

Основы педагогических измерений для учителей информатики

В стандарты педагогического образования введены новые курсы общепрофессиональной дисциплины (ОПД.Ф.08) «Современные средства оценивания результатов обучения» для подготовки будущих учителей. Данная дисциплина реализует подготовку учителя в области теории и практики педагогических измерений с использованием компьютерных технологий. Курс предполагает изучение вопросов прикладной направленности: различать контрольный материал по цели использования, проводить корректную интерпретацию результатов тестирования, понимать основные характеристики качества контрольного материала и возможности использовать компьютерные технологии при создании, проведении и обработке результатов тестирования.

Начиная разговор о педагогических измерениях в образовании необходимо сделать обзор некоторых понятий, которые будут использоваться в дальнейшем.

Понятие «педагогический контроль» можно рассматривать как единую дидактическую и методическую систему проверочной деятельности. Применительно к повседневному учебному процессу проверочная деятельность направлена на выявление и оценку результатов учебной деятельности школьников или студентов. В действительности, несмотря на существующие общие критерии выставления оценок, оценка остается субъективным мерилем учебных достижений. В педагогическом контроле нет выделенного объекта измерения, нет точки отсчета (абсолютного нуля), нет единиц измерения. Это создает определенные трудности и ряд проблем, которые специалисты пытаются решать, опираясь на математическую теорию измерений. Возникает проблема объективного измерения результатов учебной деятельности. «Измерение предполагает проведение объективного количественного сопоставления оцениваемого свойства ученика с некоторым эталоном, принятым в качестве единицы измерения. При педагогическом измерении роль оцениваемого свойства отводится знаниям, умениям и навыкам ученика, так традиционно сложилось в отечественной практике, а вместо единицы измерения используются контрольные задания или части заданий по проверяемому содержанию предмета». [3] Для педагогических измерений необходимо определить: объект измерения (переменные измерения), измерительную процедуру, средства измерения и шкалу.

Главная проблема - выделить переменную измерения. Переменные измерения в педагогике не носят явного характера. Такие переменные называют латентными (неявными), т.е. скрытыми от возможности непосредственного измерения. Переменная может быть выражена в виде совокупности заданий, и проявляться только в эмпирических результатах выполнения теста, путем анализа.

Процедура измерения – совокупность операций, позволяющая перейти от эмпирических референтов (заданий) к числовым оценкам измеряемых характеристик [3].

Измерительный инструмент состоит из двух компонент: теста и шкалы для фиксации результатов измерения. Педагогический тест – это инструмент, состоящий из квалитметрически (с использованием различных методов определения качества) выверенной системы тестовых заданий, стандартизированной процедуры проведения и заранее спроектированной технологии обработки и анализа результатов, предназначенный для измерения качеств и свойств личности, изменение которых возможно в процессе систематического обучения. [1]

Рассмотрим понятие шкалы. «Шкала – числовая система, в которой отношения между различными свойствами изучаемых явлений, процессов переведены в свойства того или иного множества, как правило – множества чисел». [2] Выделяют шкалы отношений, интервальные шкалы, порядковые (ранговые) шкалы и номинальные шкалы (шкалы наименований).

Увеличение мощности ↑	Название	Особенность
	Шкала отношений	Имеет естественное начало отсчета, но нет естественной единицы измерений.
	Интервальная шкала	Начало отсчета и интервалы выбираются по согласованию относительно выбранного параметра. Шкала растягивается и сжимается.
	Порядковая шкала	Нет начала отсчета и единицы измерения.
	Шкала наименований	Неметрическая

1. **Шкала наименований** представляет номинальный уровень, не связана с понятием «величина», т.е. неметрическая, используется для того, чтобы отличать один объект от другого (фамилии учеников, номера телефонов и пр.). Не допускается никаких арифметических действий над элементами такой шкалы. Возможен лишь подсчет количества объектов с совпадающими признаками. Для школьной практики это означает возможность подсчета количества отличников, «хорошистов» или двоечников. Возможно лишь сравнение этих групп по количеству учащихся.

2. **Порядковая шкала** (ранговая, шкала упорядоченной классификации) представляет собой ординальный уровень. Когда объекты измерения разбиваются на группы соответствующие определенной точке шкалы. Такая шкала не имеет начала отсчета и единиц измерения, она только упорядочивает объекты, приписывая им те или иные ранги. Относительно значений шкалы нельзя говорить ни о том, во сколько раз измеряемая величина больше или меньше другой, ни на сколько она меньше или больше. Не допускается никаких арифметических действий со значениями. Допустима замена чисел, характеризующих величину признака, другими, не нарушающими ранговый порядок объектов.

Для использования в школе этой шкалы корректно также следует помнить, что можно подсчитать лишь количество учащихся, получивших различные оценки. Сравнить качество подготовки вычислением среднего балла недопустимо.

3. **Интервальная шкала** (интервальный уровень измерения) основана на сравнении различия между объектами по величине измеряемого признака или свойства, характеризуется тем, что не имеет естественного начала отсчета и единиц измерения. Начало отсчета и интервалы выбираются по согласованию относительно выбранного параметра. Шкала растягивается и сжимается. Позволяет просчитать среднее арифметическое и среднее квадратичное отклонения и коэффициенты корреляции. Современная теория тестов позволяет преобразовать результаты в интервальную шкалу.

Для школьной практики это новый формат оценивания, который может прижиться при активном внедрении в практику тестовых форм контроля качества подготовки выпускников.

4. **Шкала отношений** – самая мощная шкала. Она позволяет оценивать, во сколько раз один измеряемый объект больше (меньше) другого объекта, принимаемого за эталон. Имеет естественное начало отсчета, но нет естественной единицы измерений. Шкалами отношений измеряют почти все физические величины. Но ею нельзя пользоваться в социальных измерениях.

Говоря о качестве педагогических измерений, подразумевают, прежде всего, объективность. Существует несколько градаций объективности.

Процедурная объективность отождествляется с объективностью процедуры тестирования (удален педагог, одинаковые условия для всех, единые критерии проверки и пр.).

Классическая объективность – описывается следующей формулой:

$$X_{\text{(наблюдаемый балл)}} = T_{\text{(истинный балл)}} + E_{\text{(ошибка измерения)}}$$

T (истинный балл) – это (оценка) параметр испытуемого, не зависящий от средства и способа измерения, но меняющийся в процессе обучения. **T** постоянна в момент измерения. Чем меньше ошибка измерения, тем измерение объективнее. **T** – абсолютно объективная оценка. Совокупность заданий должна работать на оценивание переменной измерения. Оцениваемая переменная должна быть доминирующей. Классическая объективность связана с величиной ошибки измерения, с содержанием и процедурными проблемами. [3]

Понятийный аппарат тестологии только формируется, поэтому возможно использование специалистами различных трактовок одного и того же понятия. Понятие «педагогический тест» может использоваться в двух смыслах, как метод педагогического измерения и как результат применения теста.

В данной статье не будут рассматриваться различные классификации тестов, поскольку речь идет только о тех, которые используются в системе образования.

По целям использования в системе образования различают:

(классификация по книге Нормана Е. Гронлунда, названия даны в переводном варианте, возможно, по мере внедрения в педагогическую практику они претерпят изменения).

Тесты (классификация по целям использования)			
Определяющий	Формирующий	Диагностический	Суммирующий
<ul style="list-style-type: none"> • Входной; • Предварительный. 	<ul style="list-style-type: none"> • Обучающий; • Тематический; • Текущий. 	<ul style="list-style-type: none"> • Корректирующий; • Промежуточный. 	<ul style="list-style-type: none"> • Итоговый; • Аттестационный.
Использование в педагогической практике			
На всех ступенях обучения.	На всех ступенях обучения.	Практически не используется.	Начиная с основного общего образования.

Следует подчеркнуть следующий факт, что каждый из этих видов тестов имеет свои особенности, которые еще не до конца изучены. Использование одних тестов вместо других может привести к серьезным негативным последствиям. Вместе с тем, тесты необходимо интегрировать в учебный процесс. Тесты обеспечат объективность контроля, разнообразят учебную деятельность учащихся, сэкономят колоссальное количество времени, что очень важно при перегрузке учителей и учащихся. Знания основ педагогических измерений необходимы каждому педагогу, только тогда возможно говорить о корректном использовании тестового материала и процедур в учебном процессе.

Эффект от использования тестов в обучении любому предмету существенен в случае использования компьютера. Компьютерное тестирование по предметам как один из элементов системы внутришкольного контроля открывает широкие возможности для индивидуализации процесса усвоения знаний. Процесс обучения и контроля за обучением должен быть един. Компьютерное тестирование – это реальная возможность решения целого ряда задач, способствующих повышению качества обучения не за счет увеличения нагрузки педагога, а путем внедрения в учебный процесс обучающих программ и тестов, которые способны изменить технологию обучения.

Все существующие на сегодняшний день системы компьютерного тестирования имеют незавершенный вид. К полноценным программным продуктам, готовым к распространению и использованию в педагогической практике их отнести нельзя. Как правило, тестирующие системы имеют возможность использования ограниченного типа тестовых заданий, в основном задания с выбором одного или нескольких правильных ответов. В тестирующих системах нет четкого разграничения обучающих и контролирующих тестов, а диагностические тесты отсутствуют совсем.

Создание качественных итоговых тестов по всему предметному курсу – дело специалистов соответствующих центров. Учитель предметник должен разбираться в многочисленном тестовом материале и уметь использовать готовые тематические тесты в соответствии с их назначением. Поскольку в нашей стране «спасение утопающих – дело

самых утопающих», учителя-предметники заинтересовались возможностью самим составлять тесты контроля знаний, поэтому весьма полезными будут знание основных критериев качества, предъявляемых тестовым заданиям. Подготовку тестовых материалов учителю следует начинать с тематических (обучающих, текущих) тестов. Это наиболее легкий этап, поскольку область проверки ограничена одной темой, содержание проще конкретизировать, задания будут разной трудности их не сложно расположить по возрастанию трудности. Для определения качества теста можно вручную подсчитать некоторые параметры.

Существуют некоторые общие правила, на которые возможно следует опираться во время работы над тестом. Различными авторами предлагаются различные классификации видов тестовых заданий. Рассмотрим наиболее употребительные в практике тестирования.

Предварительно можно выделить четыре основные группы заданий тестовой формы:

- **закрытой формы или с выбором одного правильного ответа из нескольких предложенных,**
- **задания на дополнение или открытой (свободной) формы,**
- **на установление соответствия,**
- **на установления правильной последовательности.**

Выбор формы тестового задания зависит от цели контроля (сюда же относится форма предъявления: бланковая или компьютерная) и содержания контролируемого материала. За рубежом наибольшее распространение получила закрытая форма тестовых заданий, во многом из-за преимуществ обработки (легко автоматизировать).

1. Наиболее распространенными и в нашей стране являются **задания закрытой формы или с выбором одного правильного из нескольких предложенных вариантов ответов.**

Именно задания закрытой формы отождествляют со всеми тестовыми заданиями и обрушивают часто необоснованную критику. Основным недостатком таких заданий считают возможность угадать правильный ответ. Вероятность угадывания можно снизить подбором не менее четырех ответов (используют термин дистракторы), использованием фасетов заданий (способов записи нескольких вариантов одного и того же задания), жестким временным ограничением на выполнение заданий. Важным достоинством таких заданий можно считать универсальность в использовании. Содержание практически любой дисциплины можно переложить для тематического контроля с помощью заданий закрытого типа. Измеряемый уровень знаний зависит в основном от сложности содержания контрольного задания, не от формы его представления.

Отдельного внимания заслуживает количество дистракторов (отвлекающие или неправильные ответы). Основное требование к качеству дистракторов, чтобы они были правдоподобны и привлекательны для выбора, но не допускали неоднозначного толкования и не были частично правильными.

В условиях активного распространения массового аттестационного тестирования, когда необходимо привлечь новые информационные и компьютерные технологии для проведения и обработки результатов у заданий данного типа большие перспективы. Они будут совершенствоваться и видоизменяться.

Например:

1. Сигнал, который может иметь только 2 состояния, передаёт количество информации равное

1) 1 биту

2) 2 байтам

3) 8 битами

4) 4 байтам

Задание для проверки понимания сущности единицы измерения информации. Может присутствовать в тематическом тесте, в итоговом оно нежелательно. Его место в начале

обработки результатов, поэтому в настоящее время они не используются в процедурах централизованного тестирования и ЕГЭ.

Следующий тип тестовых заданий – **с открытым (свободным) ответом или задания на дополнение**. Если задания с выбором ответа можно критиковать за возможность угадать правильный ответ, то здесь догадка исключена, поскольку в место для ответа его нужно вписать.[4]

Задания этого типа могут быть разной сложности. Первый тип заданий с ограничением на ответ по количеству и виду символов. Это важно для исключения многословности и двусмысленности в ответах, правильность выполнения заданий тогда может быть однозначно определена.

Например:

В место для ответа впишите соответствующее число.

1. Наименьшее количество двоичных разрядов, необходимое для кодирования алфавита из 5 букв, равно _____.

Задание равно хорошо использовать как в тематическом так и в итоговом тесте. Время выполнения не более 1 минуты.

Использование заданий тестовой формы с открытым (свободным) кратким ответом актуально для проведения текущих проверок усвоения предметной терминологии, определений, дат, формул, небольших логических или расчетных задач. В условии задания легко использовать принцип **фасетности** для построения нескольких вариантов. В итоговом контроле такие задания должны составлять отдельную группу, так как требуют отдельной обработки результатов. Задания этого типа могут быть обработаны компьютером при небольшой модификации ответов.

Другой тип этих заданий позволяет свободно конструировать развернутый ответ без жестких ограничений по его объему и содержанию. Это значит, по условию требуется представить полное решение задачи, обоснование ответа, написать небольшое сочинение – эссе. В большинстве своем проверяют сложную структуру знаний по предмету. Аналог традиционных письменных экзаменов по любому предмету, но с заданными критериями для проверки. Для проверки таких заданий разработчики составляют критерии, позволяющие определить правильность выполнения. Критерии это не точные образцы ответов, а перечисление некоторых обязательных условий, при выполнении которых ответ можно засчитать как правильный. Не поддаются компьютерной обработке. Решения проверяют эксперты, не менее двух, для снижения фактора субъективизма при оценке.

Например: Квадратный массив A размера $N \times N$ (где $N=5$) в программе сначала обнуляется, а затем производится некоторая дальнейшая обработка в соответствии с программой. Текст программы приведен ниже.

Требуется выяснить состояние массива A по окончании работы программы, а именно:

1) Какое число будет стоять в элементе массива $A[1,2]$?

2) Сколько всего ненулевых чисел будет в массиве A ?

а также ответить на следующий вопрос:

3) В программе после начального обнуления массив A заполняется с помощью следующих операторов

```
i:=1; j:=1; A[1,1]:=1; k:=1; while (i<N) OR (j<N) do
begin if (i+j) mod 2 = 0 then j:=j+1 else i:=i+1;
k:= -k; A[i,j]:=k;
end;
```

На какой более короткий фрагмент программы можно заменить приведенную часть программы, обеспечивая такое же заполнение массива A ?

Программа на языке Паскаль	Программа на языке Бейсик
CONST N=5;	N=5
VAR i, j: integer; {i - строка j - столбец}	DIM A(N, N) AS INTEGER

<pre> k: integer; A: array[1..N,1..N] of integer; BEGIN for i:=1 to N do for j:=1 to N do A[i,j]:=0; i:=1; j:=1; A[1,1]:=1; k:=1; while (i<N) OR (j<N) do begin if (i+j) mod 2 = 0 then j:=j+1 else i:=i+1; k:= -k; A[i,j]:=k; end; END. </pre>	<pre> FOR I=1 TO N FOR J=1 TO N A (I, J) = 0 NEXT J NEXT I I=1 J=1 A(1, 1)=1 K=1 WHILE (I<N) OR (J<N) IF (I+J) MOD 2 = 0 THEN J=J+1 ELSE I=I+1 ENDIF K=-K A(I,J)=K WEND END </pre>
---	--

Ответ (критерии оценивания):

Содержание верного ответа и указания по оцениванию (допускаются иные формулировки ответа, не искажающие его смысла)	Баллы
<p>Элементы ответа:</p> <p>1) $A[1,2] = -1$</p> <p>2) В массиве A будет 9 ненулевых элементов (5 значений "1" по главной диагонали и 4 значения "-1" над главной диагональю)</p> <p>3) Доработку можно провести следующим образом:</p> <pre> for i:=1 to N do begin A[i,i]:=1; if i<N then A[i,i+1]:=-1; end; </pre>	
Ответ правильный и полный, включает все названные выше элементы	3
Правильно записаны только два из названных выше элементов ответа	2
Правильно записан один из названных выше элементов ответа	1
Все элементы ответа записаны неверно	0
<i>Максимальный балл</i>	3

Примечание: баллы выставляются независимо за каждый из правильно выполненных 3-х пунктов задания.

Задания на установление соответствия позволяют проверить знания взаимосвязи определений и фактов, сущности и явлений, соотношения между объектами и их свойствами, законами и формулами. Следует отметить, что эти задания используются в курсе информатики редко из-за громоздкой технологии подготовки самих заданий и из-за невозможности использовать подготовленные бланки с заданиями многократно. Поэтому их лучше всего применять в текущем контроле знаний или для самоконтроля. В настоящее время они все же стали входить в итоговые контрольно-измерительные материалы (пока не по информатике), но отдельным блоком в конце работы.

Например:

Соедините стрелками числа, обозначающие одно и то же количество, но в разных системах счисления. Время выполнения 3-5 минут. Исправления исключаются.

1 вариант

2 вариант

X_2	X_{10}	X_2	X_{10}
1000	12	11	5
10	8	1001	3
1100	2	101	9
110	4	1110	7
1111	6	111	13
100	15	1101	14

Задания на установление правильной последовательности проверяют знания определенной последовательности действий для получения конечного результата или знание хронологии событий. Ниже приведены примеры из демонстрационной версии единого государственного экзамена.

Например: Доступ к файлу http.txt, находящемуся на сервере www.net осуществляется по протоколу ftp. В таблице фрагменты адреса файла закодированы буквами от А до Ж. Запишите последовательность этих букв, кодирующую адрес указанного файла.

А	://
Б	http
В	ftp
Г	.net
Д	.txt
Е	/
Ж	www

Ответ: _____

или ответ может быть занесен в таблицу, тогда его можно проверить в автоматизированном режиме (с помощью компьютера):

--	--	--	--	--	--

Литература:

1. Майоров А.Н. Теория и практика создания тестов для системы образования. – М.: Интеллект-центр. 2002. – 296 с.
2. Новиков Д.А. Статистические методы в педагогических исследованиях. М.; МЗ-Пресс, 2004. - 160 с.
3. Чельшкова М.Б. Теория и практика конструирования педагогических тестов. Учебное пособие, М., Логос, 2002. - 432 с.
4. Самылкина Н.Н. Построение тестовых заданий по информатике. Методическое пособие. М.: Бином, 2003, 176 с.
5. Самылкина Н.Н. Методические рекомендации по оцениванию заданий с развернутыми ответами. Информатика. ЕГЭ 2005. - М.: "Уникум-Центр", 2005, 160 с.