

Грязина Галина Ивановна,
преподаватель ОПД, высшая категория,
ГАОУ СПО СО «Училище олимпийского резерва №1» (колледж),
г. Екатеринбург,
Стаж педагогической работы – 22года

ПЕДАГОГИЧЕСКОЕ ТЕСТИРОВАНИЕ В СИСТЕМЕ ОЦЕНКИ И УПРАВЛЕНИЯ КАЧЕСТВОМ ОБРАЗОВАНИЯ (РАЗРАБОТКА И АНАЛИЗ ИТОГОВОГО ТЕСТА ПО ДИСЦИПЛИНЕ «ГИГИЕНА»)

В соответствии с Законом РФ «Об образовании» в практику современного профессионального образования вошло понятие государственного образовательного стандарта (ФГОС СПО 4). Однако, без эффективной системы контроля выполнения его требований адекватное функционирование образовательного стандарта практически невозможно. Объективным средством определения параметров качества образовательного процесса может выступать тестовый контроль как один из элементов педагогического контроля.

Тестирование является более качественным и объективным способом оценивания, что достигается путем стандартизации процедуры проведения, проверки показателей качества заданий и тестов целиком. Тестирование – более справедливый метод, оно ставит всех учащихся в равные условия, как в процессе контроля, так и в процессе оценки.

Тесты позволяют обеспечить своевременное выявление уровня знаний, практических навыков и умений обучаемых в соответствии с образовательными стандартами, вносить коррективы в содержание, организацию и методику преподавания учебных дисциплин.

Тема моей работы «Разработка и анализ оригинального (авторского) тестового материала по дисциплине «Гигиена», сформированного по классической тестовой технологии.

Дисциплина «Гигиена» изучается на 2курсе. Целью «Гигиены» является обоснование гигиенических норм, нормативов, правил и мероприятий, реализация которых обеспечивает оптимальные условия для жизнедеятельности, укрепления здоровья и предупреждения заболеваний. Задачами курса являются:

- изучение закономерностей воздействия факторов окружающей среды на организм;
- выявление факторов риска и проведение гигиенической диагностики;
- разработка и внедрение нормативов по безопасности и безвредности факторов окружающей среды для организма;
- использование полученных теоретических и практических знаний в физкультурно-педагогической деятельности;
- разработка и внедрение мероприятий по оздоровлению населения и окружающей среды.

Для этого применяются следующие гигиенические методы: оптимизация условий и режимов труда и отдыха; рациональное питание; оптимизация двигательной активности; закаливание.

Разработанный тест включает в себя задания по всем темам курса, что позволяет установить уровень знаний обучающихся по предмету в целом. Данный вид контроля знаний ставит всех обучающихся в равные условия, используя

единую процедуру и единые критерии оценки, что приводит к снижению предэкзаменационных нервных напряжений.

Спецификация тестовых материалов для итогового теста по «Гигиене» определялась ФГОС СПО 4 по специальности 050141 «Физическая культура» (Приказ Минобрнауки от 05.04.2013 № 266 г.); рабочей учебной программой по дисциплине «Гигиена», утверждённой предметно-цикловой комиссией УОР №1 8 сентября 2013 года; учебным пособием для СПО медико-биологического профиля «Гигиена» авторы И.И. Бурак, В.П. Филатов, С.М. Соколов, Е.С. Зятиков 2 –е издание, Минск, 2006; учебным пособием «Гигиена», для студентов высших пед. учебных заведений под редакцией Я.С. Вайнбаума, В.И. Ковалёва, Т.А. Родионова М., «Академия», 2003г; учебно-методическим комплексом «Гигиена», разработанным Грязиной Г.И., утверждённым научно-методическим советом УОР №1 в 2013г.

Зачётная работа состоит из одной части, содержащей 30 заданий с выбором одного правильного ответа из четырех. Содержание теста представлено в приложении.

Распределение заданий зачётной работы по содержанию, видам умений и способам деятельности представлено в приложении (таблица 1), все задания базового уровня.

Распределение заданий по видам проверяемой деятельности - знания (63,3%) и понимание (36,7%) представлено в приложении (таблица 2)

Система оценивания выполнения зачётной работы – один балл за верное выполнение каждого задания. Минимальное количество баллов: 30% от максимального количества вопросов. На выполнение зачётной работы отводится (40 - 60 минут). Примерное время, отводимое на выполнение отдельных заданий 1-2 минуты. Дополнительные материалы и оборудование на зачёте не используются.

Тестирование проводилось среди студентов 2 курса училища в рамках итогового зачёта по предмету. Обобщённый план работы, представленный в приложении (см. таблицу 3) позволяет сделать вывод, что в тесте охвачены все темы курса «Гигиена».

Дата проведения зачёта 6 декабря 2014 года.

Время – 8.00-9.40утра (первая пара занятий). В тестировании должно было участвовать 2 группы по 10 человек. Приняло участие 20 человек. Состояние групп во время тестирования без особенностей. Никаких психологических нагрузок накануне не было. Обстановка во время проведения теста рабочая, позитивная. Студенты информированы о зачётных баллах. Группа студентов зимних видов спорта отсутствовала из-за нахождения на сборах и соревнованиях.

Процесс научного обоснования качества теста предполагает проведение вариационного анализа, построение вариационного ряда результатов и гистограммы частотного распределения результатов, оценку его надёжности и валидности, а также определение дифференцирующей силы заданий и проведение дистракторного анализа.

Тест считается надёжным, если он обеспечивает высокую точность измерений и если дает близкие результаты при повторном выполнении с условием, что подготовка ученика не изменилась за время повторного выполнения теста.

Надёжность теста определялась методом деления теста пополам т.к. тестирование проводилось однократно и на одной группе учащихся. Существуют

также метод параллельных форм (двукратное тестирование) и метод расщепления теста (однократное тестирование), метод Кьюдера-Ричардсона для дихотомических оценок по заданиям теста.

Один из аспектов применения коэффициента надежности связан с определением стандартной ошибки измерения.

Стандартная ошибка измерения трактуется как стандартное отклонение результатов испытуемого от его истинного балла, полученное при выполнении им большого числа параллельных форм теста. Неудовлетворительная надежность теста может быть при угадывании, отсутствии логической корректности, неоправданных выборах весовых коэффициентов, при короткой длине теста, отсутствии стандартных инструкций к тесту, малом количестве тестируемых. Тест менее 0,2 – низкая надежность; от 0,2 до 0,8 – средняя надежность; 0,8 и более – высоконадёжный тест.

Валидность – это очень важная характеристика качества теста, ориентированная на оценку адекватности теста поставленной цели его создания. Валидность – это характеристика способности теста служить поставленной цели измерения. Содержательная валидность – характеристика репрезентативности и содержания теста по отношению к запланированным для проверки знаниям и умениям. Для оценки валидности используют коэффициент корреляции, определяемый точечно-бисериальным методом. Коэффициент корреляции 0,5 позволяет сделать вывод о пригодности заданий для теста. Ниже этих цифр – рекомендуется убрать эти задания из теста.

Отрицательная корреляция – слабые учащиеся лучше выполняют задания, чем сильные. Такие задания надо проанализировать.

Дифференцирующая сила определялась методом контрастных групп. Выделяют 3 группы заданий. Задания, хорошо дифференцирующие тестируемых имеют показатель дифференцирующей силы от 0,6 и более; средний – от 0,3 до 0,6; плохо дифференцирующий – менее 0,3. Последние задания следует убрать из теста.

В нашем тесте были выделены 3 группы:

1. Задания, дифференцирующая сила которых близка к нулю (0-0,15). №№ заданий 1, 7, 9, 10, 11, 13, 21, 22, 23, 24. Эти задания были выполнены и сильными и слабыми учащимися. Для критериального теста это неплохо, а для нормативно-ориентированного теста такие задания надо из теста удалить.

2. Разность между сильными и слабыми учащимися особенно хорошо видна по заданиям №№9,10,24.

3. Отрицательная дифференцирующая сила (инвертированные задания), это задания, которые слабые ученики выполнили лучше сильных - №№ 4, 6, 16, 17. В качестве примера прокомментирую задания третьей группы:

Задание 4. По влиянию на человека климат делят на:

- 1) лечебный
- 2) закаливающий
- 3) тренирующий**
- 4) оздоровительный

Задание 6. Оптимальная температура, влажность и скорость движения атмосферного воздуха:

- 1) 20° С, 50%, 2,5м/с**

2) 14° С, 70%, 4м/с

3) 25° С, 30%, 5м/с

4) 22° С, 60%, 1м/с

Задание 16. Значение витамина А

1) источник жиров

2) обеспечивают процесс зрения

3) регулируют функции половых желёз

4) регулируют рост человека

Задание 17. Оптимальное распределение калорийности пищи по отдельным приемам при 4-х разовом питании:

1) 20-50-20-10%%

2) 10-50-30-10%%

3) 25-35-15-25%%

4) 10-25-50-15%

Далее проводился дистракторный анализ. Требования к дистракторам: рядоположенность и равновероятностная привлекательность. Рядоположенность определялась с помощью экспертного анализа.

Равновероятностная привлекательность состоит в том, что из 4 дистракторов один правильный, остальные должны быть распределены равномерно. По результатам тестирования получены первичные результаты, которые были проанализированы по классической тестовой технологии.

Дистракторный анализ выявил, что равновероятностная привлекательность состоит в том, что из 4 дистракторов один правильный, остальные должны быть распределены равномерно. Равновероятностная привлекательность позволила выявить проблемные дистракторы. В 12 заданиях слабо работают по 1 дистрактору. В задании №26 4-й ответ заменить «фитнес» на «группы здоровья». В задании №15 заменить ответ 2 – «безвредный» на «полноценный»

Задание №14 содержит лёгкий вопрос – рекомендуется его исключить из теста. Коррекцию в вопросе №24 ответ «количество учащихся» заменить на «количество параллелей классов». Проблемные дистракторы в №20 вопросе 3,4 заменить на «малую интенсивность» и «малый объём».

После обработки первичных результатов был проведён вариационный анализ.

Для проведения вариационного анализа использовались индивидуальные баллы и количество учащихся, получивших этот балл (приложение таблица 4).

Из вариационного ряда видно, что тест не работает в области низких баллов (от 0 до 8). Основной контингент учащихся находится в области от 12 до 27 баллов. Учащиеся хорошо сгруппированы. Максимальный балл при тестировании 25 получили 3 тестируемых.

Более наглядно выглядят результаты тестирования на гистограмме. Тест одномодальный (мода 25), асимметрия правосторонняя. В соответствии с формулой Колмогорова – Смирнова распределение результатов не соответствует закону нормального распределения, так как не соответствует нормальному распределению Гаусса. Это в первую очередь связано с малым количеством тестируемых.

По результатам тестирования было получено частотное распределение результатов (см. приложения рис.1 «Гистограмма частотного распределения результатов»).

Рассчитаны показатели вариационного анализа, такие как среднее выборочное, размах, стандартное отклонение, асимметрия, эксцесс и надежность теста. Эти показатели представлены в приложениях (таблица 5).

Средний балл теста равен 18,75, а стандартное отклонение 4,91, т.е. $18,75 \pm 4,91$ это результаты 95% тестируемых находятся в этих пределах. Размах теста 18 – достаточный. Значения средних результатов смещены вправо – это может свидетельствовать о том, что тест лёгкий или учащиеся хорошо подготовлены в соответствии с нормативно-ориентированным тестированием. Это же подтверждает и асимметрия, смещённая в сторону больших значений. Эксцесс характеризует размах. Показатели размаха, дисперсии и эксцесса имеют небольшие значения.

Проанализировав полученные результаты, можно сделать вывод, что группа однородна, так как результаты тестирования сгруппированы. Данный тест хорошо дифференцирует группу. Надежность теста 0,75- средняя.

Характеристика заданий теста проводилась на основе анализа показателей точечно-бисериального коэффициента корреляции и анализа дифференцирующей силы заданий.

Анализ значений коэффициента точечно-бисериальной корреляции (см. приложения, таблица 6) выявил ряд заданий (№№ 10, 13, 17, 24) со значениями от 0,05 до 0,2. Эти задания из теста необходимо удалить, как неудачные. Коэффициент от 0,5 до 0,75 присутствует у большинства заданий, это позволяет сделать вывод о пригодности заданий для теста. Из теста выбиваются и задания №№ 4, 6, 7, 17, они имеют отрицательную корреляцию

(-0,1;-0,2). Это задания со сложной структурой. Более слабые учащиеся отвечают на них лучше.

Доля правильных ответов лежит в области средних величин от 0,3 до 0,7, из 30 заданий – это 26 (86%).

Задания №1, 14, 22, 23, 24 лежат в области от 0,75 до 0,9 – это лёгкие задания (14%). 4 задания являются инвертированными №4, 6, 7, 16. Слабые ученики на эти задания отвечают лучше. Эти задания в дальнейшем следует доработать.

Коэффициент дискриминативности (см. приложения, таблицу 7) составляет в 3-х заданиях 0,00, еще 2 задания имеют отрицательный коэффициент дискриминативности. В целом эффект зависит от корреляции между заданиями и от числа заданий. Мы вновь обращаем внимание на малое число тестируемых. Обобщённый анализ представленной спецификации и статистических данных показал:

1. Тест критериальный. Для аттестации по предмету необходимо ответить на 30 % вопросов (9 баллов). Все студенты аттестованы.

2. Вариативный анализ (частотное распределение результатов) учащихся не соответствует нормальному, т.к. количество тестируемых мало, всего 20 человек. Анализ характеристик теста показал, что группа однородна. Тест одномодальный, (максимальное значение 25 баллов), асимметрия смещена вправо в сторону больших баллов. Результаты тестирования сгруппированы, средняя группа с баллами от 12 до 24 – большая. Стандартное отклонение $18,75 \pm 4,91$ говорит о том, что 95% результатов лежат в этих пределах. Данный тест хорошо дифференцирует группу (размах 18). Доля правильных ответов лежит в диапазоне от 0,3 до 0,7, что

говорит о средней сложности заданий, лишь 2 задания имеют показатель 0,9 (№№1 и 14). Тест средней сложности, учащиеся хорошо подготовлены.

3. Надежность теста 0,75 – средняя.

4. Анализ показателей качества тестовых заданий и дистракторов выявил наличие лёгких, инвертированных заданий и заданий, где дистракторы не работают. Эти задания требуют коррекции.

5. В тестировании приняло участие 20 человек, что недостаточно для полноценного статистического анализа. Специфические особенности нашего учебного заведения (многочисленные учебно-тренировочные сборы и соревнования) не позволяют одномоментно протестировать весь 2 курс, а по классической тестовой технологии это является одним из условий тестирования.

6. На основании анализа была проведена коррекция отдельных заданий теста.

СПИСОК ЛИТЕРАТУРЫ

1. Аванесов В.С. *Композиция тестовых заданий*. 2-е изд. исправл. и доп. М.: Адепт. 1998. С. 217
2. Аванесов В.С. *Применение заданий в тестовой форме в новых образовательных технологиях*. testolog.narod.ru (дата обращения 15.10.2014)
3. Аванесов В.С. Проблема объективности педагогических измерений//*Педагогические Измерения*.2008. №3. - С.3- 40
4. Аванесов В.С. Проблема эффективности педагогических измерений//*Педагогические Измерения*.2008, №4. - С.3-40
5. Аванесов В.С. Тест как педагогическая система // *Педагогические Измерения*.2008. №1. - С.33-35
6. Балыхина Т.М. *Словарь терминов и понятий тестологии*. М.: Изд. РУДН. 2000. С.164
7. Беспалько В.П. *Педагогика и прогрессивные технологии обучения*. М.: Изд-во ин-та профессионального образования МО РФ. 1995. С. 338
8. Бурак И.И., Филонов В.П., Соколов С.М., Зятиков Е.С. *Гигиена учебное пособие для медицинских училищ*. Минск, Вышэйшая школа. 2006.
9. Вайнбаум Я.С., Коваль В.И., Родионова Т.А. *Гигиена физического воспитания и спорта, учебное пособие для студентов высших учебных заведений*. М., Академия. 2003.
10. Васильев В.И., Тягунова Т.Н. *Культура компьютерного тестирования. Часть II. Программно-дидактическое тестовое задание*. М.: МГУП. 2002.С. 90
11. Гласс Д., Стэнли Д. *Статистические методы в педагогике и психологии*. Пер. с англ. М.: Прогресс. 1976.
12. Ефремова Н.Ф. *Современные тестовые технологии в образовании. Учебное пособие*. М.: Логос. 2003. С.176
13. Кларин М.В. *Инновационные модели в зарубежных педагогических поисках*. М.:1994.
14. Михеев В.И. *Методы теории измерения в педагогике: Учебное пособие*. – М.: Логос. 2003. С. 64
15. Переверзев В.Ю. *Критериально-ориентированное педагогическое тестирование: Учебное пособие*. М.: Логос. 2003.С.120
16. Родионов Б.У., Татур А.О. *Стандарты и тесты в образовании*. М. МИПКРО. 1995. С68
17. Савельев Б.А., Масленников А.С. *Оценка уровня обученности студентов в целях аттестации образовательного учреждения профессионального образования: Учебное пособие*. М.: Логос. 2003. С.136
18. *Тестовая технология контроля знаний*. - Сост. А.В. Мальцев Екатеринбург, ИРРО. 1996. С.39
19. Чельщикова М.Б. *Теория и практика конструирования педагогических тестов: Учебное пособие*. М.: Логос. 2002. С.432.