

УПРАВЛЕНИЕ В СОЦИАЛЬНЫХ И ЭКОНОМИЧЕСКИХ СИСТЕМАХ

УДК [004.9+ 004.021]:[378.146+373.1]

СИСТЕМНЫЙ АНАЛИЗ ВОПРОСОВ КОМПЬЮТЕРНОГО ТЕСТИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ НЕЧЕТКИХ ОТВЕТОВ.

Часть 1. Тестовые задания категорий «выбор из набора ответов» и «ранжирование объектов»

Статья поступила в редакцию 30.12.2018, в окончательном варианте – 11.02.2019.

Брумштейн Юрий Моисеевич, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,

кандидат технических наук, доцент, ORCID <http://orcid.org/0000-0002-0016-7295>,
https://elibrary.ru/author_profile.asp?authorid=280533, e-mail: brum2003@mail.ru

Коновалова Дарья Игоревна, Астраханский государственный университет, 414056, Российская Федерация, г. Астрахань, ул. Татищева, 20а,
студент, ORCID: 0000-0002-8348-0171, e-mail: haizbq@gmail.com

Дана общая характеристика целей и технологий проведения компьютеризованного тестирования знаний. Приведена классификация типичных тестовых заданий, используемых при компьютерном тестировании, включающая четыре категории. Охарактеризованы некоторые дополнительные возможности тестирования, обеспечиваемые современными инструментами разработки и эксплуатации программных средств. Проанализированы достоинства и недостатки применения тестовых заданий с четкими ответами. Обоснована принципиальная целесообразность использования тестовых заданий с нечеткими ответами. Для двух категорий таких тестовых заданий выполнен системный анализ целей их использования, вариантов компьютерной реализации, алгоритмов обработки ответов. Для категории «тестовые задания с выбором ответов из predetermined наборов» рассмотрены следующие подкатегории: 1) с единственным верным ответом, причем об этом сказано в формулировке тестовых заданий; 2) с двумя и более верными ответами, количество которых указано в формулировке тестовых заданий; 3) с заранее неизвестным для пользователя количеством верных ответов. Разновидностью подкатегорий 1...3 может быть выбор «неверных ответов» из числа предъявленных; 4) тестовые задания с предлагаемыми бинарными ответами типа «да» или «нет». Рассмотрены возможные варианты предъявления ответов тестируемому на мониторе: все одновременно; последовательное воспроизведение ответов. Для обоих вариантов возможны подварианты предъявления ответов: в predetermined или случайном порядке. В случае нечетких ответов для каждого из сочетаний «подкатегория тестовых заданий – вариант предъявления – подвариант» рассмотрены особенности формулировок вопросов в тестовых заданиях; способы технической реализации на ПЭВМ предъявления тестовых заданий и инструментов для нечеткого выбора; методы оценки нечетких ответов тестируемых лиц. Для категории тестовых заданий «на ранжирование предложенных объектов по заданному критерию» исследованы их возможные подкатегории в случае нечетких ответов, особенности компьютерной реализации в отношении представления вопросов и получения ответов тестируемых лиц, методы обработки (оценки) таких ответов. Рассматриваемый материал проиллюстрирован примерами из предметных областей «Информатика» и «Вычислительная техника» (для уровней обучения «школа» и «вуз»).

Ключевые слова: компьютерное тестирование, типы тестовых заданий, выбор ответов, ранжирование объектов, нечеткие ответы, методы оценивания ответов, адаптивное тестирование, программные средства, использование графических объектов, информатика, вычислительная техника

SYSTEM ANALYSIS OF COMPUTER TESTING QUESTIONS WITH ILLEGIBLE ANSWERS USING.

Part 1. Test tasks of categories "the choice from a set of answers" and "ranging of objects"

The article was received by editorial board on 30.12.2018, in the final version – 11.02.2019.

Brumshhteyn Yuri M., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,

Cand. Sci. (Engineering), Associate Professor, ORCID <http://orcid.org/0000-0002-0016-7295>,
https://elibrary.ru/author_profile.asp?authorid=280533, e-mail: brum2003@mail.ru

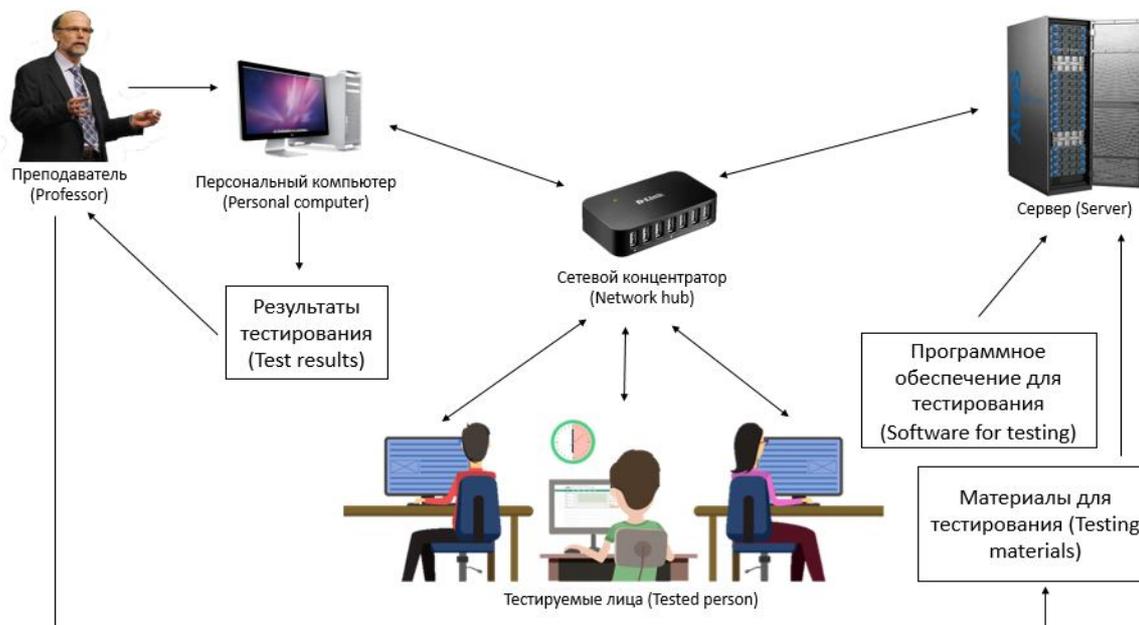
Konovalova Daria I., Astrakhan State University, 20a Tatishchev St., Astrakhan, 414056, Russian Federation,
student, ORCID: 0000-0002-8348-0171, e-mail: haizbq@gmail.com

The general characteristic of the purposes and technologies for conducting computerized knowledge testing is given. Authors are examined the classification of the typical test tasks used at computer testing, including four categories. Some

additional opportunities of testing, provided with modern development tools and operation of software, are characterized. The paper analyzes merits and demerits of test tasks application with definite answers. Authors are proved the basic expediency of test tasks with illegible answers usage. For two categories of such test tasks the system analysis of their use purposes, options of computer realization, answers processing algorithms is made. For category "Test tasks with the choice of answers from the predetermined sets" the following subcategories are considered: 1) with the only right answer, and about it is told in test tasks formulations; 2) from two and more right answers which number is specified in a formulation of test tasks; 3) with the number of right answers, which are beforehand unknown for the user. The choice of "incorrect answers" from offered answers can be a kind of subcategories for «1 ... 3»; 4) test tasks with the offered binary answers of the "Yes" or "No" type. Possible answers presentation options for the tested person on the monitor are considered: all answers are shown at the same time; consecutive reproduction of answers. For both options may be used sub-options of answers presentation at the monitor: in the predetermined or accidental order. In case of illegible answers authors are considered for each of combinations " test tasks subcategory – presentation option - a sub-option" questions formulations features in test tasks; ways of technical realization for presentation on personal computers of test tasks and instruments for illegible choice; methods of the illegible answers assessment of the tested persons. For category of test tasks "on offered objects ranging by the set criterion" authors are investigated possible subcategories in case of illegible answers, features of computer realization concerning representation of questions and obtaining answers of the tested persons, methods of processing (assessment) of such answers. The considered material are illustrated with examples from subject domains of "Information technologies" and "Computer facilities" (for training levels "school" and "higher education institution").

Key words: computer testing, test items types, choice of answers, object ranking, indistinct answers, answers evaluating methods, adaptive testing, software, graphical objects usage, computer science, computer technology

Graphical annotation (Графическая аннотация)



Введение. В настоящее время компьютерное тестирование знаний (КТЗ) широко применяется в различных целях [2, 6]: **а)** как средство «входного контроля знаний» перед началом обучения по определенной дисциплине или группе дисциплин; **б)** в процессе обучения, в том числе для адаптации [11, 16] уровня сложности и содержания изучаемого материала на основании результатов промежуточного тестирования; **в)** для итоговой оценки знаний после завершения курсов обучения; **г)** в рамках использования контрольно-обучающих систем (тестирование в сочетании с «обучающими фрагментами», в том числе и мультимедийного характера); **д)** для определения «остаточных знаний» через некоторое время после окончания обучения по определенной дисциплине; **е)** как средство оценки квалификации лиц, претендующих на занятие определенных должностей; **ж)** в качестве метода профессиональной (квалификационной) аттестации уже работающего персонала; **з)** как средство оценки некоторых психофизиологических характеристик физических лиц [7].

В типичных случаях тесты состоят из нескольких десятков тестовых заданий (ТЗ), рассчитанных на визуальное восприятие тестируемыми лицами («скрининговое» тестирование). Актуальной задачей является повышение вариативности ТЗ, включая формулировки вопросов и ответов, предъявляемые задания расчетного характера и пр. В последнем случае могут осуществляться изменения параметров расчетов (в том числе индивидуально для каждого тестируемого) с помощью программных генераторов случайных чисел, а также «штатных» средств MS Excel [20].

Предъявление ТЗ для выполнения их тестируемыми лицами возможно по различным вариантам: в детерминированном порядке, зафиксированном в сценарии теста; последовательность предъявления ТЗ (и, возможно, их сложность) определяется динамически – с учетом результатов ответов по предшествующим ТЗ, израсходованного тестируемым времени и пр.; порядок прохождения (выполнения) ТЗ в тесте выбирается самими тестируемыми – в некоторых случаях с возможностью «возврата» к ранее выполненным ТЗ. При этом совокупность ТЗ должна обладать достаточной дифференцирующей способностью [8] в отношении оценки знаний обучаемых. Ответы тестируемых лиц могут оцениваться как по каждому ТЗ, так и по их совокупности [3]. При этом в общем случае могут использоваться различные подходы [15, 19].

Современные инструментальные (программные) средства значительно расширили возможности проведения КТЗ, в том числе за счет применения графических объектов (статических и динамических) и звука; использования сенсорных экранов – это дает возможность работать без использования компьютерной мыши или тачпада; повышения «интеллектуальности» обработки ответов тестируемых лиц. При КТЗ важно соблюдение допустимых норм нагрузок на зрительный аппарат тестируемых лиц (особенно младших школьников) [4]; применение заданий с воспроизведением вопросов и, возможно, предопределенных ответов, в «звуковой» форме; в перспективе – получение ответов от тестируемых лиц в речевой форме или в виде жестов. Специально отметим, что использование мультимедийных объектов в ТЗ (включая динамическую графику) может значительно повышать мотивацию к обучению тестированию – это может быть особенно важно для младших школьников, а также лиц со «слабыми» типами нервной системы.

В настоящее время преобладает использование КТЗ для тех областей, где вопросы в ТЗ допускают формализованное представление с «четким выбором» ответов – такой вариант упрощает их программную обработку. Однако в ряде случаев более целесообразно использование «нечетких» ответов на ТЗ, которые лучше отражают реалии «предметных областей», логику принятия и реализации решений тестируемыми лицами. Кроме того, ТЗ с «нечеткими ответами» потенциально позволяют расширить «предметные области», использования КТЗ; улучшить дифференцирующую способность ТЗ [8].

Работы по различным аспектам «нечеткого тестирования» в русскоязычной литературе немногочисленны [9, 12, 13], и в них отсутствует комплексный анализ особенностей формулировок вопросов и обработки ответов для различных категорий ТЗ; подробный анализ средств компьютерной реализации соответствующих ТЗ. Поэтому целью настоящей статьи был системный анализ вопросов, связанных с использованием в тестах различных видов ТЗ, предполагающих возможность нечетких ответов тестируемых лиц. При этом вопросы ограничения времени ответов тестируемых лиц на отдельные ТЗ и на тест в целом мы, как правило, не затрагиваем – для уменьшения объема работы.

Отметим, что в некоторых случаях может быть целесообразным дать возможность самому тестируемому выбирать, какими типами ответов он хочет пользоваться – четкими или нечеткими.

Общая характеристика целей и технологий проведения компьютеризованного тестирования знаний. В настоящее время в России, по крайней мере, в рамках образовательных процессов преобладает тестирование знаний с использованием стационарного компьютерного оборудования и программных средств. При этом типичны следующие варианты: тестирующая программная система (ТПС) инсталлирована на ПЭВМ пользователя; ТПС находится на лазерном диске с некоторой обучающей системой; ТПС размещено на сервере и обслуживает локальную компьютерную сеть; размещенная на сервере ТПС допускает дистанционный доступ извне организации, в том числе доступ на основе ввода пары «логин – пароль», которая индивидуальна для каждого пользователя. Отметим, что в последнем случае возможно и тестирование «мобильных пользователей» – с применением компьютерных планшетов или даже смартфонов.

В общем случае ТПС может содержать в себе значительное количество «тестов» (т.е. совокупностей ТЗ), которые должен пройти обучаемый. При таких условиях целесообразно хранение результатов по пройденным ранее тестам, в том числе для установления уровней сложности при последующем тестировании.

В рамках одного теста уровень сложности ТЗ может быть неизменным или динамически изменяться (адаптироваться [10, 16]) с учетом результатов ответов на предшествующие ТЗ; истраченного и оставшегося времени тестирования и пр. Для такой адаптации могут быть использованы различные методы, включая и нейронные сети [14], обучаемые по ходу тестирования.

Первоначально в рамках ТПС использовались преимущественно (а иногда и исключительно) задания, в которых формулировки вопросов и ответов были представлены исключительно в виде текстовой форме (иногда в комбинации с числами). При ответе испытуемые могли использовать клавиатуру и мышь (для выбора вариантов).

Развитие инструментальных средств разработки обеспечило ряд дополнительных возможностей, обеспечивающих различные действия (манипуляции) с объектами, используемыми в ТЗ – в том числе с помощью «компьютерной мыши»: **а)** выбор отдельных частей графических объектов или точек на них в порядке ответов на вопросы ТЗ; **б)** перемещение (перетаскивание) графических объектов по отношению друг к другу; по отношению к некоторым контейнерам и пр.; **в)** установление связей между объектами, включая рисунки и «тексты в рамках», с помощью проведения линий; **г)** задание «сил связей» между объектами путем «маркировки» связывающих их линий; **д)** повороты графических объектов, в том числе вокруг произвольно задан-

ных осей; е) пропорциональное масштабирование графических и текстовых объектов; ж) изменение цветов изображений на графических объектах; з) изменение прозрачности объектов, представленных на мониторе ЭВМ; и) перемещение различного рода «экранных ползунков», вращение «экранных ручек-регуляторов» и пр. Это позволяет устанавливать «непрерывные» или дискретно изменяющиеся числовые значения без использования ввода с клавиатуры. Использование таких «экранных инструментов» особенно важно для представления нечетких ответов; к) увеличение фрагментов экрана (частей объектов) с помощью «экранный лупы»; л) «проговаривание вслух» вопросов и, возможно, ответов в ТЗ. Однако, поскольку ответы могут проговариваться только последовательно во времени, то этот вариант для тестируемого сложнее, чем одновременное восприятие всех ответов на мониторе ПЭВМ; м) получение ответов на ТЗ от испытуемых в звуковой форме, их распознавание и ввод в ТПС. Последние три возможности важны прежде всего для лиц с ограниченными возможностями по зрению.

Типичные классификации ТЗ включают их разделение на два класса: закрытые и открытые [1]. В свою очередь для «закрытых» ТЗ принято выделять [11] три категории: с выбором верных ответов из числа предъявленных; с ранжированием предъявленных объектов (текстовых или графических) по тому или иному критерию; с установлением соответствия между предъявленными объектами. Различные подкатегории и варианты модификации ТЗ для них в случае нечетких ответов рассматриваются далее.

Тестовые задания с предопределенными наборами ответов. При использовании ТЗ этой категории тестируемые лица должны выбирать один или больше верных ответов из числа предъявленных. При этом целесообразно различать несколько подкатегорий, для каждой из которых условие выбора необходимо отразить в формулировке вопроса.

Подкатегория ТЗ с единственным верным ответом из числа предъявленных – для определенности будем считать, что их количество равно N_i в i -м ТЗ. При этом предъявленные ответы могут представлять из себя следующее: только текст, в том числе и на иностранном языке [17]; только числа [5]; комбинации текста и чисел; графические объекты или их комбинации с текстом, числами.

Подход на основе «четких ответов». В типичных случаях ТПС отображает на мониторе ПЭВМ для очередного ТЗ вопрос и все ответы. Рядом с ответами стоят «пустые» кружочки, соответствующие «радиокнопкам». Если тестируемый щелкает мышью внутри контура одного из кружков, то он становится черным, а соответствующий ему ответ рассматривается как «выбранный». Поскольку «выбранным» может быть только один ответ из числа предъявленных, то при щелчке мышью на другом кружке подсвечивается уже он, а с ранее подсвеченного кружка подсветка снимается.

Если при обработке ответа на i -й ТЗ определяется, что он верный, то по этому ответу в сумматор по тесту в целом добавляется балл (B_i), определяемый по формуле

$$B_i = W_i, \quad (1)$$

где W_i – весовой коэффициент (уровень сложности) ТЗ. Отметим, что при адаптивном тестировании этот уровень в общем случае может корректироваться.

При неверном ответе на ТЗ возможны следующие варианты. 1. В сумматор ничего не добавляется. 2. Из сумматора вычитается величина B_i . (3) Из сумматора вычитается $B_i/2$. Таким образом, для тестируемого обеспечивается большая положительная мотивация к ответу, чем отрицательная. 4. Из сумматора вычитается некоторый «штраф», величина которого связана с уровнем сложности ТЗ (чем ТЗ проще, тем штраф должен быть выше). Это означает, что для низких B_i штраф должен быть выше.

Одновременное предъявление всех ответов возможно как в предопределенном (детерминированном) порядке, так и в случайном. Однако во многих случаях необходимо включать ответ типа «Ни один из приведенных ответов не является правильным». При этом целесообразно, чтобы такой ответ был последним по порядку. Следовательно, в этом случае случайный порядок воспроизведения может задаваться только для других ответов, но не для указанного.

Дополнительно отметим, что если тестируемый выбрал ответ типа «Ни один из приведенных ответов не является правильным» и одновременно указал один или большее количество ответов в качестве «верных», то это является логическим противоречием. Такой выбор ответов ТПС принимать не должна. Возможен и вариант с *последовательным предъявлением* ответов испытуемому на экране монитора (с показом ответов «по одному»). При этом целесообразно обеспечить следующее. (y1) Формулировка вопроса должна оставаться на экране при предъявлении всех ответов. (y2) Общее количество ответов, которое соответствует конкретному ТЗ, целесообразно отображать наряду с формулировкой вопроса. Отметим, что это количество при адаптивном режиме тестирования [10, 16] может меняться в порядке регулирования сложностей вопросов. При этом если используется ответ «Ни один из других ответов не является верным», то этот ответ необходимо предъявлять последним по порядку. (y3) Целесообразно указывать порядковые номера для всех ответов, в том числе и в случае предъявления ответов в случайной последовательности. (y4) Длительность показа (демонстрации) на мониторе каждого из ответов может быть либо постоянной, либо регулироваться в зависимости от длины текста ответа. (y5) После того

как тестируемый один из ответов выберет в качестве правильного, то предъявлять последующие ответы нет необходимости – в силу использования модели «четкой логики».

Для ТЗ с нечеткими ответами, предъявляемыми *одновременно*, тестируемый вместо выбора «радиокнопки» (см. выше) должен для каждого ответа ввести с клавиатуры (или задать с помощью «экранных ползунков») вероятность того, что этот ответ верный. При этом может быть целесообразным дать тестируемому возможность возвращаться к ранее установленным вероятностям ответов с целью их корректировки, а обработку ответов проводить только после нажатия кнопки типа «Далее». На рисунке 1 приведен пример представления нечеткого ответа с использованием двух стрелок-указателей, вращаемых вокруг общей оси с помощью курсора мыши. При этом синяя стрелка означает «нижнюю оценку», а красная – «верхнюю оценку». Для голубой и красной стрелок лучше сделать искривленные (по кругу) общие стрелки в двух направлениях – по часовой стрелке и против (т.е. с двумя оперениями стрелок).



Рисунок 1 – Задание нечеткого ответа с помощью стрелок-указателей, вращаемых с помощью их «зацепления и перемещения» курсором мыши

Представленный на рисунке 1 ответ может обрабатываться по-разному. Например, возможны такие варианты. **А.** Проверяется два условия: а1) указанный тестируемым для i -го ТЗ диапазон температур от максимальной до минимальной в виде $\Theta_i^{(max)} \dots \Theta_i^{(min)}$ целиком попадает внутрь «эталонного диапазона», предусмотренного разработчиком ТЗ; а2) ширина диапазона $\Theta_i^{(max)} \dots \Theta_i^{(min)}$ составляет не менее 60 % от ширины эталонного диапазона. При выполнении обоих условий ответ считается верным и тестируемому за i -е ТЗ начисляется предусмотренное количество баллов, т.е. W_i . **Б.** Этот вариант является альтернативным по отношению к «А». Указанный пользователем диапазон температур $\Theta_i^{(max)} \dots \Theta_i^{(min)}$ и эталонный диапазон перекрываются лишь частично. При этом целесообразно принять

$$B_i = W_i \times (\Theta_i^{(max)} - \Theta_i^{(min)}) / (\Xi_i^{(max)} - \Xi_i^{(min)}), \quad (2)$$

где $\Xi_i^{(max)}; \Xi_i^{(min)}$ – соответственно максимальное и минимальное значение для эталонного диапазона значений, заданного разработчиком ТЗ.

В случае формулировки ответа, относящегося к степени «уверенности» в нем тестируемого лица, может использоваться единственный «ползунковый указатель», который также перемещается с помощью курсора мыши (рис. 2).



Рисунок 2 – Интерфейс пользователя для ТЗ с единственным ответом

Возможен также вариант, когда тестируемому предлагается несколько тематически взаимосвязанных ответов, причем по каждому из них он выбирает ответ независимо от других. Тогда целесообразно использовать несколько шкал по типу показанной на рисунке 2, причем ползунки для каждой из шкал могут устанавливаться тестируемым лицом независимо друг от друга (рис. 3) с помощью курсора мыши по технологии «зацепить и переместить».

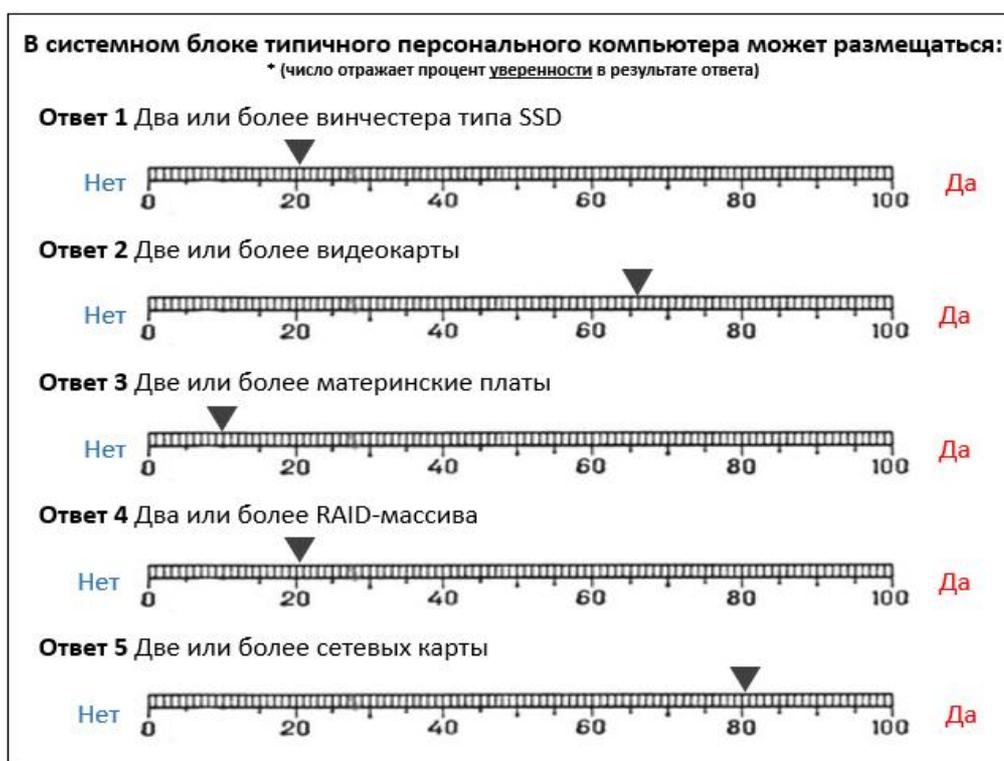


Рисунок 3 – ТЗ с несколькими нечеткими ответами

Возможность визуального сопоставления формулировок ответов и выбранных для них вероятностей того, что эти ответы правильные, облегчает работу тестируемого с тестовыми заданиями.

Обозначим величины вероятностей, установленных тестируемым для ответов, как $\{P_{n,i}\}_{n=1...N_i}$, где n – номера предъявленных ответов в i -м ТЗ; N_i – общее количество ответов в i -м ТЗ, и будем считать,

что $\{P_{n,i}\}$ выражены в процентах. При этом в общем случае сумма вероятностей $S = \sum_{n=1}^{N_i} P_{n,i}$ может быть

не равной 100 %. Это условие существенно, так как ручная «подгонка» суммы вероятностей к 100 % может быть достаточно затруднительной.

В простейшем случае по результатам обработки ответов по i -му ТЗ в сумматор баллов по тесту в целом целесообразно добавлять величину

$$B_i = W_i \left(P_{k,i} / \sum_{n=1}^{N_i} P_{n,i} \right), \quad (3)$$

где k – номер единственного верного ответа; i – номер ТЗ.

Возможна проверка и дополнительного условия

$$\left(P_{k,i} / \sum_{n=1}^{N_i} P_{n,i} \right) \leq Z_i, \tag{4}$$

где Z_i – некоторая пороговая величина, которая, вообще говоря, индивидуальна для каждого ТЗ. Если условие (3) выполняется, то $B_i = 0$ и в сумматор по тесту ничего не добавляется. Иначе – добавляется B_i , вычисленное по формуле «3».

Для i -го ТЗ «коэффициент нечеткости» ответа испытуемого (φ_i) оценим по

$$\varphi_i = 100\% \left(\left(\sum_{n=1}^{N_i} P_{n,i} \right) - P_{m,i} \right) / \sum_{n=1}^{N_i} P_{n,i}, \quad P_{m,i} = \max(P_{n,i})_{n=1 \dots N_i}. \tag{5}$$

Таким образом, индекс « m » соответствует ответу, для которого тестируемый установил (выбрал) наибольшую вероятность – для простоты считаем, что такая вероятность установлена лишь для одного ответа. При этом в общем случае индекс « m » может и не соответствовать верному (правильному) ответу.

По тесту в целом для ТЗ с вопросами рассматриваемого типа на основе формулы (5) могут быть подсчитаны статистические показатели – как при прохождении теста одним тестируемым лицом, так и группой таких лиц.

Подчеркнем, что при таком подходе с точки зрения оценивания выбранных тестируемым вероятностей для ответов в отношении того, что они правильные, все ответы имеют одинаковую «значимость». Однако в общем случае априорные вероятности выбора ответов различны – в силу их разного «правдоподобия». Примем для определенности эти априорные вероятности в виде набора $\{V_{n,i}\}_{n=1 \dots N_i}$, имеющих значения от 0 до 100 %. В то же время средние вероятности правильных ответов для каждого из них при случайном выборе для i -го ТЗ равны

$$\omega_i = 100\% / N_i. \tag{6}$$

При конструировании последующей формулы во внимание целесообразно принять два соображения. **А.** Если $V_{k,i} > \omega_i$, то выбор такого ответа более вероятен, чем при случайном выборе тестируемым одного из ответов. Следовательно, величину B_i в этом случае целесообразно снижать по сравнению с тем, что дает формула. **И,** наоборот, если верный ответ имеет меньшую априорную вероятность по сравнению с ω_i , то в сумматор по тесту целесообразно добавить увеличенное значение B_i . **Б.** В предельном (модельном) варианте, если правдоподобным является только единственный (верный) ответ (т.е. $V_{k,i} = 100\%$, а остальные вероятности равны «0»), то за этот ответ в сумматор по тесту вообще нецелесообразно добавлять баллы. Поэтому в качестве рабочей формулы вместо (3) можно предложить

$$B_i = W_i \left(P_{k,i} / \sum_{n=1}^{N_i} P_{n,i} \right) \times \left(\frac{100 - V_{k,i}}{100 - \omega_i} \right)^{\alpha_1} / (V_{k,i} / \omega_i)^{\alpha_2}, \tag{7}$$

где α_1, α_2 – некоторые константы.

Отметим следующее. **1.** Формула (7) является не единственно возможной. **2.** Сомножитель в (7) $\left(\frac{100 - V_{k,i}}{100 - \omega_i} \right)^{\alpha_1}$ позволяет удовлетворить условие «б». **3.** Делитель $(V_{k,i} / \omega_i)^{\alpha_2}$ обеспечивает удовлетворение условия «а». **4.** Соотношение коэффициентов α_1, α_2 обеспечивает некоторый «баланс» между соблюдением условий «а» и «б». **5.** Для формулы (7) может быть целесообразным принять $\{W_i\}_{i=1 \dots N_i} = 1$, так как сложность ТЗ определяется указанными сомножителем и делителем в правой части этой формулы.

Первым дополнительным решением при использовании нечетких ответов может быть их предьявление с уже установленными ненулевыми величинами вероятностей (например, при помощи положений ползунков), которые испытуемый должен изменить в порядке ответа на ТЗ.

Вторым дополнительным решением может быть использование не единственной вероятности того, что ответ верный, а двух границ для такой вероятности – нижней и верхней. При этом в простейшем случае можно брать для обработки среднее арифметическое значений для верхней и нижней границ, установленных пользователем. Другие подходы также допустимы, но требуют использования более сложного математического аппарата.

Третьим дополнительным подходом является использование не числовых, а вербальных оценок вероятностей того, что соответствующий ответ верен. Например, может быть использована шкала типа «ответ верен; высокая вероятность верного ответа; ответ скорее верен, чем неверен; вероятности

того, что ответ верен и неверен примерно одинаковы; ответ скорее неверен, чем верен; высокая вероятность того, что ответ неверен; ответ неверен). Такой шкале могут быть сопоставлены числовые оценки по специальной схеме.

Если для ввода ответов испытуемыми используются «экранные ползунки», то целесообразно сделать следующее: нанести на зону перемещения каждого «экранный ползунок» штрих-метки; снабдить их соответствующими надписями (см. предыдущий абзац); обеспечить дискретный характер перемещения указателей экранных ползунков между штрих-метками. При этом штрих-метки могут быть индивидуальными для каждого из ползунков, а надписи – общими для всех ползунков.

Если ответы для *ТЗ с нечетким оцениванием* предъявляются *последовательно* (один за другим), то это приводит к следующим изменениям. 1. Для каждого из предъявляемых ответов вероятность того, что он правильный, должна устанавливаться тестируемым независимо от других (вводом с клавиатуры или перемещением единственного «экранный ползунок»). При этом тестируемый не может исправлять «задним числом» вероятности, которые он указал для ранее уже предъявленных ответов. Поскольку тестируемый не может анализировать совокупность ответов в наборе, то этот вариант тестирования для него объективно сложнее. 2. Условие «у5» (см. выше) снимается, так как тестируемому должны быть предъявлены все вопросы.

Подкатегория ТЗ с указанным в задании количеством верных ответов (два или более). Отметим прежде всего следующее. **А.** Для этой категории ТЗ вероятность случайного выбора правильного набора верных ответов существенно ниже, чем для ТЗ с единственным верным ответом. **Б.** Последовательность выбора испытуемым верных ответов из числа предъявленных им считается «незначимой».

Для ТЗ с четкими ответами, предъявляемыми *одновременно*, ответ по «ТЗ в целом» должен считаться верным «тогда и только тогда», когда тестируемым лицом выбраны все верные ответы и не выбран ни один неверный. Для этой категории ТЗ в числе возможных ответов не должно быть варианта «Ни один из представленных ответов не является верным».

В этом случае должен использоваться алгоритм, соответствующий формуле (1). Таким образом, $B_i = 0$ принимается в следующих случаях: тестируемый выбрал не все верные ответы (в частном случае – ни одного верного), при этом количество выбранных неверных ответов уже «не значимо»; выбрал все верные ответы, но наряду с ними выбрал один или больше неверных. Таким образом, возможность отдельной оценки этих вариантов «не используется», хотя потенциально это позволяет обеспечить повышение дифференцирующей способности ТЗ [8, 18].

Отметим, что для ТЗ с «четкой» логикой в случае, если тестируемый выбрал меньшее (по сравнению с тем, что указано в формулировке вопроса) количество ответов в качестве верных, то ответ вообще не должен «приниматься» ТПС. При этом тестируемому должно быть повторно указано на необходимость выбрать заданное количество ответов.

Для ТЗ с четкими ответами, предъявляемыми *последовательно*, в рамках «четкой логики» также должен использоваться алгоритм, соответствующий формуле (1). Однако в этом случае выбор верных ответов для тестируемых значительно сложнее, так как они не могут сравнивать варианты ответов в наборе.

Для ТЗ с нечеткими ответами, предъявляемыми *одновременно*, алгоритмы оценки ответов будут иными. Примем, как и ранее, что тестируемый для каждого из предъявленных ответов указывает вероятности того, что они верные (вводом с клавиатуры или перемещением «экранных ползунков» с помощью манипулятора мышь).

При этом для выбранных неверных ответов целесообразно начислять «штрафные баллы», т.е. брать соответствующие суммы вероятностей со знаком «минус». Тогда вместо формулы (2) будем иметь

$$B_i = W_i \left(\left(\sum_{t=1}^{T_i} P_{t,i} - \sum_{u=1}^{U_i} P_{u,i} \right) / \left(\sum_{t=1}^{T_i} P_{t,i} + \sum_{u=1}^{U_i} P_{u,i} \right) \right), \quad (8)$$

где t, T_i относятся к верным ответам (true); u, U_i – к неверным ответам (untrue). При таком подходе тестируемый за ответ на конкретный ТЗ может набрать не только «положительное» и «нулевое» количества баллов, но и отрицательное. Отметим, что для формулы (8) величина $\left(\sum_{t=1}^{T_i} P_{t,i} + \sum_{u=1}^{U_i} P_{u,i} \right)$ в общем случае может быть не равна «единице».

Обобщением (8) может являться учет сумм вероятностей для «верных» и «неверных» ответов с разными весовыми коэффициентами. В содержательном плане это может означать, что ошибочный выбор тестируемым «неверных» ответов в качестве «верных» будет расцениваться более «негативно», чем «неполный выбор» им всех верных ответов.

Примем, что для дополнительных весовых коэффициентов, задающих баланс между «значимостью» выбора верных (λ^+) и неверных (λ^-), ответов выполняется соотношение

$$\lambda^+ + \lambda^- = 1. \tag{9}$$

При этом в типичных случаях $\lambda^- > \lambda^+$. Тогда вместо (6) можно использовать

$$B_i = W_i \left(\left(\lambda^+ \sum_{t=1}^{T_i} P_{t,i} - \lambda^- \sum_{u=1}^{U_i} P_{u,i} \right) / \left(\lambda^+ \sum_{t=1}^{T_i} P_{t,i} + \lambda^- \sum_{u=1}^{U_i} P_{u,i} \right) \right). \tag{10}$$

При $\lambda^+ = \lambda^- = 1$ формула (10) переходит в (8). Отметим также, что для (10) величина $\left(\lambda^+ \sum_{t=1}^{T_i} P_{t,i} + \lambda^- \sum_{u=1}^{U_i} P_{u,i} \right)$ в общем случае не равна «единице».

Для ТЗ с «нечеткими ответами» в случае, если тестируемый выбрал меньшее (по сравнению с тем, что указано в формулировке вопроса) количество ответов в качестве верных, то ответ может «приниматься» ТПС и обрабатываться по формулам (8) или (10). При этом причиной «неполного» выбора тестируемым заданного количества ответов может быть вероятность получить за «сомнительный» ответ больший отрицательный балл, чем положительный (если $\lambda^- > \lambda^+$).

Дополнительными вариантами обобщения (8) и (10) может быть учет различной «правдоподобности» приводимых ответов – аналогично тому, что описано в предыдущем разделе для ТЗ с единственным верным ответом.

Для данной подкатегории ТЗ также можно использовать показатель «коэффициент нечеткости ответа тестируемого» по i -му ТЗ (обозначим его как φ_i^*). По аналогии с формулой (5) примем

$$\varphi_i^* = 100\% \left(\left(\sum_{n=1}^{N_i} P_{n,i} \right) - \sum_{m=1}^{M_i} P_{m,i} \right) / \sum_{n=1}^{N_i} P_{n,i}, \tag{11}$$

где M – количество выбранных тестируемым лицом ответов с максимальными вероятностями. Отметим, что в общем случае этот набор может не полностью соответствовать совокупности верных ответов. Пусть тестируемый указал вероятности только для « M » из предложенных « N » ответов, а по остальным ответам установил «нулевые» вероятности. Тогда будет $\varphi_i^* = 0$, даже если все или часть установленных тестируемым для « M » ответов «ненулевых» вероятностей не равны 100 %.

Для оценки результатов в случае ТЗ с нечеткими ответами, предъявляемыми последовательно, могут использоваться формулы (8) и (10). Однако такие варианты ТЗ сложнее для тестируемых лиц, так как они последние не могут одновременно анализировать совокупность всех предъявленных ответов путем сравнения их друг с другом.

Подкатегория ТЗ с неопределенным количеством верных ответов в интервале от «нуля» до общего количества предъявляемых. Таким образом, в случае ТЗ, относящихся к этой подкатегории, «предельными» вариантами могут быть: нулевое количество верных ответов в числе предъявленных тестируемому лицу; все предъявленные ответы являются верными.

Для этой подкатегории ТЗ в составе предлагаемых ответов потенциально могут быть такие: **а)** «Все другие ответы являются неверными»; **б)** «Все другие ответы являются верными». В отношении «сочетаемости» таких вариантов ответов с другими типами ответов возможны «логические противоречия», рассматриваемые ниже.

Для ТЗ с четкими ответами, предъявляемыми одновременно, рассмотрим сначала некоторые специальные случаи сочетаний формулировок вопросов и ответов тестируемых лиц. Примем, что в формулировке вопроса (задания) по ТЗ требуется выбрать все верные ответы (варианты ТЗ, в которых требуется выбирать «неверные» ответы/утверждения, на практике встречаются, но достаточно редко).

Пусть в составе предъявленных ответов отсутствует формулировка типа «а», а тестируемый считает все ответы неверными. Тогда он должен не выбирать (не отмечать) вообще ни одного ответа из числа предъявленных.

Пусть теперь в составе предъявленных ответов есть ответ с формулировкой типа «а». Одновременный выбор тестируемым ответа с такой формулировкой и любого другого ответа (или нескольких других ответов) должен рассматриваться в качестве «технической» ошибки – такой выбор ответов ТПС не должна принимать.

Пусть в составе предъявленных ответов отсутствует формулировка типа «б», а тестируемый считает все ответы верными. Тогда он должен в явной форме выбирать (отмечать) все предъявленные ответы.

Пусть теперь в составе предъявленных ответов есть ответ с формулировкой типа «б». Одновременный выбор тестируемым ответа с такой формулировкой и любого другого ответа (или нескольких ответов) является логически противоречивым – такой выбор ответов ТПС не должна принимать.

Примем, что «ответ» тестируемого на ТЗ не относится ни к одному из перечисленных выше «особых» вариантов, которые содержат «логическое противоречие» и потому не должны «приниматься»

ТПС. Тогда обработка ответа возможна по формулам типа (1) – в том числе в таких случаях: все предъявленные ответы являются «неверными» или «верными»; из числа предъявленных верными являются один или большее количество ответов (но не все).

Для ТЗ, предполагающих четкие ответы, предъявляемые *последовательно*, справедливо все, что сказано выше для совокупностей ответов, предъявляемых одновременно. Однако отвечать на такие ТЗ тестируемым значительно сложнее.

Для ТЗ с нечеткими ответами, предъявляемыми *одновременно*, предполагается (как и ранее), что тестируемый может устанавливать (задавать) вероятности того, что предъявленные ответы верные. При этом возможны две ситуации.

1. ТПС не принимает ответ тестируемого на ТЗ в качестве допустимого в силу наличия логических противоречий. Например, тестируемый установил для ответа «Ни один из других предъявленных ответов не является верным» вероятность в 100 %. Однако он одновременно для одного или большего количества других ответов установил ненулевые вероятности того, что ответы являются верными.

Другой вариант логического противоречия. Для ответа типа «Все другие ответы являются верными» установлена 100 % вероятность. Однако по крайней мере для некоторых из других ответов

2. ТПС принимает ответ в качестве допустимого. При этом обработка ответа может осуществляться по формулам типа (8) или (10).

При отсутствии ответов с формулировками типа «Все предъявленные ответы являются верными» или «Все предъявленные ответы являются неверными» логические противоречия в выбранных совокупностях ответов менее вероятны. Однако такие противоречия потенциально могут возникать, если в ТЗ в качестве правильных используются два ответа, формулировки которых полностью или частично противоречат друг другу.

Для ТЗ с нечеткими ответами, предъявляемыми *последовательно*, анализ ответов на отсутствие логических противоречий и обработка ответов могут осуществляться и при одновременном предъявлении всех ответов (см. предыдущий абзац). Однако фактическая сложность ТЗ для тестируемых лиц при этом выше.

Подкатегория ТЗ с вопросами, предполагающими ответы только типа «Да» или «Нет». Такие ТЗ целесообразно рассмотреть отдельно.

Вариант ТЗ с четкими ответами, предъявляемыми *одновременно* или *последовательно*, применяется весьма ограниченно – в силу высокой вероятности правильного ответа тестируемыми лицами при «случайном выборе». При этом последовательность предъявления ответов «Да» и «Нет» может выбираться случайным образом в каждом ТЗ. Может также предъявляться только один ответ (либо «Да», либо «Нет»), и если тестируемый его не выбирает, то ответ считается противоположным предъявленному.

Для варианта ТЗ с нечетким ответом в типичных случаях может использоваться «экранный ползунок», перемещаемый на «оси вероятностей» между «Да» и «Нет» (это соответствует *одновременному* предъявлению обоих ответов). Фактически положение ползунка задает вероятность оценки (в интервале от 100 до 0 %) тестируемым того, что утверждение, сформулированное в вопросе, верное. Указанная ось может быть дополнена вертикальными штрихами, соответствующими вербальным оценкам из совокупности {«Да»; С высокой вероятностью «Да»; Скорее «Да», чем «Нет»; Затрудняюсь выбрать между «Да» и «Нет»; Скорее «Нет», чем «Да»; с высокой вероятностью «Нет»; «Нет»}. При этом штрихи, соответствующие этим вариантам ответов, не обязательно должны быть расположены равномерно на «оси вероятностей».

Альтернативное решение (задание вероятности для ответа типа «Да» вводом числа с клавиатуры) представляется менее удобным.

Формально возможно и *последовательное* получение оценок тестируемым вероятностей для ответов типа «Да» и «Нет». Если сумма этих вероятностей окажется не равной 100 %, то необходимо провести «масштабирование» введенных вероятностей – делением их на эту сумму и умножением на 100 %.

Тестовые задания на ранжирование предложенных объектов. Ранжируемыми объектами могут быть фрагменты текста, статическая графика, комбинации текста и статической графики, объекты динамической графики, в том числе и в сочетании с текстом. Отметим, что понятие «объект динамической графики» в общем случае может относиться не только к пространственному смещению объекта или изменению его формы, но и к изменению цвета объекта или его частей.

Для этой категории ТЗ необходимо все объекты, предназначенные для ранжирования, предъявлять одновременно. При этом в рамках «четких ответов» (четкого оценивания) типично следующее: объекты ранжируются только по одному критерию, причем рангом является место в «ряду предпочтительности»; для одного места в таком ряду может быть задано два и более объектов.

В техническом плане установка места объекта в ряду может обеспечиваться щелчком левой кнопкой мыши при положении курсора в пределах соответствующей клетки. При этом в клетке будет появляться значок «+» или «V».

Распределите накопители информации в порядке убывания их типичной (средней) ёмкости на 01.01.2019 г. («1» соответствует наибольшей ёмкости, «7» – наименьшей)

	1	2	3	4	5	6	7
CD-ROM диск					+		
DVD-ROM диск				+			
Дискета 3.5"						+	
Дискета ZIP Iomega 100Mb							+
HDD		+					
SDD	+						
Флэш- накопитель			+				

Рисунок 4 – Пример выбора тестируемым ответов в ТЗ на ранжирование объектов с использованием «чёткой логики»

Качество (правильность) ранжирования тестируемым объектов для «четких ответов» проще всего оценивать с помощью рангового коэффициента корреляции (по Спирмену) – обозначим его для i -го ТЗ как « C_i ».

$$C_i = 1 - \frac{6 \sum_{n=1}^{N_i} d_{n,i}^2}{N_i(N_i^2 - 1)}; \quad d_{n,i} = R_{n,i}^{(etal.)} - R_{n,i}^{(fact.)}, \tag{12}$$

где N_i – для i -го ТЗ количество объектов в сравниваемых рядах рангов (это количество в обоих рядах одинаково и равно количеству предъявленных тестируемому лицу объектов); $\{R_{n,i}^{(etal.)}\}_{n=1...N_i}$ – ранги объектов в эталонном ряду, созданном разработчиком; $\{R_{n,i}^{(fact.)}\}_{n=1...N_i}$ – ранги объектов в ряду, которые установил тестируемый.

Для C_i справедливо:

$$-1 \leq C_i \leq +1. \tag{13}$$

Поэтому если нужно получить положительную величину оценки в баллах для i -го ТЗ (на ранжирование объектов), то целесообразно использовать формулу:

$$B_i = W_i(C_i + 1). \tag{14}$$

Недостатком такого подхода является то, что для объектов в «начале ряда» (т.е. более важных) и «конце ряда» (т.е. менее важных) отличия от эталона (т.е. предполагаемого «правильного» ранжирования объектов) рассматриваются на одинаковых основаниях. Поэтому может быть целесообразным оценивать качество ответа по i -му ТЗ (S_i) с использованием «взвешенной» суммы квадратов отклонений по рангам, причем весовые коэффициенты могут быть частными от деления «1» на места в эталонном ряду. Предположим для простоты, что объектов с одинаковыми рангами в «эталонном» и «фактическом» рядах нет. Тогда вместо рангового коэффициента корреляции по (13) можно предложить

$$S_i = \left(\sqrt{\sum_{n=1}^{N_i} \left((1/n) \times (R_{i,n}^{(etal.)} - R_{i,n}^{(fact.)})^2 \right)} \right) / N_i. \tag{15}$$

В отношении использования «*нечетких ответов*» по ранжированию объектов возможны два подхода.

1-й подход. Тестируемый указывает степень уверенности (в %) в том, что ответ правильный. В техническом плане это можно реализовать, например, с помощью двух шагов. 1. Тестируемый переводит курсор в необходимую ячейку и щелкает левой кнопкой мыши. В ячейке появляется зеленая полудуга (вся полудуга соответствует 100 % уверенности). Начальное положение риски-указателя устанавливается на 50 % (посередине полудуги). 2. Тестируемый с помощью курсора «зацепляет» риску и перемещает ее в нужную позицию. Одновременно рядом с полукругом высвечивается «% уверенности» в числовой форме (рис. 5). Такое техническое решение представляется авторам более наглядным по сравнению с вводом «% уверенности» с клавиатуры в соответствующую клетку.

Удаление полудуги-указателя из соответствующей клетки может осуществляться путем наведения на нее курсора и щелчка правой кнопкой мыши.

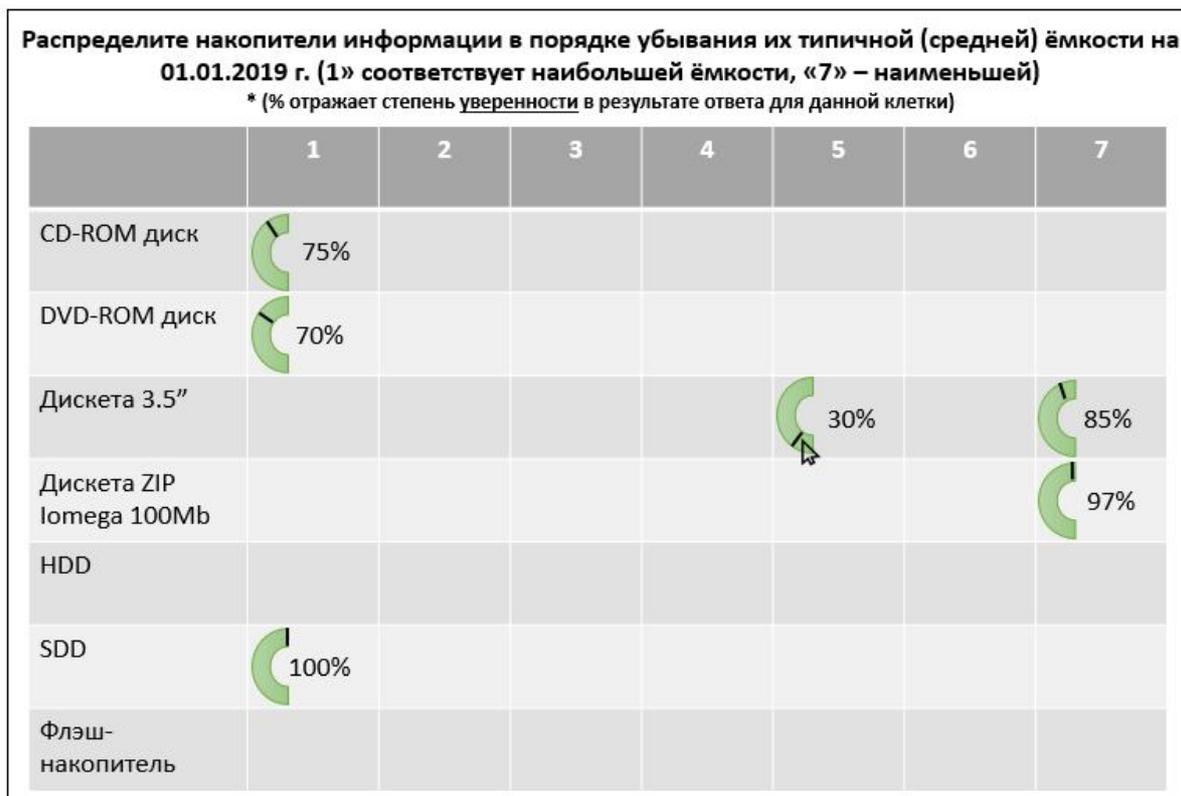


Рисунок 5 – Пример ответа на ТЗ, в котором необходимо проранжировать объекты по определенному критерию (вариант с «нечёткими ответами»)

2-й подход. Тестируемые могут *нечетко оценивать ранги предъявленных объектов*. Для определенности будем считать, что они задают эти ранги в виде пар чисел (в общем случае возможно и не целых), представляющих собой нижнюю и верхнюю границы для рангов по каждому предъявленному объекту. Если тестируемый считает целесообразным дать четкую оценку ранга объекта, то нижняя и верхняя границы для оценки указываются одинаковыми.

Такой подход к нечеткому оцениванию рангов в наглядной форме может быть реализован с помощью горизонтальных экранных ползунков, которые расположены выше и ниже «оси рангов» для каждого объекта. Если необходимо задать одинаковые нижнюю и верхнюю границы для рангов, то ползунки будут размещены друг против друга. Однако в общем случае между ползунками будет некоторое расстояние. Если необходимо задать набор дискретных положений для перемещаемых ползунков, то эти положения целесообразно указывать в виде вертикальных штрихов на оси рангов.

Пример такого решения для одного объекта показан на рисунке 6.

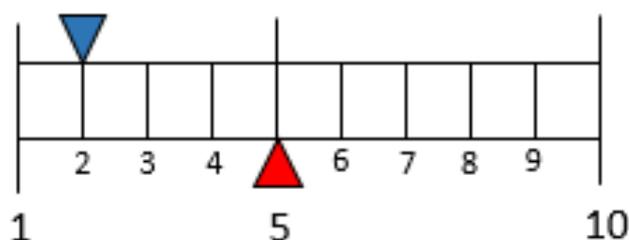


Рисунок 6 – Нечеткое задание ранга для одного предъявленного объекта

В простейшем случае при нечетком оценивании объектов на основе описанных выше подходов для этого типа ТЗ можно брать для каждого из оцениваемых объектов среднее арифметическое нижних и верхних оценок. При этом такие средние арифметические могут быть «нецелыми». Следовательно, разницы рангов в формулах (12) и (15) также могут быть нецелыми.

Альтернативным подходом (по сравнению с использованием «средних арифметических» для верхней и нижней оценок) может быть следующий алгоритм.

** Выполнить расчеты по формулам (12) или (15) для «нижних» оценок рангов для каждого из объектов в предъявленном наборе.

** Затем выполнить расчеты по формулам (12) или (15) для «верхних» оценок рангов по тому же набору объектов.

** Взять среднее арифметическое для двух полученных результатов.

Для категории ТЗ «на ранжирование предъявленных объектов» нечеткость ответа тестируемого лица для i -го ТЗ можно оценить по формуле:

$$\varphi_i^{**} = 100\% \left(\sum_{n=1}^{N_i} \left(\frac{P_{n,i}^{(\max)} - P_{n,i}^{(\min)}}{N_i - 1} \right) \right) / N_i, \quad (16)$$

где N_i – количество предъявленных для ранжирования объектов в i -м ТЗ; $P_{n,i}^{(\max)}$, $P_{n,i}^{(\min)}$ – верхняя и нижняя границы рангов, установленных тестируемым для n -го объекта в i -м ТЗ. Формула (15) обеспечивает следующее: нулевое значение для φ_i^{**} при четких оценках (т.е. $\left\{ P_{n,i}^{(\max)} = P_{n,i}^{(\min)} \right\}_{n=1 \dots N_i}$); значение

$\varphi_i^{**} = 100\%$ в случае, если для всех n объектов в i -м ТЗ тестируемый указал диапазон рангов от «1» до N_i , т.е. до максимального значения.

Могут быть также даны и некоторые оценки «перекрытия» диапазонов интервалов для рангов, выбранных тестируемым для каждого из объектов (по совокупности всех объектов, предъявленных в ТЗ). На основе (16) может быть рассчитана статистика по тесту в целом – для одного тестируемого лица и для группы таких лиц.

Выводы. 1. Использование вариантов ТЗ с четкими ответами имеет как несомненные достоинства, так и некоторые функциональные недостатки. 2. Поэтому применение ТЗ с нечеткими ответами (в том числе в сочетании с традиционными ТЗ, допускающими только четкие ответы) расширяет существующие возможности использования методов компьютеризованной оценки знаний. 3. Использование ТЗ с нечеткими ответами в принципе возможно для всех традиционных типов тестовых заданий, используемых в компьютеризованных системах тестирования настоящее время. 4. Для категории ТЗ с нечетким выбором верных (или неверных) ответов из числа предъявленных обсуждены особенности формулировок ТЗ, варианты предъявления ТЗ тестируемым лицам, средства «нечеткого выбора» ответов тестируемыми, возможные подходы и конкретные алгоритмы обработки «нечеткого выбора». 5. Для категории ТЗ на ранжирование объектов по заданному критерию (включая графические объекты) рассмотрены особенности формулировок заданий, методы представления таких ТЗ на экране ПЭВМ, способы действий испытуемых при нечетком ранжировании, включая использование технологии «Drag and Drop». 6. Представленный теоретический материал проиллюстрирован конкретными примерами.

Библиографический список

1. Аверченков В. И. Система тестирования знаний с поддержкой ответов открытого типа на естественном языке / В. И. Аверченков, Е. А. Белов // Известия Орловского государственного технического университета. Серия: Информационные системы и технологии. – 2006. – № 1–5. – С. 6–10.
2. Артюшкин О. В. Эволюция моделей педагогического тестирования обучающихся в системе высшего профессионального образования / О. В. Артюшкин, Т. А. Артюшкина // Вестник Хакасского государственного университета им. Н.Ф. Катанова. – 2013. – № 3. – С. 119–123.

3. Баин А. М. Математическая модель оценки ответа при дистанционном тестировании знаний / А. М. Баин // Известия высших учебных заведений. Электроника. – 2009. – № 5 (79). – С. 87–90.
4. Близнюк О. Н. Оптимизация тестовых заданий с учетом здоровьесберегающих технологий / О. Н. Близнюк, О. Ю. Русинова, А. Н. Ундозерова // Ученые записки ИУО РАО. – 2018. – № 1 (65). – С. 23–26.
5. Брумштейн Ю. М. Анализ возможностей и технологий применения тестовых заданий с числовыми ответами / Ю. М. Брумштейн, М. В. Иванова, Н. В. Хлопкова // Известия ВолГТУ. Серия «Актуальные проблемы управления, вычислительной техники и информатики в технических системах»: межвуз. сб. науч. ст. / ВолГТУ. – Волгоград, 2012. – № 4 (91). – С. 92–98.
6. Брумштейн Ю. М. Анализ направлений и возможностей использования компьютерных технологий для различных видов тестирования / Ю. М. Брумштейн, Н. В. Хлопкова, Ю. Ю. Аксенова, Ю. Н. Неживая // Теоретические и прикладные вопросы современных информационных технологий: материалы XI Всероссийской научно-технической конференции. – Улан-Удэ: Изд-во ВСГТУ, 2012. – С. 74–79.
7. Брумштейн Ю. М. Тестовые материалы специального назначения, использующие только текст / Ю. М. Брумштейн, Ю. Ю. Аксенова // Интеллектуальная собственность. Авторское право и смежные права. – 2012. – № 8. – С. 33–40.
8. Брумштейн Ю. М. Оценка дифференцирующей способности одного класса тестовых заданий с predeterminedными наборами ответов / Ю. М. Брумштейн, С. В. Окладникова // Прикаспийский журнал: управление и высокие технологии. – 2008. – № 1. – С. 65–70.
9. Буров И. П. Нечеткое оценивание знаний в педагогическом тестировании по информатике / И. П. Буров // Математика. Компьютер. Образование: тезисы Шестнадцатой Международной конференции. – 2009. – С. 478.
10. Гданский Н. И. Адаптивные алгоритмы тестирования знаний информационно-тестирующих приложений на основе стандартных LMS-систем / Н. И. Гданский, М. Л. Рысин, Д. Ю. Альтиментова, Д. Н. Леванов // Ученые записки Российского государственного социального университета. – 2013. – Т. 2, № 5 (120). – С. 77–81.
11. Залан К. В. Виды тестовых заданий и распространенные ошибки при их составлении / К. В. Залан // Вестник Южно-Уральского профессионального института. – 2014. – № 1 (13). – С. 45–51.
12. Ковтуненко А. С. Нечёткие алгоритмы формирования оценок по результатам педагогического тестирования / А. С. Ковтуненко, Г. Ф. Низамова, Е. А. Смирнова // Системы управления и информационные технологии. – 2007. – Т. 27, № 1–1. – С. 163–165.
13. Марков В. В. Организация тестового контроля знаний на основе нечёткой модели экзаменатора / В. В. Марков, М. В. Луцан // Известия ЮФУ. Технические науки. – 2010. – № 7 (108). – С. 242–248.
14. Мицель А. А. Методы тестирования знаний на основе применения аппарата нейронной сети / А. А. Мицель, А. А. Погуда, К. А. Семенов, А. Е. Утешева // Открытое образование. – 2013. – № 2 (97). – С. 34–41.
15. Печников А. Н. Модели и процедуры оценки результатов компьютерного тестирования знаний / А. Н. Печников, А. О. Туровская, Р. Р. Туктаров // Образовательные технологии и общество. – 2013. – Т. 16, № 4. – С. 365–371.
16. Фарфоров А. С. Алгоритмы адаптивного тестирования знаний на основе теории нечётких множеств / А. С. Фарфоров, С. П. Дударов // Успехи в химии и химической технологии. – 2008. – Т. 22, № 1 (81). – С. 63–67.
17. Фурсина Г. А. Различные виды тестов в системе Moodle при обучении иностранным языкам / Г. А. Фурсина, И. В. АLEXИНА // Наука и культура России. – 2014. – Т. 1. – С. 389–392.
18. Черепанов В. С. О необходимости применения различных подходов к оцениванию результатов тестирования с использованием заданий с выбором нескольких правильных ответов / В. С. Черепанов, Т. А. СНИГИРЁВА, Т. Г. СТАНКЕВИЧ // Педагогический журнал Башкортостана. – 2010. – № 1 (26). – С. 85–91.
19. Шестова Е. А. Разработка моделей и методов принятия решений в задачах тестирования знаний: автореф. дис. ... канд. техн. наук / Е. А. Шестова. – Таганрог, 2012.
20. Широков А. С. Возможность автоматизации формирования разных видов тестовых заданий с использованием средств Ms Excel / А. С. Широков, Е. В. Марчук // Проблемы и перспективы развития образования в России. – 2016. – № 38. – С. 60–64.

References

1. Averchenkov V. I., Belov Ye. A. Sistema testirovaniya znaniy s podderzhkoy otvetov otkrytogo tipa na estestvennom yazyke [The system of testing of knowledge with support of answers of public type in a natural language]. *Izvestiya Orlovskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya: Informatsionnye sistemy i tekhnologii* [News of the Oryol State Technical University. Series: Information systems and technologies], 2006, no. 1–5, pp. 6–10.
2. Artyushkin O. V., Artyushkina T. A. Evolyutsiya modeley pedagogicheskogo testirovaniya obuchayushchikhsya v sisteme vysshego professionalnogo obrazovaniya [Evolution of models of pedagogical testing of students in the system of higher education]. *Vestnik Khakasskogo gosudarstvennogo universiteta im. N.F. Katanova* [Bulletin of the Khakass State University named after N.F. Katanov], 2013, no. 3, pp. 119–123.
3. Bain A. M. Matematicheskaya model otsenki otveta pri distantsionnom testirovanii znaniy [A mathematical model of assessment of the answer at remote testing of knowledge]. *Izvestiya vysshikh uchebnykh zavedeniy. Elektronika* [News of higher Educational Institutions. Electronics], 2009, no. 5 (79), pp. 87–90.
4. Bliznyuk O. N., Rusinova O. Yu., Undozerova A. N. Optimizatsiya testovykh zadaniy s uchetom zdorovesberegayushchikh tekhnologiy [Optimization of test tasks taking into account health saving technologies]. *Uchenye zapiski IUO RAO* [Scientific notes of IUO of Russian Joint Stock Company], 2018, no. 1 (65), pp. 23–26.
5. Brumshteyn Yu. M., Ivanova M. V., Khlopkova N. V. Analiz vozmozhnostey i tekhnologiy primeneniya testovykh zadaniy s chislovyimi otvetami [The analysis of opportunities and technologies of application of test tasks with numerical answers]. *Izvestiya VolGTU. Seriya «Aktualnye problemy upravleniya, vychislitelnoy tekhniki i informatiki v tekhnicheskikh sistemakh»*: mezvuz. sb. науч. ст. / ВолГТУ. – Волгоград, 2012. – № 4 (91). – С. 92–98.

eskikh sistemakh» [Izvestiya VOLGTU. Current Problems of Management, ADP Equipment and Information Science in Technical Systems series: interhigher education institution]. Volgograd, 2012, no. 4 (91), pp. 92–98.

6. Brumshcheyn Yu. M., Khlopokova N. V., Aksenova Yu. Yu., Nezhivaya Yu. N. Analiz napravleniy i vozmozhnostey ispolzovaniya kompyuternykh tekhnologiy dlya razlichnykh vidov testirovaniya. [Analysis of the directions and opportunities of use of computer technologies for different types of testing]. *Teoreticheskie i prikladnye voprosy sovremennykh informatsionnykh tekhnologiy : materialy XI Vserossiyskoy nauchno-tekhnicheskoy konferentsii* [Theoretical and applied questions of modern information technologies : materials of the XI All-Russian Scientific and Technical Conference]. Ulan-Ude, Izd-vo VSGUTU, 2012, pp. 74–79.

7. Brumshcheyn Yu. M., Aksenova Yu. Yu. Testovye materialy spetsialnogo naznacheniya, ispolzuyushchie tolko tekst [The special purpose test materials using only the text]. *Intellektualnaya sobstvennost. Avtorskoe pravo i smezhnye prava* [Intellectual property. Copyright and allied rights], 2012, no. 8, pp. 33–40.

8. Brumshcheyn Yu. M., Okladnikova S. V. Otsenka differentsiruyushchey sposobnosti odnogo klassa testovykh zadaniy s predopredelennymi naborami otvetov [Otsenka of the differentiating ability of one class of test tasks with the predetermined sets of answers]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2008, no. 1, pp. 65–70.

9. Burov I. P. Nechetkoe otsenivanie znaniy v pedagogicheskom testirovanii po informatike [Indistinct estimation of knowledge in pedagogical testing of information science]. *Matematika. Kompyuter. Obrazovanie : tezisy Shestnadtsatoy Mezhdunarodnoy konferentsii* [Mathematics. Computer. Education : theses of the Sixteenth International Conference], 2009, p. 478.

10. Gdanskiy N. I., Rysin M. L., Altimentova D. Yu., Levanov D. N. Adaptivnye algoritmy testirovaniya znaniy informatsionno-testiruyushchikh prilozheniy na osnove standartnykh LMS-sistem [Adaptive algorithms of testing of knowledge of the information testing applications on the basis of standard LMS systems]. *Uchenye zapiski Rossiyskogo gosudarstvennogo sotsialnogo universiteta* [Scientific Notes of the Russian State Social University], 2013, vol. 2, no. 5 (120), pp. 77–81.

11. Zalan K.V. Vidy testovykh zadaniy i rasprostranennyye oshibki pri ikh sostavlenii [Types of test tasks and widespread errors by their drawing up]. *Vestnik Yuzhno-Uralskogo professionalnogo institute* [Bulletin of the Southern Ural Professional Institute], 2014, no. 1 (13), pp. 45–51.

12. Kovtunenkov A. S., Nizamova G. F., Smirnova Ye. A. Nechetkie algoritmy formirovaniya otsenok po rezul'tatam pedagogicheskogo testirovaniya [Indistinct algorithms of forming of estimates by results of pedagogical testing]. *Sistemy upravleniya i informatsionnye tekhnologii* [Management System and Information Technologies], 2007, vol. 27, no. 1–1, pp. 163–165.

13. Markov V. V., Lutsan M. V. Organizatsiya testovogo kontrolya znaniy na osnove nechetkoy modeli ekzamenatora [The organization of test control of knowledge on the basis of indistinct model of the examiner]. *Izvestiya YuFU. Tekhnicheskije nauki* [News of SFU. Technical Science], 2010, no. 7 (108), pp. 242–248.

14. Mitsel A. A., Poguda A. A., Semenov K. A., Utesheva A. Ye. Metody testirovaniya znaniy na osnove primeneniya apparata neyronnoy seti [Methods of testing of knowledge on the basis of use of the office of neural network]. *Otkrytoe obrazovanie* [Open Education], 2013, no. 2 (97), pp. 34–41.

15. Pechnikov A. N., Turovskaya A. O., Tuktarov R. R. Modeli i protsedury otsenki rezul'tatov kompyuternogo testirovaniya znaniy [Models and assessment procedures of results of computer testing of knowledge]. *Obrazovatelnye tekhnologii i obshchestvo* [Educational Technologies and Society], 2013, vol. 16, no. 4, pp. 365–371.

16. Farforov A. S., Dudarov S. P. Algoritmy adaptivnogo testirovaniya znaniy na osnove teorii nechetkikh mnozhestv [Algorithms of adaptive testing of knowledge on the basis of the theory of indistinct sets]. *Uspekhi v khimii i khimicheskoy tekhnologii* [Progress in chemistry and chemical technology], 2008, vol. 22, no. 1 (81), pp. 63–67.

17. Fursina G. A., Alekhina I. V. Razlichnye vidy testov v sisteme Moodle pri obuchenii inostrannym yazykam [Different types of tests in the Moodle system when training in foreign languages]. *Nauka i kultura Rossii* [Science and Culture of Russia], 2014, vol. 1, pp. 389–392.

18. Cherepanov V. S., Snigireva T. A., Stankevich T. G. O neobkhodimosti primeneniya razlichnykh podkhodov k otsenivaniyu rezul'tatov testirovaniya s ispolzovaniem zadaniy s vyborom neskol'kikh pravilnykh otvetov [About need of application of different approaches to estimation of results of testing with use of tasks with the choice of several correct answers]. *Pedagogicheskij zhurnal Bashkortostana* [Pedagogical Journal of Bashkortostan], 2010, no. 1 (26), pp. 85–91.

19. Shestova Ye. A. *Razrabotka modeley i metodov priyatiya resheniy v zadachakh testirovaniya znaniy* [Development of models and methods of decision-making in problems of testing of knowledge]. Taganrog, 2012.

20. Shirokov A. S., Marchuk Ye. V. Vozmozhnost avtomatizatsii formirovaniya raznykh vidov testovykh zadaniy s ispolzovaniem sredstv Ms Excel [A possibility of automation of forming of different types of test tasks with use of means of MS of Excel]. *Problemy i perspektivy razvitiya obrazovaniya v Rossii* [Problems and Perspectives of Education Development in Russia], 2016, no. 38, pp. 60–64.