

АНАЛИЗ РЕЗУЛЬТАТОВ ИТОГОВЫХ ЭКЗАМЕНАЦИОННЫХ РАБОТ ПО ИНТЕНСИВНОМУ КУРСУ ЧТЕНИЯ НА КИТАЙСКОМ ЯЗЫКЕ

Турсынгали Н. преподаватель,

Марасулова Г.А. преподаватель,

Южно-Казахстанский университет имени

М.Ауэзова, ШЫМКЕНТ, Казахстан

nursaule.tursyngali@gmail.com,

gulbakhyt.marasulova@mail.ru

Аннотация: Тесты - это измерения экспериментального характера, то есть сочетание измерений и испытаний. Преподаватели китайского языка часто задают свои собственные квазитестовые вопросы. Обычно мы проводим тесты для закрепления знаний студентов после изучения лексики или текстов на китайском языке. В данной работе мы используем тесты, отражающие реальную ситуацию обучения студентов. Мы также анализируем надежность и валидность тестовых заданий. Наконец, мы сформулировали несколько требований к составлению тестовых заданий.

Annotatsiya: Testlar eksperimental tabiatning o'lchovlari, ya'ni o'lchovlar va testlar kombinatsiyasi. Xitoylik o'qituvchilar ko'pincha o'zlarining kvazi-test savollarini berishadi. Biz odatda xitoy tilida so'z boyligi yoki matnlarni o'rgangandan so'ng talabalar bilimni mustahkamlash uchun testlar o'tkazamiz. Bu ishda biz o'quvchilarning o'zlashtirishining haqiqiy ahvolini aks ettiruvchi testlardan foydalanamiz. Shuningdek, test fanlarining ishonchliligi va haqiqiyiligini tahlil qilamiz. Nihoyat, test topshiriqlarini tuzish uchun bir nechta talablarni ishlab chiqdik.

Abstract: Tests are measurements of an experimental nature, i.e. a combination of measurements and tests. Chinese teachers often ask their own quasi-test questions. We usually do tests to strengthen students' knowledge after learning vocabulary or texts in Chinese. In this work, we use tests that reflect the true state of student learning. We will also analyze the reliability and validity of test subjects. Finally, we have developed several requirements for drawing up test tasks.

Ключевые слова: Китайский язык, тест, обучение, надежность и валидность тестовых заданий

Ключевые слова: Китайский язык, тест, обучение, надежность и валидность тестовых заданий

Tayanch iboralar: Xitoy tili, test, o'rganish, test topshiriqlarining ishonchliligi va haqiqiyiligi

I. Объект исследования

Объектом исследования являются иностранные студенты первого курса бакалавриата, изучающие китайский язык в Южно-Казахстанском университете имени М.Ауэзова. В качестве объекта исследования были взяты их интенсивные курсы чтения на китайском языке в первом семестре 2023-2024 учебного года, всего их было 14 человек. Учебник, используемый в классе интенсивного чтения китайского языка для иностранных студентов - "Самоучитель китайского языка".

II. Инструменты и задачи исследования

В качестве инструмента исследования используется экзаменационная работа по курсу интенсивного чтения на китайском языке для иностранных студентов первого курса программы бакалавриата по китайскому языку Шанхайского международного исследовательского университета (SISU), которая составлена в форме субъективных вопросов, всего 103 вопроса из 100 баллов, время экзамена - 90 минут, режим экзамена - письменный тест. Образец экзаменационной работы состоит из 10 основных вопросов.

Подробная информация приведена в таблице ниже:

Таблица 2.1 Типы экзаменационных вопросов

Номер экзаменационного вопроса	Название темы	Всего вопросов и оценок
I	диктовать	5 вопросов / 10 балл
II	Выберите правильное произношение слова с пунктуацией	8 вопросов / 4 балл
III	расширение слова	12 вопросов / 6 балл
IV	Выберите правильные слова, чтобы заполнить пропуски	24 вопросов / 12 балл
V	Выберите наиболее подходящий ответ	15 вопросов / 15 балл
VI	Завершите предложение.	8 вопросов / 12 балл
VII	исправить ошибку	8 вопросов / 12 балл
VIII	составление предложений	8 вопросов / 12 балл
IX	Интегрированный заполнитель	14 вопросов / 7 балл
X	сочинение (студенческое эссе)	1 вопрос / 10 балл

III. Тестовые задания

Мы провели предварительную статистику по итоговым экзаменационным тестам 14 студентов-международников Шанхайского международного исследовательского университета (SISU), изучавших китайский язык в течение одного года на интенсивном курсе чтения, и вывели оценки студентов по всем основным вопросам, а также итоговые суммарные баллы и их обычные оценки следующим образом, при этом имена студентов скрыты и представлены только их номерами. Мы используем n1, n2 n10 для обозначения 10 основных вопросов.

Таблица 3.1 Таблица исходных достижений

ID	SCORE	GROUP	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10
1	80.00	4.00	8.50	2.50	5.00	10.00	12.00	11.50	9.00	10.00	4.50	7.00
2	83.00	4.00	9.00	3.00	5.50	12.00	13.00	12.00	7.50	10.00	4.00	7.00
3	77.00	3.00	9.50	3.00	4.00	9.50	10.00	9.50	10.50	9.50	4.50	7.00
4	92.00	5.00	9.00	3.50	6.00	12.00	14.00	12.00	10.50	11.50	5.50	8.00
5	38.00	1.00	3.50	2.00	1.50	3.00	6.00	4.50	4.50	4.00	3.00	6.00
6	67.00	2.00	9.50	3.00	3.50	11.00	10.00	7.50	7.50	4.50	4.50	6.00
7	24.00	1.00	2.50	1.50	1.00	3.00	5.00	2.00	1.50	.00	3.50	4.00
8	83.00	4.00	9.50	4.00	6.00	12.00	13.00	9.00	6.00	10.50	5.00	8.00
9	78.00	3.00	8.00	2.00	5.00	10.00	12.00	8.00	9.50	9.50	6.00	8.00
10	66.00	2.00	5.00	2.50	3.00	11.00	9.00	6.00	7.50	11.00	5.00	6.00
11	68.00	2.00	5.50	3.50	4.00	8.50	10.00	11.50	8.50	6.00	3.50	7.00
12	85.00	4.00	9.50	3.00	4.50	12.00	11.00	11.00	11.00	10.50	5.50	7.00
13	72.00	3.00	8.00	3.00	4.50	10.50	8.00	10.00	8.00	11.00	4.00	5.00
14	77.00	3.00	8.50	3.00	5.00	11.00	10.00	9.00	10.50	9.00	5.00	6.00

IV. Описательный статистический анализ результатов экзамена по интенсивному курсу чтения на китайском языке

Средний балл данного экзамена составляет 70,7, средний балл (медиана) - 77, основной балл кандидатов - 77, стандартная разница (стандартное отклонение) между разделенными средними баллами кандидатов - 18,5, параболическая структура (перекос) баллов имеет отрицательный перекос с наклоном -1,675, подробности можно увидеть в следующей таблице.

Frequencies

		SCORE	GROUP
N	Valid	14	14
	Missing	0	0
Mean		70.7143	2.9286
Median		77.0000	3.0000
Mode		77.00(a)	3.00(a)

Std. Deviation	18.55316	1.20667
Variance	344.220	1.456
Skewness	-1.675	-.150
Std. Error of Skewness	.597	.597
Kurtosis	2.577	-.748
Std. Error of Kurtosis	1.154	1.154
Minimum	24.00	1.00
Maximum	92.00	5.00
Sum	990.00	41.00
Percentiles	25	66.7500
	50	77.0000
	75	83.0000

HISTOGRAM

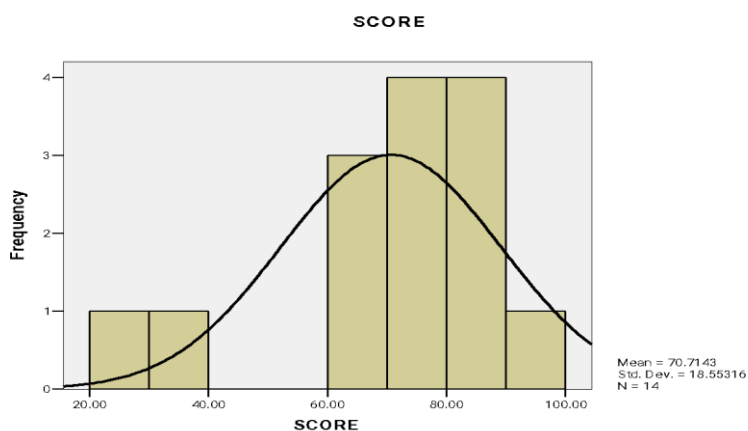


Рисунок 4.1 Столбчатое распределение баллов итогового экзамена по интенсивному курсу чтения на китайском языке

В следующей таблице показано частотное распределение баллов итогового экзамена по интенсивному курсу чтения на китайском языке. Частота - это частота каждого балла, Процент - процент балла, а кумулятивный процент - это кумулятивный процент балла, который обозначает место балла кандидата в общей группе. Например, если кандидат набрал 77 баллов, а кумулятивный процент этого балла составляет 57,1%, то итоговая оценка этого студента на

интенсивном курсе китайского чтения лучше, чем у 57,1% кандидатов. Исходя из количества людей, вошедших в каждую группу баллов, и их процентного соотношения, можно определить количество и долю людей в каждой группе баллов.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 24.00	1	7.1	7.1	7.1
38.00	1	7.1	7.1	14.3
66.00	1	7.1	7.1	21.4
67.00	1	7.1	7.1	28.6
68.00	1	7.1	7.1	35.7
72.00	1	7.1	7.1	42.9
77.00	2	14.3	14.3	57.1
78.00	1	7.1	7.1	64.3
80.00	1	7.1	7.1	71.4
83.00	2	14.3	14.3	85.7
85.00	1	7.1	7.1	92.9
92.00	1	7.1	7.1	100.0
Total	14	100.0	100.0	

Мы выяснили, что среди 14 человек в этом классе есть 2 человека с неудовлетворительными оценками (0-59), 3 человека с проходными оценками (60-69), 4 человека со средними оценками (70-79), 4 человека с хорошими оценками (80-89) и 1 человек с отличными оценками (90-100). Подробную информацию можно найти в следующей таблице.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 不及格	2	14.3	14.3	14.3
及格	3	21.4	21.4	35.7

中	4	28.6	28.6	64.3
良好	4	28.6	28.6	92.9
优秀	1	7.1	7.1	100.0
Total	14	100.0	100.0	

I. Анализ пунктов экзаменационных работ по интенсивному курсу чтения на китайском языке

Таблица 5.1 Статистический анализ пунктов оригинальных экзаменационных работ

	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10
N Valid	14	14	14	14	14	14	14	14	14	14
Missing	0	0	0	0	0	0	0	0	0	0
Mean	6.964	2.82	4.178	9.678	10.21	8.821	8.000	8.357	4.535	6.571
Median	3	14	6	6	43	4	0	1	7	4
Mode	8.250	3.00	4.500	10.75	10.00	9.250	8.250	9.750	4.500	7.000
Std. Deviation	0	00	0	00	00	0	0	0	0	0
Variance	9.50	3.00	5.00	12.00	10.00	9.00(a)	7.50(a)	9.50(a)	4.50(a)	7.00
Skewness	2.804	.668	1.514	3.016	2.607	2.997	2.638	3.404	.8652	1.157
Std. Error of Skewness	28	15	13	67	05	48	47	91	3	87
Kurtosis	7.864	.446	2.293	9.100	6.797	8.985	6.962	11.593	.749	1.341
Std. Error of Kurtosis	-.933	-.357	-.961	-1.765	-.607	-1.028	-1.209	-1.448	-.127	-.722
Range	.597	.597	.597	.597	.597	.597	.597	.597	.597	.597
Minimum	-.627	.052	.400	2.228	-.059	.574	1.602	1.391	-.676	.442
Maximum	1.154	1.15	1.154	1.154	1.154	1.154	1.154	1.154	1.154	1.154
	4									
Range	8.00	2.50	5.00	9.00	9.00	10.00	9.50	11.50	3.00	4.00
Minimum	1.50	1.50	1.00	3.00	5.00	2.00	1.50	.00	3.00	4.00
Maximum	9.50	4.00	6.00	12.00	14.00	12.00	11.00	11.50	6.00	8.00

Sum		97.50	39.50	58.50	135.50	143.00	123.50	112.00	117.00	63.50	92.00
Percentiles	25	4.6250	2.3750	3.3750	9.2500	8.7500	7.1250	7.1250	5.6250	3.8750	6.0000
	50	8.2500	3.0000	4.5000	10.7500	10.0000	9.2500	8.2500	9.7500	4.5000	7.0000
	75	9.1250	3.1250	5.1250	12.0000	12.2500	11.5000	10.5000	10.6250	5.1250	7.2500

5.1 Сложность вопросов, дифференциация и коэффициент надежности тестовых вопросов

Item Statistics

	Mean	Std. Deviation	N
n1	7.5357	2.38960	14
n2	2.8214	.66815	14
n3	4.1786	1.51413	14
n4	9.6786	3.01667	14
n5	10.2143	2.60705	14
n6	8.8214	2.99748	14
n7	8.0000	2.63847	14
n8	8.3571	3.40491	14
n9	4.5357	.86523	14
n10	6.5714	1.15787	14

Из таблицы видно, что среднее значение n1 равно 7,5357, так как наш метод подсчета баллов - правильно - 1, неправильно - 0. Этот средний балл и есть сложность первого вопроса, также называемая индексом легкости (item facility), который показывает легкость вопроса. Принято считать, что сложность вопроса находится в диапазоне от 0,3 до 0,7, при этом вопросы ниже 0,3 являются более сложными и требуют доработки, вопросы выше 0,7 - более легкими, а вопросы со сложностью 0,5 - самыми лучшими. Однако цель данного теста - проанализировать вопросники и выяснить степень достижения

целей обучения и степень освоения студентами целей обучения. Если студенты освоили все знания и навыки, представленные в вопросах, то сложность каждого вопроса будет равна 1. В этом случае вопрос со сложностью 0,7 или выше означает, что студенты освоили знания или навыки, представленные в вопросе, очень хорошо; вопрос со сложностью 0,5 или ниже означает, что студенты освоили знания или навыки, представленные в вопросе, очень плохо. StdDev представляет собой стандартное отклонение вопросов, и вопросы с наименьшим стандартным отклонением среди 10 вопросов - это вопросы 2 и 9 (StdDev = 0,66815, 0,86523), что означает, что разница в оценках испытуемых по этим трем вопросам невелика. n означает количество испытуемых. Но в ходе исследования мы обнаружили, что стандартное отклонение вопросов превышает 0,5, что говорит о том, что вопросы легкие.

Corrected Item-Total Correlation - это дифференциация вопроса, которая показывает, насколько вопрос способен дифференцировать отличников и неуспевающих студентов. Из 10 вопросов лучше всего дифференцирован вопрос 3, с дифференциацией 0,934, а хуже всего - вопрос 2, с дифференциацией 0,694.

Коэффициенты надежности - это коэффициенты надежности вопросов теста. Чем выше коэффициент, тем меньше статистическая ошибка. Принято считать, что коэффициент надежности вопросов объективного теста 0,8 и выше означает, что погрешность вопросов теста мала и надежность хорошая; коэффициент надежности 0,9 и выше - отличная; коэффициент надежности 0,7-0,8 - средняя; коэффициент надежности 0,6-0,7 - приемлемая; коэффициент надежности 0,6 и ниже - неприемлемая, и вопросы должны быть пересмотрены. Надежность вопросов объективного теста зависит от сложности и дифференцированности вопросов, а дифференцированность - от вариативности испытуемых; чем больше разница в уровне способности испытуемых к интенсивному чтению на китайском языке, тем выше дифференцированность вопросов и тем выше надежность вопросов теста. Чем больше вопросов, тем,

естественно, выше надежность. Количество испытуемых для данного набора вопросов составляет 14 человек, количество вопросов - 10, а коэффициент альфа равен 0,937.

Reliability Statistics

Cronbach's Alpha	N of Items
0.937	10

Основным основанием для дифференциации компромиссов является Alpha if Item Deleted (коэффициент надежности после удаления темы), если надежность всего набора тестовых вопросов уменьшается по сравнению с исходной после удаления темы, это означает, что тема очень важна для всего набора тестовых вопросов и дифференциация идеальна, и наоборот, это означает, что дифференциация темы не идеальна и должна быть изменена или устранена. Как видно из приведенного ниже графика, надежность тестовых вопросов снижается после удаления каждого вопроса, что говорит о том, что дифференциация этих вопросов идеальна.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
n1	63.1786	271.446	.852	.925
n2	67.8929	327.007	.694	.941
n3	66.5357	293.479	.934	.927
n4	61.0357	249.364	.900	.923
n5	60.5000	262.885	.882	.923
n6	61.8929	255.238	.835	.927
n7	62.7143	268.412	.796	.928
n8	62.3571	244.516	.827	.930
n9	66.1786	322.639	.670	.939
n10	64.1429	312.747	.736	.936

VI. Резюме

Из проанализированных данных вопросника следует, что надежность и валидность данного набора вопросников соответствуют требованиям и

достигают ожидаемого эффекта, который может отражать реальную учебную ситуацию студентов. Однако данные также показали, что некоторые вопросы в работе были слишком простыми по сложности, что снизило дифференциацию этой части теста, что требует нашего внимания при подготовке тестовой работы.

В то же время, анализируя данные тестовой работы, можно увидеть, что для того, чтобы сделать подготовку тестовой работы более разумной, будущее направление усилий по подготовке тестов достижений в целом следующее:

- 1) предлагающий тест учитель должен попытаться отрегулировать соотношение между степенью надежности и валидности в соответствии с различными целями теста, и контролировать определенную степень дифференциации и полезности;
- 2) каждый тест должен быть составлен после исчерпывающей и научной статистической обработки, соответственно, по сложности теста, дифференциации, проверке надежности и валидности, всесторонней проверке кандидатов и валидности теста, соответственно, для проверки сложности, дифференциации, надежности и валидности теста, а также для систематического анализа качества тестовой работы путем интеграции результатов кандидатов и проверки разрыва между фактическим и ожидаемым;
- 3) улучшение средств тестирования и типов тестовых вопросов, например, увеличение количества пунктов субъективного вопроса и установка нефиксированных типов вопросов.

Список использованной литературы

1. Чжан Янь Ли. «Курс РРТ для языкового тестирования и оценки», Шанхайский университет международных исследований. 2016г.
2. Чэнь Вэй. «Исследование обоснованности преподавания китайского языка в качестве теста на знание иностранного языка», Шанхайский университет международных исследований. 2010г.
3. Материал WORD «Применение SPSS в исследовании языкового тестирования».