

## **Методология организации тестирования на основе алгоритмов планирования и обработки двухуровневых многофакторных экспериментов\***

**Р. А. Нейдорф, Е. Н. Обухова**

*Исследуется декомпозиция знаний на множество «микрозваний», допускающих однозначную двоичную тестовую оценку. В таком случае задание теста формулируется как набор микрозваний ответов, совокупно раскрывающих уровень знания определённого подраздела. Это позволяет представить разработку набора ответов на тестовый вопрос как задачу исследования многофакторной зависимости. Таким образом, процесс разработки набора ответов можно рассматривать с позиций планирования многофакторного двухуровневого эксперимента. При таком подходе создание теста и оценка его результатов хорошо формализуются и становятся доступны преподавателям, не специализирующимся на разработках тестовых заданий. Методика проиллюстрирована примером построения и применения тестов для оценки уровня знаний по теории автоматического управления. Статистическая обработка результатов тестирования показала уровень знаний испытуемых, а также степень доверия тестирующего к осознанности ответа.*

**Ключевые слова:** факторы знания, планирование эксперимента, двухуровневые многофакторные эксперименты, варьирование факторов, матрица планирования.

**Введение.** Несмотря на определённые достижения в развитии тестирования, разработка методики составления тестовых материалов является на данном этапе сложной междисциплинарной проблемой [1]. Трудность внедрения тестовых методик в практику работы образовательных учреждений объясняется, в первую очередь, отсутствием у педагогов достаточной профессиональной подготовки в составлении и применении тестовых заданий. Не всегда у педагогов есть возможность освоить эти технологии. В связи с этим весьма актуальной становится задача формализации построения тестовых заданий, позволяющих сохранить высокую оценочную способность тестов. При этом важно, чтобы самостоятельно разрабатывать тесты могли преподаватели, не обладающие высокой квалификацией в данной области [2, 3].

**Постановка задачи.** Одной из характеристик тестового задания является количество допустимых ответов из множества предложенных. Если возможен один ответ на вопрос, тестируемый имеет оценку, выраженную долей полного знания. При этом повышается вероятность угадывания правильного ответа, вследствие чего невозможно достоверно оценить результаты тестирования.

Таким образом, для эффективной оценки подготовки по изучаемой дисциплине испытуемому с неуверенными знаниями необходимо дать возможность выбора нескольких вариантов ответов, кажущихся ему близкими к правильному. При этом не исключается возможность выбора одного ответа. В этом случае оценка знания формируется как выборочная совокупная из оценок содержательности указанных тестируемым ответов. Такой подход даёт возможность оценить степень случайности выбора ответов, которая характеризует неуверенность знаний тестируемого. Это обусловлено тем, что возникает возможность оценить близость по содержанию выбранных ответов. Слишком разнородные по оценкам ответы будут указывать на неуверенность знаний, а однородные — на близость испытуемого к истинным знаниям.

**Сущность предлагаемого метода решения проблемы.** Процесс разработки тестового вопроса рассматривается как построение строки, включающей в символьных переменных, обозначающих различные микрозвания [4, 5]. По введённой ранее парадигме, каждому микрозванию

---

\* Работа выполнена по теме № 2.3.13 «Метод сбалансированной дискретизации для задач имитационного моделирования динамических процессов в распределённых объектах» в рамках выполнения госзадания Минобрнауки России в части НИР.

может быть сопоставлено два варианта ответов: правильный — Yes и неправильный — No. Тогда возможный набор вариантов ответов на тестовый вопрос становится аналогичным матрице планирования многофакторного двухуровневого эксперимента, если положительный ответ — Yes заменить на +1, а отрицательный ответ — No на -1 (табл. 1).

Таблица 1

Кодовые строки-ответы и их оценки для  $n$ -факторного теста

№ строки	$x_1$	$x_2$	...	$x_m$	$y$
1.	$x_{11} = 1$	$x_{12} = 1$	...	$x_{1m} = 1$	$n(1)$
2.	-1 (0)	1	...	1	$n - 1(0, \dots)$
⋮	⋮	⋮	⋮	⋮	⋮
$i$	$x_{i1}$	$x_{i2}$	$x_{ij}$	$x_{im}$	$n - i \left( \frac{n-i}{m} \right)$
⋮	⋮	⋮	⋮	⋮	⋮
$n$	-1 (0)	-1 (0)	...	-1 (0)	0 (0)

Таким образом, варьируя ответы на микрозвнания на двух уровнях, можно получить набор строк-ответов на полный вопрос теста, реализующий все их возможные сочетания по аналогии с полным факторным экспериментом. Если количество микрозвнаний  $n$  выбрано, то можно найти количество различных полных ответов по формуле [6]:

$$N = 2^k, \quad (1)$$

где  $N$  — число опытов,  $k$  — число факторов, 2 — число уровней.

При оценке знаний тестируемого удобнее неправильный ответ оценивать нулём (нулевое знание), поэтому в таблице значение -1 удобно заменить на нули (указаны в скобках). Таким образом, из микрозвнаний как из факторов составляется вопрос-строка, а множество ответов на неё представляется как набор строк, составленных из элементов-ответов, принимающих значения 0 или 1. Варьирование ответов на этих двух уровнях позволяет реализовать полное множество строк-ответов возможных сочетаний элементов-ответов [7]. В связи с этим для микрозвнаний удобно, используя терминологию планирования эксперимента, применять термин «фактор знания».

Оценку  $y_i$  (обозначение  $y$  взято по аналогии с символом отклика в плане факторного эксперимента [6, 7]) для каждой строки возможного ответа на тестовый вопрос можно вычислять двумя способами. Первый — как абсолютную сумму оценок элементов-ответов по микрозвнаниям. Второй — как взвешенную относительно количества факторов сумму (указана в скобках), т. е. абсолютную сумму, делённую на количество элементов. Поэтому в более общем варианте оценки ответа взвешенной суммой по формуле

$$y_i = \frac{\sum_{j=1}^m x_{ij}}{m} \quad (2)$$

полностью правильный ответ на вопрос теста будет оценён как «1», полностью неправильный — «0», а частично правильный — дробным числом  $0 < y_i < 1$ .

Факторный подход к построению теста помогает доступно и однозначно формализовать процедуру построения теста и оценки его результатов. Однако с ростом числа составляющих вопрос факторов знаний число вариантов ответов растёт согласно (1) в показательной зависимости.

Это ограничивает объём вопроса по составу включённых факторов знаний (следует избегать построения излишне громоздкого теста, запутывающего испытуемого).

Рассмотренные условия формирования ответов на вопросы теста, содержащие  $n$  факторов знания, можно записать в виде матрицы планирования многофакторного эксперимента, где строки соответствуют различным вариантам ответов, а столбцы — значениям факторов знаний [6–8]. Если вопросы в teste отображают заложенные в них микрозвнания, то построенная таким образом матрица является универсальным инструментом построения вариантов ответов на весь тест в целом.

Необходимо заметить, что перечень подобного рода вопросов организован регулярным образом. Поэтому испытуемый, многократно проходивший тест, может понять закономерность порядка вопросов. Для устранения этого недостатка следует использовать другой приём метода планирования эксперимента — рандомизацию, т. е. перестановку в произвольном порядке строк табл. 1 — ответов теста [6, 7]. Тогда последовательность строк при повторных тестирований будет отличаться от расположения строк в табл. 1. Например,  $i = 7, 4, 9, 2, 5, 1, 6\dots$ . Этот стандартный приём теории планирования экспериментов позволяет уменьшить влияние на результаты опытов временного тренда, а в практике построения факторных тестов он даёт возможность снизить влияние адаптируемости испытуемых к тестам с регулярной структурой.

Кроме того, используя рандомизацию, можно увеличить количество комбинаторных вариантов последовательностей одинаковых ответов на один и тот же тестовый вопрос. В результате для небольшого набора ответов можно использовать полный комбинаторный набор вариантов ответов на каждый вопрос. Тогда количество одинаковых по содержанию, но различных по структуре вариантов тестовых ответов будет определяться по формуле вычисления количества перестановок ( $P$ ) из  $n$  по  $n$  —  $P_n^n = n!$ . В этом случае рандомизацию можно использовать для организации случайной последовательности вариантов при запросе очередного теста.

Для  $n = 4$  количество комбинаций из набора одинаковых ответов составит 24, однако для  $n = 8$  — уже 40320. Таким образом, для  $n \geq 5$  целесообразно всё-таки прибегать к рандомизации факторной таблицы.

#### **Примеры применения факторного подхода к составлению тестов.**

**Пример составления двухфакторного тестового задания.** По передаточной функции типового динамического звена  $W(p) = 6,8 / p$  найти значения вещественной и мнимой составляющих частотной передаточной функции.

Данное тестовое задание базируется на знании освоенных ранее разделов дисциплины. Необходимо знать следующее.

1. Способ перевода передаточной функции из операторной в частотную форму представления. Для этого производится замена  $p = j\omega$ .

2. Способ приведения дробно-рациональной функции комплексной переменной к стандартной форме записи. Для этого числитель и знаменатель умножаются на функцию, сопряжённую знаменателю. В числителе выделяются вещественная и мнимая части, и они почленно делятся на ставший вещественным знаменателем.

При этом условии в приведённом выше вопросе можно выделить два фактора знания.

1. Нахождение вещественной составляющей полученной частотной передаточной функции (ЧПФ) путём освобождения от мнимой составляющей в знаменателе и деления вещественной части числителя на вещественный знаменатель.

2. Нахождение мнимой составляющей той же ЧПФ делением мнимой части числителя на вещественный знаменатель.

Выделенным факторам знаний сопоставляются правильные ответы:

- вещественная составляющая ЧПФ —  $\operatorname{Re}\{W(j\omega)\} = \operatorname{Re}\left\{\frac{j \cdot 6,8 \cdot \omega}{-\omega^2}\right\} = \operatorname{Re}\{0 - j \cdot 6,8 \cdot \omega^{-1}\} = 0$ ;
- мнимая составляющая ЧПФ —  $\operatorname{Im}\{W(j\omega)\} = \operatorname{Im}\left\{0 - j \cdot \frac{6,8}{\omega}\right\} = -6,8 \cdot \omega^{-1}$ .

Неправильные ответы формируются с учётом характерных ошибок, допускаемых студентами по данному материалу. Важно также придерживаться принципа правдоподобия, когда в ответе присутствуют структуры и константы, связь которых с содержанием вопроса хорошо просматривается, например:

- вещественная составляющая ЧПФ —  $\operatorname{Re}\{W(j\omega)\} = 6,8$ ;
- мнимая составляющая ЧПФ —  $\operatorname{Im}\{W(j\omega)\} = \omega^{-1}$ .

Комбинируя факторы знания — правильные и неправильные ответы, можно составить все возможные строки ответов тестового задания, представленные ниже.

1. Вещественная составляющая:  $\operatorname{Re}\{W(j\omega)\} = 0$ ; мнимая составляющая:  $\operatorname{Im}\{W(j\omega)\} = -6,8j\omega^{-1}$ .
2. Вещественная составляющая:  $\operatorname{Re}\{W(j\omega)\} = 6,8$ ; мнимая составляющая:  $\operatorname{Im}\{W(j\omega)\} = -6,8j\omega^{-1}$ .
3. Вещественная составляющая:  $\operatorname{Re}\{W(j\omega)\} = 0$ ; мнимая составляющая:  $\operatorname{Im}\{W(j\omega)\} = -6,8j\omega^{-1}$ .
4. Вещественная составляющая:  $\operatorname{Re}\{W(j\omega)\} = 6,8$ ; мнимая составляющая:  $\operatorname{Im}\{W(j\omega)\} = -6,8j\omega^{-1}$ .

В этих строках содержится абсолютно правильный ответ на вопрос, содержащий два верных ответа по факторам знаний и абсолютно неправильный ответ, содержащий два неверных ответа и два полуправильных ответа, содержащие один верный и один неверный ответы по факторам знаний. Если варианты правильных ответов по факторам знаний обозначить +1, а неправильный — -1, то сведённые в таблицу варианты ответов образуют структуру, аналогичную матрице планирования полного факторного эксперимента (табл. 2).

**Таблица 2**

**Таблица кодовых строк-ответов и их оценок для двухфакторного теста**

№ строки	$x_1$	$x_2$	$y_i$
1.	1	1	1
2.	-1 (0)	-1 (0)	0
3.	1	-1 (0)	0,5
4.	-1 (0)	1	0,5

Имея ввиду, что отрицательного знания быть не может, а минимально возможное знание может быть нулевым, естественно сместить весовые оценки отрицательных ответов к нулевому значению (в таблице — в скобках). Тогда каждой строке ответов можно сопоставить коэффициент, определяемый, например, как среднее от совокупности ответов (обозначенный в табл. 2 как  $y_i$ ).

Возможны и другие способы количественной оценки результата такого тестирования, но здесь выбран наиболее простой, естественный и понятный.

Аналогичным образом можно подойти к составлению трёхфакторного тестового задания, которое будет содержать строку с тремя факторами знаний [4, 5, 8].

**Пример составления полного трёхфакторного тестового задания.** По передаточной функции апериодического звена  $W(p) = 1/(5p + 1)$  найти значения модуля и фазы частотной передаточной функции, а также представить её в показательной форме.

В данном тестовом задании присутствуют три фактора знания.

1. Умение аналитически находить амплитудно-частотную характеристику. Для этого выделяются вещественная и мнимая составляющие ЧПФ. Затем они вводятся в формулу, определяющую модуль ЧПФ  $A(\omega) = |W(j\omega)| = \sqrt{Re^2 + Im^2}$ .

2. Умение аналитически находить фазочастотную характеристику. Для этого выделяются вещественная и мнимая составляющие ЧПФ. Затем они вводятся в формулу, определяющую аргумент ЧПФ  $\phi(\omega) = \arg W(j\omega) = \arctg(Im/Re)$ .

3. Умение представить частотную передаточную функцию в показательной форме с использованием рассчитанных значений модуля и аргумента ЧПФ по формуле  $W(j\omega) = A(\omega) \cdot e^{j\phi(\omega)}$ .

Соответственно факторам знаний сопоставляются правильные ответы:

- амплитудно-частотная характеристика ЧПФ —  $A(\omega) = 1/\sqrt{1 + 25\omega^2}$ ;
- фазочастотная характеристика ЧПФ —  $\phi(\omega) = -\arctg(5\omega)$ ;
- показательная форма ЧПФ —  $W(j\omega) = \left(1/\sqrt{1 + 25\omega^2}\right) \cdot e^{-j\arctg 5\omega}$ .

Аналогично приведённому выше двухфакторному тестовому заданию составляются неправильные ответы — также на основе распространённых неправильных ответов и по принципу правдоподобия:

- амплитудно-частотная характеристика ЧПФ  $A(\omega) = 1/(1 + 5\omega^2)$ ;
- фазочастотная характеристика ЧПФ  $\phi(\omega) = \arctg(5\omega)^2$ ;
- показательная форма ЧПФ  $W(j\omega) = \left(5\omega/(1 + 5\omega^2)\right) \cdot e^{j\arctg \omega}$ .

Комбинируя между собой верные и неверные ответы по факторам знания всеми возможными способами, получаем полный трёхфакторный тест, состоящий из 8 строк ответов.

1) АЧХ —  $A(\omega) = 1/\sqrt{1 + 25\omega^2}$ ; ФЧХ —  $\phi(\omega) = -\arctg(5\omega)$ ; показательная форма —  $W(j\omega) = \left(1/\sqrt{1 + 25\omega^2}\right) \cdot e^{-j\arctg 5\omega}$ .

2) АЧХ —  $A(\omega) = 1/(1 + 5\omega^2)$ ; ФЧХ —  $\phi(\omega) = \arctg(5\omega)^2$ ; показательная форма —  $W(j\omega) = \left(5\omega/(1 + 5\omega^2)\right) \cdot e^{j\arctg \omega}$ .

3) АЧХ —  $A(\omega) = 1/\sqrt{1 + 25\omega^2}$ ; ФЧХ —  $\phi(\omega) = \arctg(5\omega)^2$ ; показательная форма —  $W(j\omega) = \left(1/\sqrt{1 + 25\omega^2}\right) \cdot e^{-j\arctg 5\omega}$ .

4) АЧХ —  $A(\omega) = 1/(1 + 5\omega^2)$ ; ФЧХ —  $\phi(\omega) = -\arctg(5\omega)$ ; показательная форма —  $W(j\omega) = \left(1/\sqrt{1 + 25\omega^2}\right) \cdot e^{-j\arctg 5\omega}$ .

5) АЧХ —  $A(\omega) = 1/\sqrt{1 + 25\omega^2}$ ; ФЧХ —  $\phi(\omega) = -\arctg(5\omega)$ ; показательная форма —  $W(j\omega) = \left(5\omega/(1 + 5\omega^2)\right) \cdot e^{j\arctg \omega}$ .

6) АЧХ —  $A(\omega) = 1/(1 + 5\omega^2)$ ; ФЧХ —  $\phi(\omega) = \arctg(5\omega)^2$ ; показательная форма —  $W(j\omega) = (1/\sqrt{1+25\omega^2}) \cdot e^{-j\arctg 5\omega}$ .

7) АЧХ —  $A(\omega) = 1/(1 + 5\omega^2)$ ; ФЧХ —  $\phi(\omega) = -\arctg(5\omega)$ ; показательная форма —  $W(j\omega) = (5\omega/(1 + 5\omega^2)) \cdot e^{j\arctg \omega}$ .

8) АЧХ —  $A(\omega) = 1/\sqrt{1 + 25\omega^2}$ ; ФЧХ —  $\phi(\omega) = \arctg(5\omega)^2$ ; показательная форма —  $W(j\omega) = (5\omega/(1 + 5\omega^2)) \cdot e^{j\arctg \omega}$ .

Аналогично предыдущему примеру на вопрос по каждому фактору знания можно дать правильный и неправильный ответы. Поэтому варьирование ответов по трём факторам знания на двух уровнях также позволяет применить матрицу планирования полного факторного эксперимента для полного трёхфакторного теста, представленную в табл. 3.

**Таблица 3**

**Таблица кодовых строк-ответов и их оценок для 3-факторного теста**

№ строки	$x_1$	$x_2$	$x_3$	$y_i$
1.	1	1	1	1
2.	-1 (0)	-1 (0)	-1 (0)	0
3.	1	-1 (0)	1	0,67
4.	-1 (0)	1	1	0,67
5.	1	1	-1 (0)	0,67
6.	-1 (0)	-1 (0)	1	0,33
7.	-1 (0)	1	-1 (0)	0,33
8.	1	-1 (0)	-1 (0)	0,33

Таким образом, методика составления тестов на основе планирования факторного эксперимента позволяет конструировать трёхфакторные тестовые задания, содержащие 8 строк ответов. Необходимо отметить важное преимущество таких тестов: увеличение количества факторов знаний приводит к общему увеличению числа ответов — и вероятность угадывания правильного ответа уменьшается. Тест получается более содержательным в смысле градации уровня знаний тестируемого, так как включает уже две дробные оценки знаний, соответствующие двум правильным ответам из трёх (0,67) и одному правильному ответу (0,33). В таком случае при пороге знания 0,6 ответивший на два вопроса получит зачёт по тесту, т. е. вероятность сдачи зачёта (прохождения теста) составляет 0,5. В двухфакторных тестах такая вероятность составляет 0,25.

Однако трёхфакторные тесты имеют и недостатки. Они получаются достаточно громоздкими. Это увеличивает время на ответ, и тестируемый испытывает повышенную психологическую нагрузку. Соответственно, растёт вероятность случайной ошибки, ведь внимание рассредоточено на восемь строк ответов.

В связи с этим для упрощения процедуры тестирования был предложен подход [6], при котором каждое полное трёхфакторное тестовое задание декомпозируется на усечённые трёхфакторные тестовые задания (аналог дробных реплик в теории планирования экспериментов). Сущность такого подхода кратко освещена ниже.

**Разбиение полного трёхфакторного теста на усечённые дробные реплики.** Принцип формирования усечённого дробного тестового задания заключается в том, что необходимо составить дробную реплику из четырёх вариантов ответов [6]. При этом реплика обязательно должна содержать правильный и неправильный ответы, т. к. категорически нельзя лишать тестируемого

возможности продемонстрировать полное знание вопроса или его отсутствие. В связи с этим строки ответов, имеющие различные оценки по сумме факторов знаний, комбинируются друг с другом таким образом, что в каждом усечённом тестовом задании присутствуют строки ответов с оценками 0,33 и 0,67. Такой принцип также кажется обязательным, поскольку наличие двух неполных ответов с оценками 0,67 повышает вероятность завышения оценки за счёт угадывания, а наличие двух неполных ответов с оценками 0,33 гарантированно и, возможно, неоправданно занижает оценку выбирающего второй ответ в дополнение к правильному. Иными словами, необходимо чтобы в teste содержался весь спектр ответов. Так, в усечённом тестовом задании, приведённом ниже, помимо абсолютно правильной и неправильной строки ответа имеются строки с оценками, составляющими 0,33 и 0,67 балла, которые соответствуют срокам № 5 и № 7 в полном трёхфакторном тестовом задании.

**Пример составления усечённого трёхфакторного тестового задания.** По передаточной функции апериодического звена  $W(p) = 1/(5p + 1)$  найти значения модуля и фазы комплексной передаточной функции и представить её в показательной форме.

$$1) \text{ АЧХ} — A(\omega) = 1/\sqrt{1+25\omega^2}; \text{ ФЧХ} — \phi(\omega) = -\arctg(5\omega); \text{ показательная форма} — W(j\omega) = \left(1/\sqrt{1+25\omega^2}\right) \cdot e^{-j\arctg 5\omega}.$$

$$2) \text{ АЧХ} — A(\omega) = 1/(1+5\omega^2); \text{ ФЧХ} — \phi(\omega) = \arctg(5\omega)^2; \text{ показательная форма} — W(j\omega) = \left(5\omega/(1+5\omega^2)\right) \cdot e^{j\arctg \omega}.$$

$$3) \text{ АЧХ} — A(\omega) = 1/\sqrt{1+25\omega^2}; \text{ ФЧХ} — \phi(\omega) = -\arctg(5\omega); \text{ показательная форма} — W(j\omega) = \left(5\omega/(1+5\omega^2)\right) \cdot e^{j\arctg \omega}.$$

$$4) \text{ АЧХ} — A(\omega) = 1/(1+5\omega^2); \text{ ФЧХ} — \phi(\omega) = -\arctg(5\omega); \text{ показательная форма} — W(j\omega) = \left(5\omega/(1+5\omega^2)\right) \cdot e^{j\arctg \omega}.$$

Строки ответов сводятся в табл. 4, где правильные ответы на факторы знания  $x_1$ ,  $x_2$  и  $x_3$  кодируются +1, а неправильные — -1. Каждая строка ответов имеет определённый весовой коэффициент, обозначенный  $y_i$ .

**Таблица 4**

**Таблица кодовых строк-ответов и их оценок для усечённого З-факторного теста**

№ строки	$x_1$	$x_2$	$x_3$	$y_i$
1.	1	1	1	1
2.	-1 (0)	-1 (0)	-1 (0)	0
3.	1	1	-1 (0)	0,67
4.	-1 (0)	1	-1 (0)	0,33

При повторном использовании данного тестового задания строки табл. 5 в результате рандомизации с большой вероятностью будут расположены в последовательности, отличной от табл. 4 согласно вышеизложенной теории.

**Примеры использования факторного тестирования.** По составленным двухфакторным заданиям были проведены тестирования в группах, изучающих дисциплину «Теория автоматического управления». Перед испытуемыми ставилась задача выбора правильной строки ответов. Не-

уверенные в собственных знаниях учащиеся могли выбрать два наиболее правильных, на их взгляд, ответа. По результатам тестирования рассчитывались:

- оценки (в среднем) знаний, продемонстрированных учащимися;
- степень доверия тестирующего к знаниям учащихся, с учётом их неуверенности.

*Таблица 5*

**Рандомизированная таблица 3-факторного теста**

№ строки	$x_1$	$x_2$	$x_3$	$y_i$
3.	1	1	-1 (0)	0,67
1.	1	1	1	1
4.	-1 (0)	1	-1 (0)	0,33
2.	-1 (0)	-1 (0)	-1 (0)	0

**Пример двухфакторного тестирования.** В табл. 6 приведён результат тестирования с использованием двухфакторных тестовых заданий. Статистическая обработка тестовой выборки производилась в соответствии с известным математическим аппаратом, используемым для обработки результатов экспериментов [9]. В различных формах статистическую обработку используют при сдаче единых экзаменов в России [1] и за рубежом [2, 3].

*Таблица 6*

**Результаты тестирования и расчёт средних оценок знаний**

		Номер тестируемого ( $j$ )										
		1	2	3	4	5	6	7	8	9	10	11
Тест. задание 1	$y_1$	1	0,5	0,5	1	1	0	0,5	0,5	1	0	1
	$y_2$		1	1		0,5						
	Балл 1	1	0,75	0,75	1	0,75	0	0,5	0,5	1	0	1
	Дисп. 1	—	0,125	0,125	—	0,125	—	—	—	—	—	—
	СКО1	0	0,35	0,35	0	0,35	0	0	0	0	0	0
	$y_1$	1	0,5	1	1	1	1	0,5	0	0,5	0,5	1
	$y_2$		0,5									
Тест. задание 2	Балл 2	1	0,5	1	1	1	1	0,5	0	0,5	0,5	1
	Дисп. 2	—	0	—	—	—	—	—	—	—	—	—
	СКО 2	—	0	0	0	0	0	0	0	0	0	0
	$y_1$	1	0,5	1	0,5	0	1	0,5	0	0,5	0,5	1
	$y_2$					1		0		0		
	Балл 3	1	0,5	1	0,5	0,5	1	0,25	0	0,25	0,5	1
	Дисп. 3	—	—	—	—	0,5	—	0,125	—	—	—	—
Тест. задание 3	СКО 3	0	0	0	0	0,7	0	0,35	0	0,35	0	0
	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
	$y_1$	1	0	1	1	0,5	1	1	0	0,5	0	0,5
	$y_2$							0,5	0,5			
	Балл 7	1	0	1	1	0,5	1	0,75	0,25	0,5	0	0,5
	Дисп. 7	—	—	—	—	—	—	—	0,125	—	—	—
	СКО 7	0	0	0	0	0	0	0,35	0,35	0	0	0
Тест. задание 7	Сред. оценка	0,92	0,46	0,96	0,85	0,82	0,78	0,57	0,25	0,39	0,35	0,78
	Дисп. по тесту	—	0,017	0,017	—	0,089	—	0,035	0,017	0,017	—	—
	СКО по тесту	0	0,13	0,13	0	0,29	0	0,18	0,13	0,13	0	0

В табл. 6 введены следующие обозначения:  $y_i$  — оценка выбранной строки ответа; балл — среднее значение оценок выбранных тестируемым строк ответов; дисп. — дисперсия по тестовому заданию; СКО — среднее квадратическое отклонение от оценки по тестовому заданию; сред. оценка — средняя суммарная оценка за тест.

По результатам двухфакторного тестирования были построены столбчатые гистограммы, представленные на рис.

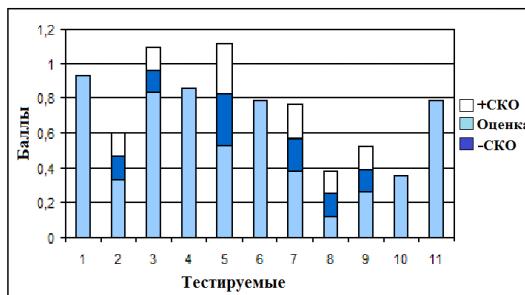


Рис. Иллюстрация оценок знаний и степени их неуверенности, полученных с использованием двухфакторных тестовых заданий

Столбчатые диаграммы построены по полученным усреднённым оценкам тестируемых по всему тесту. От средней оценки, полученной за тест, вверх и вниз отложено значение среднеквадратического отклонения. Сформированный диапазон указывает на область возможных значений оценки знаний и характеризует неуверенность ответа тестируемого студента. Так, например, у испытуемых, выбравших в качестве ответов по одной строке в каждом тестовом задании, дисперсия и соответствующее ей СКО, показывающее степень неуверенности знаний, равны нулю [10]. Это указывает на уверенность частичных знаний как по отдельному тестовому заданию, так и по тесту в целом, которую наглядно иллюстрируют столбчатые диаграммы под номерами 1, 4, 6, 10, 11 (см. рис.).

По результатам тестируемых, проявивших неуверенность и указавших две строки ответов, построены столбчатые гистограммы под номерами 2, 3, 5, 7, 8, 9. В данном случае неуверенность знаний характеризует значение дисперсии или её среднеквадратического отклонения от ожидаемой усреднённой оценки. Из построенных столбчатых диаграмм видна обратная зависимость статистических характеристик и степени доверия, т. е. чем больше среднеквадратическое отклонение от усреднённой оценки знаний по тесту, тем меньше доверия к таким знаниям.

**Выводы.** Рассмотренные примеры демонстрируют хороший уровень формализации и удобство составления тестов на основе методики многофакторного планирования двухуровневых экспериментов. Кроме того, хорошо формализуется оценка результатов теста, что повышает его объективность.

Важным нововведением является допущение своеобразной «нечёткости» ответа на вопрос теста. Его следствие — двухпараметрическая оценка результата теста: оценка неполного знания и степени уверенности этого знания.

Результаты тестирований с использованием двухфакторных и трёхфакторных заданий подтвердили положительную роль факторности теста, помогающей ранжировать испытуемых с неуверенными знаниями.

#### **Библиографический список**

1. Ефремова, Н. Ф. Тестовый контроль в образовании / Н. Ф. Ефремова. — Москва : Логос : Университетская книга, 2007. — 386 с.
2. Curriculum Frameworks for Mathematics and Science / general editor D. Robitaille // TIMMS Monograph № 1. — Vancouver : Pacific Educational Press, 2003. — 616 р.

3. Measuring student knowledge and skills. A New Framework for Assessment / Organization for Economic Co-operation and Development (OECD). — Paris : OECD Publications Service, 1999. — 84 p.
4. Нейдорф, Р. А. Методология построения тестов на основе алгоритма планирования двухуровневых многофакторных экспериментов [Электронный ресурс] / Р. А. Нейдорф, Е. Н. Обухова // Информационные технологии в управлении: сб. тр. V Российской междунар. конф. — Санкт-Петербург, 2012. — 1 электрон. опт. диск (CD-ROM).
5. Нейдорф, Р. А. Методика составления тестовых заданий на основе алгоритма планирования двухуровневых многофакторных экспериментов и анализ их использования в образовательном процессе [Электронный ресурс] / Р. А. Нейдорф, Е. Н. Обухова // Методики и технологии обеспечения и оценки качества образования: сб. мат-лов междунар. науч. конф. (Киев, 26–28 июня). — Киров : МЦНИП, 2013. — 1 электрон. опт. диск (CD-ROM).
6. Чернова, Н. А. Статистические методы планирования экстремальных экспериментов / Н. А. Чернова, В. В. Налимов. — Москва : Наука, 1965. — 340 с.
7. Адлер, Ю. П. Планирование эксперимента при поиске оптимальных условий / Ю. П. Адлер, Е. В. Маркова, Ю. В. Грачевский. — Москва : Наука, 1976. — 274 с.
8. Нейдорф, Р. А. Анализ применения трёхфакторных дробных тестов в учебном процессе / Р. А. Нейдорф, Е. Н. Обухова // Математические методы в технике и технологиях — ММТТ-26 : сб. тр. XXVI междунар. науч. конф. / под общ. ред. А. А. Большакова. — Нижний Новгород, 2013. — Т. 9, секция 11. — С. 320–323.
9. Вентцель, Е. С. Теория вероятностей и её инженерные приложения / Е. С. Вентцель, Л. А. Овчаров. — 3-е изд. — Москва : Академия, 2003. — 464 с.
10. Нейдорф, Р. А. Методика статистического расчёта балла тестируемого с выделением области неуверенности знаний [Электронный ресурс] / Р. А. Нейдорф, Е. Н. Обухова // Системный анализ, управление и обработка информации : сб. тр. IV Междунар. науч. семинара (п. Дивноморское, 29 сент. — 3 окт. 2013). — 1 электрон. опт. диск (CD-ROM).

Материалы поступили в редакцию 02.07.2013.

## **References**

1. Yefremova, N. F. Testovyy kontrol v obrazovanii. [Test check in education.] Moscow: Logos: Universitetskaya kniga, 2007, 386 p. (in Russian).
2. Robitaille, D., gen. ed. Curriculum Frameworks for Mathematics and Science. TIMSS Monograph № 1. Vancouver: Pacific Educational Press, 2003, 616 p.
3. Measuring student knowledge and skills. A New Framework for Assessment. Organization for Economic Co-operation and Development (OECD). Paris: OECD Publications Service, 1999, 84 p.
4. Neydorf, R. A., Obukhova, E. N. Metodologiya postroyeniya testov na osnove algoritma planirovaniya dvukhurovnevykh mnogofaktornykh eksperimentov. [Methodology of test construction based on planning algorithm for two-level multifactor experiments.] Informatzionnye tekhnologii v upravlenii: sb. trudov V Rossiyskoy mezhdunar. konf. [Information technologies in management: Proc. V Rus. Int. Sci. Conf.] 1 electronic optical disc (CD-ROM). Saint Petersburg, 2012 (in Russian).
5. Neydorf, R. A., Obukhova, E. N. Metodika sostavleniya testovykh zadaniy na osnove algoritma planirovaniya dvukhurovnevykh mnogofaktornykh eksperimentov i analiz ikh ispolzovaniya v obrazovatelnom protsesse. [Methods of test task preparation based on planning algorithm for two-level multifactor experiments, and analysis of their application in educational process.] Metodiki i tekhnologii obespecheniya i otsenki kachestva obrazovaniya: sb. materialov mezhdunar. nauch. konf. (Kiyev, 26–28 iyunya). [Methods and technologies of ensuring and assessing quality of education: Proc. Int. Sci. Conf. Kiev, 26–28 June.] 1 electronic optical disc (CD-ROM). Kirov: MTsNIP, 2013 (in Russian).

6. Chernova, N. A., Nalimov, V. V. Statisticheskiye metody planirovaniya ekstremalnykh eksperimentov. [Statistical methods of extreme experiment planning.] Moscow: Nauka, 1965, 340 p. (in Russian).
7. Adler, Y. P., Markova, E. V., Granovsky, Y. V. Planirovaniye eksperimenta pri poiske optimalnykh usloviy. [Experiment design in search for optimal conditions.] Moscow: Nauka, 1976, 274 p. (in Russian).
8. Neydorf, R. A., Obukhova, E. N. Analiz primeneniya trekhfaktornykh drobnykh testov v uchebnom protsesse. [Analysis of three-factor fractional test application in educational process.] Matematicheskiye metody v tekhnike i tekhnologiyakh — MMTT-26: sb. trudov XXVI mezhdunar. nauch. konf. [Bolshakov, A. A., ed. Mathematical methods in engineering and technologies — MMTT-26: Proc. XXVI Int. Sci. Conf.] Nizhny Novgorod, 2013, vol. 9, sec. 11, pp. 320–323 (in Russian).
9. Ventzel, E. S., Ovcharov, L. A. Teoriya veroyatnostey i yeye inzhenernyye prilozheniya. [Probability theory and its engineering applications.] 3<sup>rd</sup> ed. Moscow: Akademiya, 2003, 464 p. (in Russian).
10. Neydorf, R. A., Obukhova, E. N. Metodika statisticheskogo rascheta balla testiruyemogo s vydeleniem oblasti neuvarennosti znaniy [Statistical methods of testee's credit points calculation with selection of knowledge uncertainty]. Sistemnyy analiz, upravleniye i obrabotka informatsii: sb. trudov IV Mezhdunar. nauch. seminara (p. Divnomorskoye, 29 sent. — 3 okt. 2013) [System analysis, management, and information processing: Proc. IV Int. Sci. Seminar (Divnomorskoye, 29 Sept. — 3 Oct.)] 1 electronic optical disc (CD-ROM) (in Russian).

## **TESTING DESIGN METHODOLOGY BASED ON TWO-LEVEL MULTIFACTOR EXPERIMENT PLANNING AND DATA PROCESSING\***

**R. A. Neydorf, E. N. Obukhova**

*The decomposition of knowledge into an array of "micro-knowledge" permitting a single-valued binary test score is studied. In this case, the test task is formulated as a set of micro-knowledge responses jointly revealing the knowledge level of the specific subsection. It allows presenting the development of a set of responses to a test question as the objective of investigating the multifactor dependence. Thus, the development process of a set of responses can be considered from the perspective of planning a multifactor two-level experiment. This approach allows a good formalization of the test construction, and its result estimation which becomes available to the teachers who do not specialize in the test task development. The method is illustrated by an example of the test construction, and application to assess the knowledge level on the automatic control theory. The data statistics has shown the respondents' knowledge level, as well as the degree of a tester's belief to the response awareness.*

**Keywords:** knowledge factors, experiment planning, two-level multifactor experiments, factor variability, planning matrix.

---

\* The work is done on 2.3.13 theme "Balanced discretization method for tasks of dynamic process simulation modeling in distributed objects" within the frame of the Ministry of Education and Science of the Russian Federation thematic research plan.