

УДК 378.091.398

О КОМПЕТЕНТНОСТИ И ОБЪЕКТИВНОСТИ ПРИ ТЕСТИРОВАНИИ ЗНАНИЙ СЛУШАТЕЛЕЙ ФАКУЛЬТЕТА ПОВЫШЕНИЯ КВАЛИФИКАЦИИ

В. П. Симонов

Аннотация. Статья посвящена актуальной проблеме взаимосвязи достоверности тестирования и компетентности лиц, готовящих и проводящих данный вид контроля за степенью обученности слушателей Факультета повышения квалификации. В ней впервые характеризуется процесс тестирования с позиции взаимосвязи получаемых оценок с компетентностью и объективностью разработчиков тестовых материалов. Приведен пример итога экспериментальной работы в вузе по данной проблеме.

В статье сделана попытка выработать рекомендации по созданию единой технологии, процедур и инструментария для более достоверной оценки учебных достижений обучающихся на ФПК с помощью тестирования, для чего осуществлен анализ основных недостатков. Обращается внимание на то, что самое главное – это высокая компетентность (обученность, эрудиция и личный опыт) составителей тестов, которые должны в первую очередь определиться с точностью шкал измерения. В связи с подписанием нашей страной Болонского соглашения, в статье приводится сравнительная таблица соответствия десятибалльной шкалы с существующими на западе оценочными шкалами и обосновывается их совместное применение в дальнейшем.

The article is devoted to the topical problem of the relationship of reliability testing and competence of the persons preparing and implementing this type of control for the degree of command of students of the Faculty of professional skill improvement. The process of testing is considered as interrelation of marks, competence and objectivity of the developers of test materials for the first time. The example of the outcome of experimental work in the University is given in the article.

The author of the article makes an attempt to develop recommendations on the creation of a unified technology, procedures and tools for a more reliable assessment of academic achievements of

students at the Faculty of professional skill improvement with the help of testing, based on the analysis of the major drawbacks. It is paid special attention to the fact that the most important thing is a high level of expertise (training, knowledge and experience) of tests drafters, which should first of all determine the accuracy of the measurement scales. In connection with the signing of our country of the Bologna agreement, the article gives a comparative table of compliance of the ten point scale with the existing in the West estimation scales and justification of their joint use in the future.

Ключевые слова: обученность, распознавание (различение), запоминание, понимание, виды тестирования, достоверность оценки, полностью положительная десятибалльная шкала, Болонский процесс.

Training, recognition (differentiation), memory, understanding, kinds of testing, reliability of the assessment, completely positive ten point scale, the Bologna process.

Широчайшее распространение, которое за последние десятилетия в образовании России получило тестирование, актуализировало проблему достоверности его результатов, а также степени компетентности и объективности лиц, создающих и использующих тестовые материалы.

Цель тестирования в образовании – осуществлять текущий или итоговый контроль хода и результатов образовательного процесса одновременно у любого количества обучаемых на основе вопросов (заданий) на уровнях «запоминания» и «понимания», и практических заданий – на репродуктивном уровне, а также вносить в него на этой основе изменения, способствующие повышению качества общеобразовательной или специальной подготовки личности, т. е. максимально достоверная оценка уровня и качества сформированных у нее знаний, умений и навыков (образовательных компетенций) в конкретной области, необходимых

либо для дальнейшего обучения, либо для осуществления какой-либо профессиональной деятельности.

Основные задачи тестирования – это, прежде всего, достоверное определение уровня и качества усвоенных знаний, умений и навыков в конкретной области, а также оценка сформированных уровней компетентности, необходимых для дальнейшего обучения или будущей профессиональной деятельности. Достоверное тестирование не только выявляет пробелы, но и образует рефлексивную основу для самооценки; способствует обобщению и систематизации знаний данной личности, а также позволяет выделить главное, основное в изучаемом материале. Анализ практики последних лет показывает, что чаще всего наблюдаются недостатки при подготовке и проведении тестирования, выражающиеся в нарушении обучающей функции проверки знаний при проведении тестирования проявляется в том, что часто составители тестов наряду с существующими фактами и научными положениями придумывают свои несуществующие, т. е. антинаучные, для того, чтобы уменьшить вероятность угадывания правильного ответа и, тем самым, искажают принцип научности обучения. Выход из

ний и умений не только не нарушается, а наоборот – реализуется в должной мере.

Рассмотрим основные недостатки при формировании тестов:

– во-первых, многие их составители слабо представляют, какой показатель степени обученности они хотят проверить с помощью этого теста. Часто вообще не указывается – какую оценку можно выставить за выполнение того или иного задания на основе единой, например, полностью положительной, математически обоснованной, ($Y=kX^2$) 10-балльной шкалы:

– во-вторых, большинство тестов с выборочным ответом «работают» на первом уровне «различения», «распознавания» – уровень «знакомства» по нашей концепции, что не позволяет говорить об обученности человека более чем на 4% от возможного. Или такие тесты, когда надо, например, указать одну правильную дату среди четырех-пяти предъявленных, ибо в этом случае проверяется лишь чисто механическое «запоминание» – второй уровень (обученность максимально 16%) (рис. 1);

– в-третьих, при предъявлении подобных тестов нарушается важнейший принцип проверки и оценки знаний – это принцип научности, суть которого заключается в том, что нель-

З Н А Н И Я			IV. Элементарн	V. Перенос
ВОСПРОИЗВЕДЕНИЕ			умения и навыки	
		III. Понимание		(творческий уровень)
	II. Запоминан.			
I. Различение			УМЕНИЯ И НАВЫКИ	
4%	12%	20%	28%	36%
1	3	5	7	9
Обученность от 1 до 8%	Обученность от 9 до 24%	Обученность от 25 до 48%	Обученность от 49 до 80%	Обученность от 81 до 100%
Плохое качество (ПК)	Низкое качество (НК)	Среднее качество (СК)	Хорошее качество (ХК)	Высокое качество (ВК)
Оценки в баллах 10-ти балльной шкалы и их соответствие степени обученности				
1 б. – 1% и >	3 б. – 9% и >	5 б. – 25% и >	7 б. – 49% и >	9 б. – 81% и >
2 б. – 4% и >	4 б. – 16% и >	6 б. – 36% и >	8 б. – 64% и >	10 б. – 100%

Рис. 1. Уровни теста

этого положения, заключается в том, чтобы давать задание по нахождению одного неверного ответа среди других трех-четырех верных. В этом случае тестируемый обучается только верному, а обучающая функция проверки зна-

зя предъявлять для «запоминания», а такое вполне реально особенно при текущем тестировании, неверные, антинаучные сведения.

Ниже мы приводим примеры тестов на нахождение одного неправильного среди других

четырёх правильных ответов, а также тест на соответствие, позволяющий оценивать не «различение» и «распознавание», а «понимание» сути изученной теории, что и является вершиной такого понятия, как «знания» и не нарушает принцип научности при проверке и оценке степени обученности слушателя ФПК.

Рассмотрим с этих позиций существующую классификацию тестовых заданий [1, 3, 4, 5]:

1. Тестовые задания открытого типа, где не используются готовые варианты ответов и которые формулируются в форме вопросов или высказывания, а тестируемым предлагается продолжить или вставить недостающую часть утверждения. Например, «Раздел педагогики занимающийся проблемами обучения называется...» (надо дописать «дидактика»).

Достоинства данного вида тестирования в том, что здесь нет готового ответа, который тем самым лишает проверяемых возможности угадать правильный ответ (думают сами). В этом случае проверяются такие показатели обученности как «запоминание» или «понимание» (максимальная оценка до «4» или «6» баллов соответственно, по 10-балльной шкале).

К недостаткам здесь можно отнести краткие формулировки. Задания этого типа проверяют умение самостоятельно сформулировать краткий ответ, который необходимо выразить соответствующей записью в виде слова (как в приведенном примере); или словосочетания (например, дистанционное обучение); набора букв или цифр (например, 1380 – дата Куликовской битвы).

Такие тесты можно разделить:

– на задания с кратким ответом (проверяется второй уровень усвоения – «запоминание», максимальная оценка до четырех баллов). Пример: «Когда в российском образовании применялось полностью безотметочное обучение во всех типах учебных заведений?» Ответ: «С 1918 по 1935 год»;

– и с развернутым ответом, которые являются заданиями более высокого уровня сложности (в этом случае проверяется третий уровень усвоения – «понимание» – максимальная оценка до шести баллов по 10-балльной шкале). Пример: «Что невозможно проверить с помощью тестирования?» Ответ: «Умение применять теорию на практике творчески».

Включение таких заданий в тест повышает защищенность экзамена от возможности подсказок и угадывания.

Задания с развернутым ответом включают в себя:

– задания на перечисление признаков какого-либо явления, объектов одного класса и т. д. Например, «понятие «знания» включает в себя следующие составляющие...». (Надо дописать: различение (узнавание), запоминание и понимание). В этом случае проверяется «запоминание» (максимальная оценка до четырех баллов, по 10-балльной шкале);

– задания, требующие раскрытия какого-либо теоретического положения (понятия) на конкретном примере. Например, «Дайте определение понятию „воспитание“ в широком его толковании». Ответ: «Воспитание в широком значении подразумевает передачу опыта старших поколений подрастающему». Здесь проверяется такой показатель обученности, как «понимание», максимальная оценка в этом случае до шести баллов;

– задания-задачи, содержащие условие в виде предписания: например, решить уравнение $4 \times 2 = 1$. При таком виде задания проверяется, помимо «понимания» и четвертый уровень – «элементарные умения и навыки». Или другой пример подобного задания: «Вставьте недостающие буквы в слово „инт_лл_генция“». В обоих случаях максимальная оценка до восьми баллов.

2. Тестовые задания закрытого типа, состоящие из основного текста (может быть дополнен чертежами, графиками и т. п.) и различных вариантов одного или нескольких ответов. По типу высказывательные или повелительные, они легко формулируются, а также легко понимаются испытуемыми. Пример: «В какой стране не применяется десятибалльная шкала оценки обученности?»

- а) Республика Беларусь;
- б) Латвия;
- в) Молдова;
- г) Украина;
- д) Эстония. (Ответ – Украина)».

Основной недостаток – возможность угадывания ответа, к тому же в этом случае проверяется максимально только второй уровень усвоения – «запоминание» (оценка до четырех баллов), а чаще всего лишь «различение,

распознавание» – т. е. первый уровень – «знакомство», максимальная оценка до двух баллов по 10-балльной шкале. Например: «Какой из нижеуказанных способов привлечения внимания обучающихся в ходе занятий К. Д. Ушинский не рассматривал?

- а) усиление впечатления;
 - б) прямое требование внимания;
 - в) наказание за невнимательность;
 - г) занимательность преподавания;
 - д) меры против рассеянности
- (Правильный ответ – позиция *в*)».

Такое задание предпочтительнее, т. к. здесь не предъявляется для усвоения неправильное, т. е. не нарушается принцип научности обучения (проверка знаний входит в понятие «обучение»).

3. Задания на установление соответствия (элементы одного множества требуется поставить в соответствие элементам другого множества). Эти задания активизируют самостоятельную работу испытуемых, а также исключают условия для списывания и позволяют проверить такой важный (третий) показатель обученности, как «понимание», т. е. осознанность имеющихся у обучающегося знаний – максимальная оценка до шести баллов.

Пример: «Покажите соответствие эффективности обучения применяемым методам и приемам:

- а) до 20%; 1) наглядные методы обучения;
 - б) до 50%; 2) практические методы обучения;
 - в) до 70%; 3) словесные методы обучения;
 - г) до 90%; 4) чтение учебника;
 - 5) дискуссионность в ходе занятия.
- (Ответы: а3, б1, в5, г2)».

Данное тестовое задание также носит обучающий характер и не нарушает принцип научности. Для снижения вероятности автоматического угадывания последнего ответа, здесь введен дополнительный элемент – дистрактор, чтобы и у последнего (четвертого) действия был вариант для выбора. В этом случае тестируемый в принципе может исключить именно неверный ответ и еще раз повторить все правильное.

4. Задания на установление правильной последовательности (цель – формирование логического мышления, а также проверка степени сформированности осознанных знаний, т. е. третьего уровня усвоения – «понимания»).

Пример: «Определите последовательность этапов формирования знаний, умений и навыков:

- 1-й этап а) «запоминание»;
- 2-й этап б) «воспроизведение»;
- 3-й этап в) «элементарные умения и навыки»;
- 4-й этап г) «понимание»;
- 5-й этап д) «перенос» (творческий уровень);
- е) «различение» (уровень знакомства).

Ответы: 1-й этап – е, 2-й этап – г, 3-й этап – а, 4-й этап – в, 5-й этап – д». Здесь также введен дополнительный элемент – дистрактор.

Рассмотрим критерий качества содержания тестовых заданий, который подразумевает определенность педагогического измерения, т. е. тестовое задание должно быть четко направлено:

– на контроль степени усвоения теоретических знаний второго уровня – «запоминания» или третьего уровня – «понимания» как высшей характеристики обобщенного понятия «знания»;

– на выявление степени готовности к практическим действиям (четвертого уровня обученности) – «элементарные умения и навыки», т. е. к реальному применению теории на практике в простейших (алгоритмизированных) ситуациях.

Творческий (пятый) уровень обученности – перенос – с помощью тестирования, как показывает практика, проверить не удается, что ограничивает область применения и степень доверия к тестам. Выход здесь один – применять дополнительно традиционные проверочные работы: диктанты (в том числе математические, физические и пр.), изложения, сочинения, контрольные и устные собеседования в их традиционном виде и т. п.

Общая цель проверки определяет и содержание теста. Если, например, цель тестирования – отобрать определенное количество учащихся (студентов) на олимпиаду, то ясно, что необходимо составить тестовые задания повышенной трудности, так как только с помощью подобных заданий можно отобрать лучших. Если стоит цель – оценить обучающихся во всем диапазоне знаний, то в тест включаются задания на проверку показателя «знания» и задания на проверку показателя «умения и навыки», т. е. – на уровне «запоминания», «понимания» и «элементарных умений и навыков» соответственно (примерное соотношение 20%–40%–40%). В основу данных рекомендаций положена наша пятиуровневая модель обученности личности и 10-балльная шкала оценки реального качества обучения в целом (см. вышеприведенный рисунок) [2, 7, 8].

Таблица 1

Пример тестирования по теме «Оценка качества обучения» на ФПК преподавателей МГОУ в мае 2012 года

Полученные баллы полностью положительной 10-и балльной шкалы									
1 балл	2 балла	3 балла	4 балла	5 баллов	6 баллов	7 баллов	8 баллов	9 баллов	10 баллов
Результаты тестирования									
-	-	1 чел 2,2%	3 чел 6,6%	2 чел 4,4%	11 чел 24,2%	22 чел 48,4%	5 чел 11%	1 чел 2,2%	-

Главной целью при разработке тестирования по конкретной дисциплине является создание единой технологии, процедур и инструментария для более достоверной оценки учебных достижений обучающихся. Здесь самое главное – высокая компетентность (обученность, эрудиция и личный опыт) составителей тестов, которые должны в первую очередь определиться со шкалами измерения. Порядковые и ранговые шкалы неприемлемы для оценки тестирования и должны быть исключены.

Многолетний опыт и практика показали, что можно также использовать наши «Критерии количественного подхода к оценке ре-

зультатов тестирования на основе полностью положительной 10-балльной шкалы», только при условии отсутствия в тестах заданий (вопросов) на низший уровень усвоения – различие (распознавание):

- 1 балл – за 1–3% правильных ответов;
- 2 балла – за 4–8% правильных ответов;
- 3 балла – за 9–15% правильных ответов;
- 4 балла – за 16–24% правильных ответов;
- 5 баллов – за 25–35% правильных ответов;
- 6 баллов – за 36–48% правильных ответов;
- 7 баллов – за 49–63% правильных ответов;
- 8 баллов – за 64–80% правильных ответов;
- 9 баллов – за 81–99% правильных ответов;
- 10 баллов – за 100% правильных ответов.

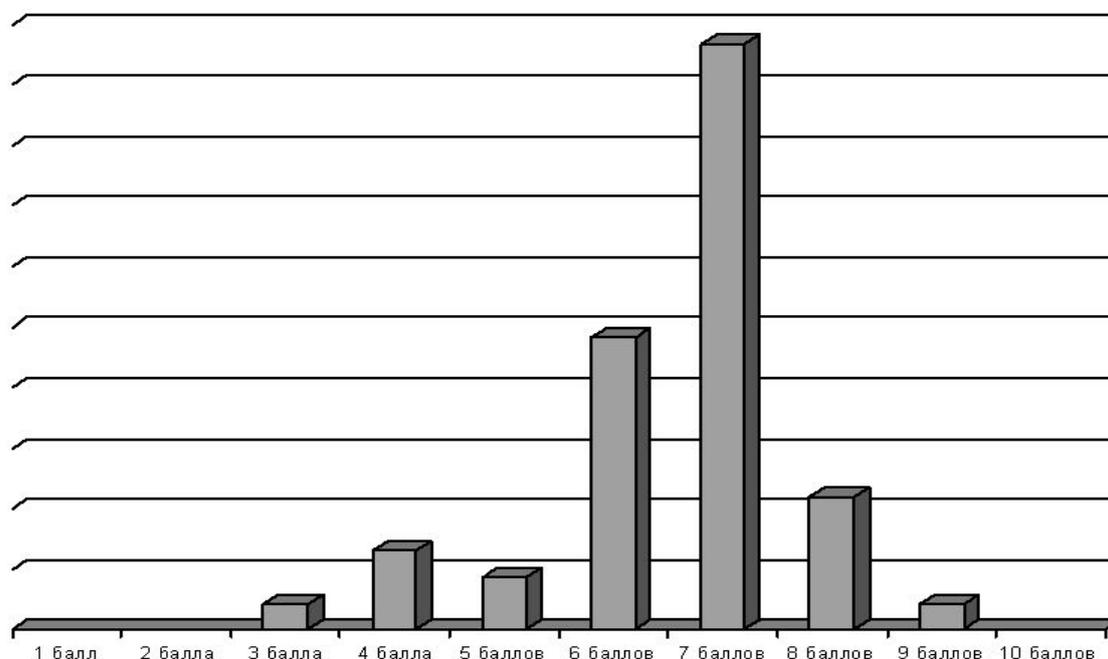


Рис. 2. Результаты тестирования по теме «Оценка качества обучения» на ФПК преподавателей МГОУ в мае 2012 года

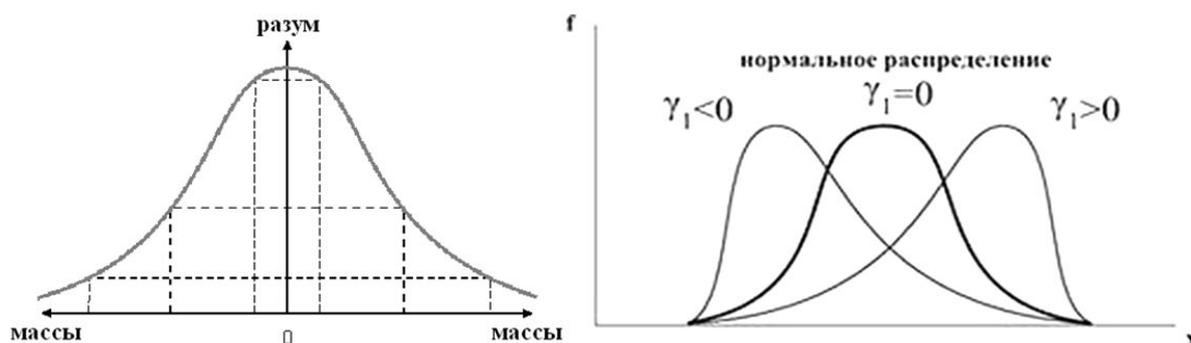


Рис. 3. Кривая им К.Ф. Гаусса (слева) и кривая, построенная на основе полученных баллов (справа)

Рассмотрим пример проведенного тестирования по теме «Оценка качества обучения» на ФПК преподавателей МГОУ в мае 2012 года. Приняли участие 45 преподавателей вузов. Отвечали на 59 вопросов в сорока семи заданиях (среди них было два задания на соответствие – десять правильных позиций и одно задание на установление последовательности – пять правильных позиций) (табли-

ца 1). Представим эти результаты на диаграмме (см. рис. 2).

Отсюда видно, что диапазон полученных оценок значительно шире, а значит и точнее, чем это было бы при использовании т. н. «пяти-балльной шкалы» (семь вместо возможных четырех разных баллов при традиционном подходе). Из диаграммы также видно, что, если

Таблица 2

Унификация оценочных шкал стран, входящих в Болонский процесс

Баллы Tenths	Обученность Training		Суррогатные шкалы Surrogate scale	Европейская шкала The European scale
1 балл – очень слабо – Very weak	от 1% from 1%	I. Узнавание Discernment	«2» “D”	F – неудовлетворительный ответ – unsatisfactory answer
2 балла – слабо – Weak	от 4% from 4%		«3-» “D+”	
3 балла – посредственно – Mediocre	от 9% from 9%	II. Запоминание Memory	«3» “C-”	FX – можно добрать до мини- мального удовлетворительного ответа – you can reach the minimum satisfactory answer
4 балла – удовлетворительно – Satisfactory	от 16% from 16%		«3+» “C”	
5 баллов – недостаточно хоро- шо – not good Enough	от 25% from 25%	III. Понимание Understanding	«4-» “C+”	E – отвечает минимальным требованиям удовлетвори- тельного ответа – meet the minimum require- ments of a satisfactory response
6 баллов – хорошо – Good	от 36% from 36%		«4» “B-“	
7 баллов – очень хорошо – Very good	от 49% from 49%	IV. Репродуктивное применение Reproductive application	«4+» “B”	D – достаточно удовлетвори- тельный ответ – quite satisfactory answer
8 баллов – отлично – Excellent	от 64% from 64%		«5-» “B+”	
9 баллов – великолепно – Excellent	от 81% from 81%	V. Перенос Transfer – creative level	«5» “A-“	B – очень хороший ответ – very good answer
10 баллов – прекрасно – Beautiful	100%		«5+» “A”	

построить кривую на основе полученных баллов, то она имеет отличия от нормального распределения с положительной асимметрией в сторону более высоких оценок (пример $\gamma_1 > 0$) (см. рис. 3).

Для оценки итогов тестирования были применены следующие нормативы на основе 10-балльной шкалы:

- 1 балл – от 1% – 1 верный ответ,
- 2 балла – от 4% – 2 верных ответа и более,
- 3 балла – от 9% – 5 верных ответов и более,
- 4 балла – от 16% – 9 верных ответов и более,
- 5 баллов – от 25% – 15 верных ответов и более,
- 6 баллов – от 36% – 22 верных ответа и более,
- 7 баллов – от 49% – 29 верных ответов и более,
- 8 баллов – от 64% – 38 верных ответов и более,
- 9 баллов – от 81% – 48 верных ответов и более,
- 10 баллов – 100% – 59 верных ответов.

Анализ использования вышеприведенных критериев оценки результатов тестирования в течение ряда лет показал: баллы «1», «2» и «10» появляются крайне редко, что свидетельствует о широких возможностях и большей достоверности данной оценочной шкалы (здесь проявились семь баллов соответствующих большей точности и достоверности, что нереально при так называемой «пятибалльной шкале»). Примечательный факт: количество хороших (6 баллов) и высоких (7, 8, 9 баллов) оценок коррелирует с посещаемостью занятий слушателями по данному курсу в целом.

В связи со входом нашей страны в Болонский процесс, мы должны думать и об унификации нашего оценочного процесса с учебными заведениями стран, вошедших в него. Существующие на западе оценочные шкалы невозможно «состыко-

вать» с так называемой пятибалльной шкалой, существующей в нашей стране с 1944 года, однако, это вполне возможно на основе предложенной нами и широко апробированной 10-балльной шкале. Все это и отражено в таблице 2.

Литература

1. Аванесов, В. С. Теоретические основы разработки заданий в тестовой форме : пособие для проф. преп. состава высшей школы / В. С. Аванесов. – М., МГТА, 1995. – 95 с.
2. Симонов, В. Законы и парадоксы асимметричного мира. Инновации в менеджменте. Сборник лекций и практических заданий / В. Симонов. – Lambert Academic Publishing, 2011. – С. 123–142.
3. Гулюкина, Н. А. Педагогический тест: этапы и особенности конструирования и использования : уч. пособие / Н. А. Гулюкина. – Новосибирск : НГТУ, 2001. – 131 с.
4. Клайн, П. Справочное руководство по конструированию тестов: Введение в психометрическое проектирование / П. Клайн ; пер. с англ. Е. П. Савченко. – Киев : ПАН Лтд, 1994. – 282 с.
5. Матушанский, Г. У. Педагогическое тестирование в России / Г. У. Матушанский // Педагогика, 2002. – № 2. – С. 15–21.
6. Переверзев, В. Ю. Критериально-ориентированные педагогические тесты для итоговой аттестации студентов / НМЦ СПО Минобразования РФ [Электронный ресурс]. URL : <http://www.russ.ru/edu/99-05-20/shishov.htm>.
7. Симонов, В. П. Виды, критерии и достоверность тестирования в образовании / В. П. Симонов // Педагогический менеджмент. НОУ-ХАУ в образовании. – М. : Высшее образование, 2006. – С. 278–284.
8. Симонов, В. П. Педагогическая диагностика в образовательных системах / В. П. Симонов. – М. : Перспектива, 2010. – С. 155–172.