным выбором (MCQ) являются наиболее широко используемым форматом тестирования сегодня. Эффективность MCQ как эффективного инструмента оценки зависит исключительно от их качества, которое лучше всего оценивается с помощью анализа индексов сложности и дискриминации.

**Цель.** Оценить качество тестовых заданий MCQ по дисциплине «Менеджмент научных исследований» в резидентуре по показателям сложности и дискриминативности и когнитивных уровней MCQ.

**Методы.** Исследование проводилось среди 215 резидентов первого года обучения в первом и втором семестрах 2018-2019 учебного года в медицинском университете Ақтөбе. 73 тестовых задания MCQ были проанализированы по индексам сложности и дискриминации с использованием программы «StartExam».

**Результаты.** Средняя сложность составила 0,27 [95% ДИ 0,20;0,34], средний показатель дискриминативности 0,32 [95%ДИ 0,19;0,40]. Более половины тестовых заданий имели низкую дискриминативность. Высокую дискриминативность показали 42 % тестовых заданий. 40% тестовых заданий имели оптимальную сложность с хорошей долей правильных и неправильных ответов. Почти половина тестовых заданий были слишком легкие, тогда как, 13,7% тестовых заданий были слишком сложными. Индекс дискриминативности в нашем исследовании плохо коррелировал с индексом сложности ( $r = 0,1$ ).

**Выводы.** Тестовые задания, имеющие среднюю оптимальную сложность и высокую дискриминативность, соответствующего когнитивного уровня должны быть включены в будущем для улучшения качества теста.

**Ключевые слова:** индекс сложности, индекс дискриминации, предметный анализ, вопросы с множественным выбором.

#### Введение

Вопросы с множественным выбором Multiple-Choice Questions, (MCQ) обычно признаются как наиболее широко применимый и полезный тип объективных тестовых заданий. MCQ обычно используются для оценки, потому что они могут предоставить большое количество элементов исследования, которые охватывают много областей содержимого, могут быть введены за относительно короткий период [1]. В целом, MCQ являются самыми сильными предикторами

общей успеваемости студентов по сравнению с другими формами оценки, такими как участие в классе, тематические экзамены, письменные задания и имитационные игры. Разработка качественного банка валидных и надежных заданий для вопросов множественного выбора, направленных на оценку когнитивных способностей и соответствующих рекомендациям, касающимся создания задания, является непростой задачей для разработчика [2].

Анализ предмета оценивает надежность и обоснованность экзамена, проверяя успеваемость учащихся

в отношении каждого MCQ и применяя статистический анализ, чтобы определить, следует ли хранить, проверять или исключать предмет из теста [3,4].

Инструментами для проверки правильности построения MCQ являются индексы сложности и дискриминации. Сложность – отношение количества неправильно ответивших участников на задание к общему количеству участников, ответивших на задание. Дискриминативность – показатель, который демонстрирует, какая из групп: сильная или слабая, лучше решает задание теста. Сильная группа – это 25% участников, показавшие лучшие результаты по тесту. Слабая группа – это 25% участников, показавшие худшие результаты по тесту. Дискриминативность – это разница между процентом правильных ответов в сильной и слабой группах [5]. Несоблюдение стандартных правил написания тестов может сделать экзаменационные вопросы легче или сложнее, чем предполагалось, о чем свидетельствуют результаты исследований оценивающих качество тестов [6,7]. Вопросы с множественным выбором (MCQ) должны быть не только нацелены на оценку воспоминания знаний, но и на измерение других целей обучения в рамках таксономии обучения Блума, таких как понимание, применение, анализ, синтез и оценка [8]. Исследований, направленных на оценку качества тестов для студентов-медиков по показателям сложности и дискриминативности и когнитивных уровней MCQ, в странах СНГ и в нашей стране недостаточно. Существует необходимость оценки качества применяемых тестов.

Цель исследования оценить качество тестовых заданий MCQ по дисциплине «Менеджмент научных исследований» в резидентуре по показателям сложности и дискриминативности и когнитивных уровней MCQ.

### Методы

Исследование поперечного дизайна проведено в первом и втором семестрах 2018-2019 учебного года при прохождении дисциплины «Менеджмент научных исследований» среди резидентов различных специальностей первого года обучения (n=215). Проводилось тестирование текущего контроля с использованием онлайн-формы тестовых заданий, созданных в программе «StartExam» (продукт компании Милдсофт, Россия). Тестовые задания (10 вопросов) были в форме MCQ с дистракторами и с единственным или множественным правильными ответами. А – тип – это вопрос, с одним наилучшим вариантом ответа; N-тип – это тестовый вопрос расширенного выбора, но с указанием точного числа правильных ответов из множества. Выделенное время составило 15 минут. Резидентам различных специальностей первого года обучения (n=215) предлагалось пройти по ссылке, отправляемой им в группу в социальной сети «ВКонтакте», об опыте применения «ВКонтакте» в образовательном процессе и об удовлетворенности обучающихся мы сообщали ранее [9,10].

Резиденты могли выбрать проходить ли тестирование с использованием компьютера или с использованием телефона с выходом в интернет. Резиденты предварительно проходили регистрацию (ФИО, номер группы, выбирая из всплывающего окна специальность, пол) и начинали тестирование (10 вопросов). Внутри теста существует навигация. Каждому резиденту выпадал свой набор вопросов. По завершении тестирования каждый обучающийся видел свой результат. Время и попытки сдачи теста были ограничены. За тест ставилось зачтено, если резидент правильно ответил на 60% и более вопросов.

В связи с отсутствием вмешательства одобрения этического комитета не требовалось.

Статистические методы: Обработка данных проводилась с использованием программы «Statistica 10» (StatSoft USA). Вычислены простые пропорции, в долях и частотах, среднее значение, с расчетом 95 % доверительных интервалов. Нормальность распределения проверена с использованием теста Колмогорова-Смирнова. Сложность и дискриминативность автоматически рассчитывалась программой «StartExam» (продукт компании Милдсофт, Россия). Карточка задания состоит из номера задания, ссылки на задание, статистических характеристик, текста задания и вариантов опций с указанной сложностью опций. Если задание является проблемным, то приводятся рекомендации относительно того, изменить ли задание (например, заменой нескольких опций) или полностью его исключить. Тесты были классифицированы в соответствии с их индексом сложности, индексом дискриминации. Сложность рассчитывается при наличии 3 и более ответов и имеет 3 уровня: Оптимальное задание [от 0,15 до 0,70], хорошая доля правильных и неправильных ответов. Слишком легкое задание [от 0,00 до 0,15], его рекомендуется изменить. Слишком сложное задание [от 0,70 до 1,00], его рекомендуется изменить. Дискриминативность – это разница между процентом правильных ответов в сильной и слабой группах. Рассчитывается при наличии 5 и более ответов и имеет 3 уровня: Высокая дискриминативность [от 0,30 до 1,00]. Сильная группа выполняет задание лучше, чем слабая группа. Задание хорошо разделяет сильных и слабых участников. Средняя дискриминативность [от 0,10 до 0,30]. Задание не очень эффективно разделяет сильных и слабых участников. Если это не связано с низкой или высокой сложностью, рекомендуется пересмотреть задание. Низкая дискриминативность [от -1,00 до 0,10]. Задание «не работает», т.к. не может отличить сильных или слабых участников. Отрицательная дискриминативность означает, что слабая группа выполняет задание лучше, чем сильная, что может указывать на ошибки в вариантах ответов [5].

Связь между индексом сложности и индексом дискриминации для каждого тестового элемента была определена корреляционным анализом, с расчетом коэффициента корреляции Спирмена [11].

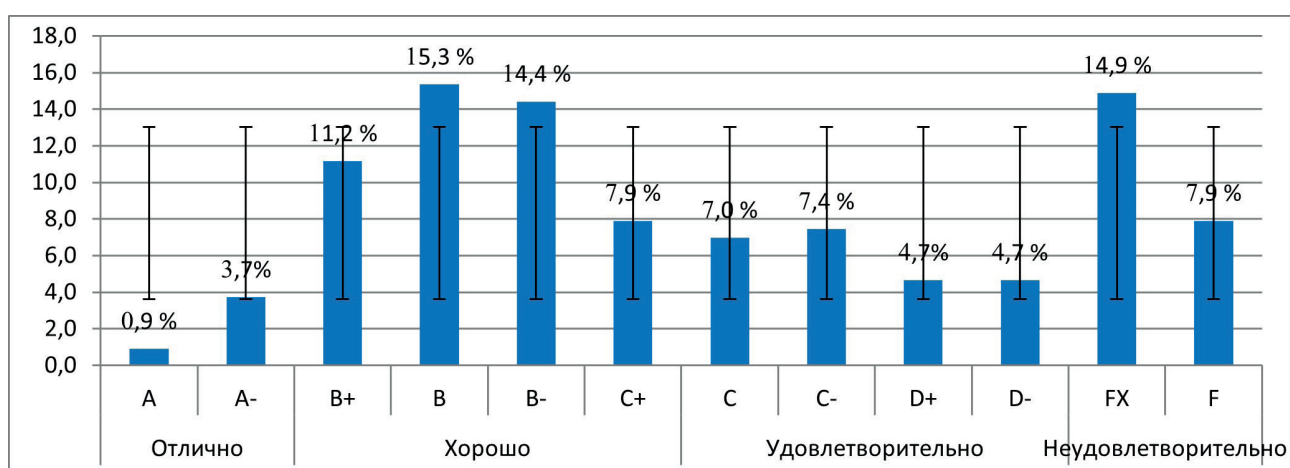


Рисунок 1. Распределение оценок обучающихся согласно бально-рейтинговой системе (%).

## Результаты

Всего тестирование прошли 215 обучающихся. Из них 144 женщины, 71 мужчин. Распределение оценок обучающихся (n=215) в % согласно бально-рейтинговой системе представлено на рисунке 1.

На отлично ответили 4,6% (n=11), на хорошо 49% (n=104), на удовлетворительно 23,8% (n=51), на неудовлетворительно 22,8% (n=49) резидентов.

Всего было проанализировано 73 тестовых задания. Средняя сложность составила 0,27 [95%ДИ 0,20;0,34], средний показатель дискриминативности 0,32 [95%ДИ 0,19;0,40].

Как видно из представленных данных (таблица 1.) 40% тестовых заданий имели оптимальную сложность с хорошей долей правильных и неправильных ответов, почти половина тестовых заданий имели сложность в диапазоне [от 0,00 до 0,15], то есть они были слишком легкие, их рекомендуется изменить. Тогда как, 13,7% тестовых заданий были слишком сложными [от 0,70 до 1,00], также с рекомендацией пересмотра.

В таблице 1 представлено распределение показателей дискриминативности тестовых заданий по уровням. Более половины тестовых заданий имели низкую дискриминативность [от -1,00 до 0,10]. 42 % тестовых заданий показали высокую дискриминативность [от 0,30 до 1,00]. На рисунке 2 показана взаимосвязь между индексом сложности и индексом дискриминации. В нашем исследовании между индексом сложности

и индексом дискриминации выявлена слабая взаимосвязь ( $r=0,1$ ).

По когнитивному уровню 36 % проанализированных заданий были на применение и 64 % на знание и понимание.

Обсуждение. Индексы сложности и дискриминации являются одними из инструментов для проверки правильности построения MCQ[5]. В настоящем исследовании средняя сложность составила 0,27 [95%ДИ 0,20;0,34], что соответствует оптимальному уровню сложности (0,15-0,70) [12]. Наши результаты согласуются с результатами Gajjar S, и Karelia BN [13,14]. Однако в другом исследовании сообщалось о более высоких средних значениях сложности [15]. Средний показатель дискриминативности в нашем исследовании составил 0,32 [95%ДИ 0,19;0,40], то есть задания хорошо разделяют сильных и слабых участников. Более 40% тестов имели высокую дискриминативность [от 0,30 до 1,00], то есть сильная группа выполняет задание лучше, чем слабая группа, как и в исследовании, проведенном в Индии, где индекс дискриминации по 12 (48%) пунктам был превосходным [16]. Однако, наличие низкой дискриминативности в половине тестовых заданий, требует их пересмотра. Это является возможным объяснением наличия почти у четверти резидентов неудовлетворительных оценок. Одно задание имело отрицательную дискриминативность, что означает, что учащиеся с более низкими способностями отвечали более правильно, чем уча-

Таблица 1. Показатели сложности и дискриминативности тестовых заданий

	уровень	Процент	95% ДИ
Сложность задания	Легкое	46,6	35,1-58
	Оптимальное	39,7	28,5-51
	Сложное	13,7	5,8-21
Дискри-минатив-ность	Низкая	56,7	45-69
	Средняя	1,5	1,4-4,4
	Высокая	41,8	30-53,6



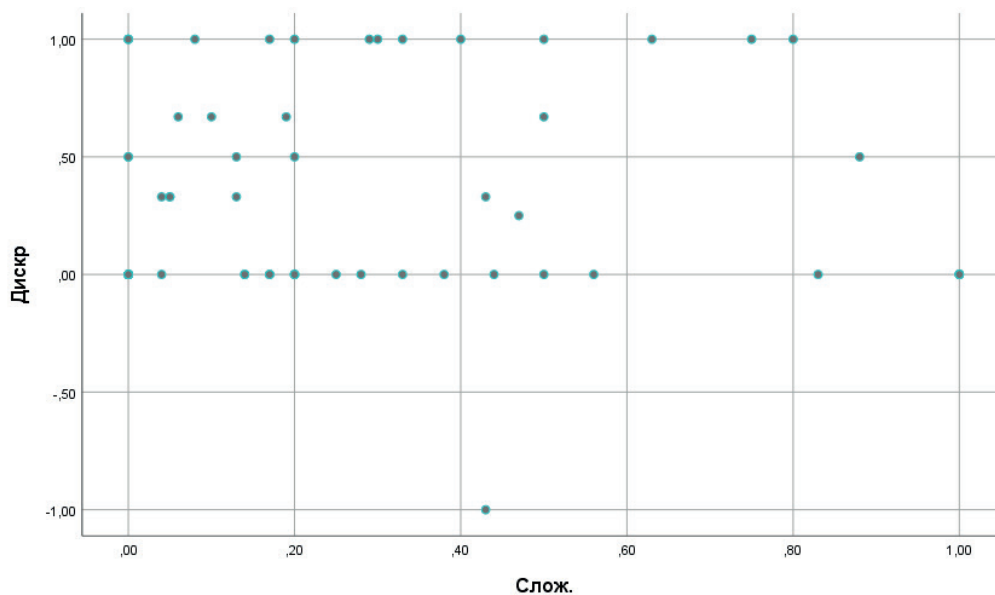


Рисунок 2. Точечная диаграмма, показывающая взаимосвязь между индексом сложности и индексом дискриминации тестов ( $r=0,1$ ).

щиеся с более высокой способностью. Причинами отрицательного показателя дискриминативности могут быть неверный ключ, неоднозначное формулирование вопросов или общая плохая подготовка студентов. Элементы с отрицательной дискриминативностью снижают валидность теста и должны быть удалены из набора вопросов. Индексы сложности и дискриминации обычно взаимно связаны, но их взаимосвязь часто считается куполообразной и нелинейной [17]. Индекс дискриминативности в нашем исследовании плохо коррелировал с индексом сложности ( $r = 0,1$ ), как и в исследовании Karelia BN и соавт., при анализе тестов по фармакологии у студентов второго курса [14]. Этот вывод свидетельствует о том, что вопросы с высоким значением дискриминативности плохо различаются, и наоборот, за исключением случаев, когда дискриминативность является либо чрезвычайно высокой, либо низкой [18,19]. Требуется пересмотр соотношения распределения заданий по когнитивному уровню, увеличение доли заданий на применение и включение заданий на анализ. Существуют индивидуальные и институциональные барьеры для подготовки качественных тестов. Реализация программ развития преподавателей, нацеленных на написание тестов, является одним из институциональных вмешательств, которые, как было установлено, улучшают качество тестов. Однако улучшение качества предметов может по-прежнему не производить достаточно качественные тесты со скоростью, необходимой для медицинских программ [20,21].

Применение онлайн-формы тестовых заданий, созданных в системе «StartExam» у резидентов в образо-

вательном процессе позволяет преподавателю проводить по статистике теста анализ (в деталях увидеть сколько тестов выполнено, пропущено, % верных ответов); проверить ответы на антиплагиат (эссе); создавать отчет по качеству тестов (уровень сложности и дискриминативности), с рекомендациями, что дает возможность преподавателю улучшить их; экономит время, бумагу, (переход на безбумажную форму); обеспечивает прозрачность и объективность оценки. Все это, в конечном счете, повышает качество образовательного процесса. Необходимо проведение в будущем сравнительных исследований для оценки эффективности системы «StartExam».

Наше исследование имеет следующие ограничения: не все задания были использованы, либо были использованы менее 5 раз, что не дает возможности расчета показателей сложности и дискриминативности в программе «StartExam». Кроме того, мы не рассчитывали эффективность дистрактора.

#### Выводы

Тестовые задания, имеющие среднюю оптимальную сложность и высокую дискриминативность, соответствующего когнитивного уровня должны быть включены в будущее для улучшения качества теста. Необходимо проанализировать качество, включая содержание инструментов оценки. Структурированная программа развития преподавателей рекомендуется для разработки улучшенных инструментов оценки, которые соответствуют результатам обучения и измеряют компетентность студентов-медиков.

Список литературы / References:

- Baig M, Ali SK, Ali S, Huda N. Evaluation of Multiple Choice and Short Essay Question items in Basic Medical Sciences. Pak J Med Sci. 2014 Jan;30(1):3–6. doi: 10.12669/pjms.301.4458.
- Coughlin PA, Featherstone CR. How to Write a High Quality Multiple Choice Question (MCQ): A Guide for Clinicians. Eur J Vasc Endovasc Surg. 2017 Nov;54(5):654–658. doi: 10.1016/j.ejvs.2017.07.012. Epub 2017 Sep 1.
- Pawluk SA, Shah K, Minhas R, Rankie D, Wilby KJ. A psychometric analysis of a newly developed summative, multiple choice question assessment adapted from Canada to a Middle Eastern context. Curr Pharm Teach Learn. 2018;10(8):1026–1032. doi:10.1016/j.cptl.2018.05.003.
- Kheyami D, Jaradat A, Al-Shibani T, Ali FA. Item Analysis of Multiple Choice Questions at the Department of Paediatrics, Arabian Gulf University, Manama, Bahrain. Sultan Qaboos Univ Med J. 2018;18(1):e68–e74. doi:10.18295/squmj.2018.18.01.011.
- Hingorjo MR, Laleel F. Analysis of One -Best MCQs: the Difficulty Index, Discrimination Index and Distractor Efficiency. J Pak Med Assoc. 2012; 62:142–47.
- Senecal EL, Askew K, Gorney B, Beeson MS, Manthey DE. Anatomy of a clerkship test. Acad Emerg Med. 2010;17(S2):S31–7. doi: 10.1111/j.1553-2712.2010.00880.x. [PubMed] [CrossRef] [Google Scholar].
- Rush BR, Rankin DC, White BJ. The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. BMC Med Educ. 2016;16(1):250. Published 2016 Sep 29. doi:10.1186/s12909-016-0773-3.
- Muhaissen SA, Ratka A, Akour A, Alkhatib HS. Quantitative analysis of single best answer multiple choice questions in pharmaceuticals. Curr Pharm Teach Learn. 2019;11(3):251–257. doi:10.1016/j.cptl.2018.12.006.
- Жамалиева ЛМ, Кошмаганбетова ГК, Калбагаева ГХ. Опыт использования социальной сети «ВКонтакте» в преподавании Менеджмента научных исследований в резидентуре. Батыс Қазақстан медицина журналы. 2017;2 (54):70–74. *Jamalieva LM, Koshmaganbetova GK, Kalbagaeva GH. Opyt ispolzovania sotsialnoi seti «Vkontakte» v prepodavanii Menejmenta nauchnyh issledovaniy v rezidentyre. Batys Qazaqstan medisina jyrnaly. 2017;2 (54):70–74.*
- Кошмаганбетова ГК. Оценка удовлетворенности обучающихся по дисциплине «Менеджмент научных исследований» в резидентуре ЗКГМУ им. М. Оспанова. Вестник АГИУ 2018;2:68–70. *Koshmaganbetova GK. Osenka ydovletvorennosti obuchayushhsya po discipline «Menejment nauchnyh issledovaniy» v rezidentyre ZKGMÝ im. M. Ospanova. Vestnik AGIÝ. 2018;2:68–70.*
- Гржибовский АМ. Корреляционный анализ // Экология человека. 2008. №9. URL: <https://cyberleninka.ru/article/n/korrelyatsionnyy-analiz> (дата обращения: 01.06.2020). *Grijbovskiy AM. Korrelatsionnyy analiz // Ekologiya cheloveka. 2008. №9. URL: https://cyberleninka.ru/article/n/korrelyatsionnyy-analiz (data obrasheniya: 01.06.2020).*
- The relationship between item difficulty and discrimination indices in multiple-choice tests in a Physical science course [MSc thesis]. Boca Raton, Florida: Florida Atlantic University; 2006 [cited 2015 June 14]. Available from: [http://www.physics.fau.edu/research/education/A.Hotiu\\_thesis.pdf](http://www.physics.fau.edu/research/education/A.Hotiu_thesis.pdf)
- Gajjar S, Sharma R, Kumar P, Rana M. Item and test analysis to identify quality multiple choice questions (MCQS) from an assessment of medical students Ahmedabad, Gujarat. Indian J Community Med. 2014;39:17–20.
- Karelia BN, Pillai A, Vegada BN. The levels of difficulty and discrimination indices and relationship between them in four response type multiple choice questions of pharmacology summative tests of year II MBBS students. IeJSME 2013; 7(2):41–6.
- Mehta G, Mokhasi V. Item analysis of multiple choice questions-an assessment of the assessment tool. Int J Health Sci Res. 2014; 4(7):197–202.
- Namdeo, Surya Kumar, & Bandy Sahoo. “Item analysis of multiple choice questions from an assessment of medical students in Bhubaneswar, India.” International Journal of Research in Medical Sciences [Online], 4.5 (2016):1716–1719. Web. 18 Feb. 2020.
- Menon AR, Kannambra PN. Item analysis to identify quality multiple choice questions. Nat J Lab Med. 2017;6:MO07–10. doi: 10.7860/NJLM/2017/25690:2215. [CrossRef] [Google Scholar]
- Kowash M, Hussein I, Al Halabi M. Evaluating the Quality of Multiple Choice Question in Paediatric Dentistry Postgraduate Examinations. Sultan Qaboos Univ Med J. 2019;19(2):e135–e141. doi:10.18295/squmj.2019.19.02.009
- Sim SM, Rasiah RI. Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a Para-clinical multidisciplinary paper. Ann Acad Med Singap. 2006;35(2):67–71.
- Karthikeyan S, O’Connor E, Hu W. Barriers and facilitators to writing quality items for medical school assessments - a scoping review. BMC Med Educ. 2019;19(1):123. Published 2019 May 2. doi:10.1186/s12909-019-1544-8
- HM, Irshad M, Haque S, Ahmad T, Sattar K, Khalil MS. Effectiveness of longitudinal faculty development programs on MCQs items writing skills: a follow-up study. PLoS One. 2017;12(10):e0185895.

Авторство

КГК, ЖЛМ,КАР,ДГС соответствуют критериям авторства Международного комитета редакторов медицинских журналов: 1) внес существенный вклад в концепцию и дизайн исследования, сбор, анализ и интерпретацию данных; 2) подготовил первый вариант статьи или её существенно переработал на предмет важного интеллектуального содержания; 3) окончательно утвердил присланную в редакцию рукопись.