

value lies in applying the Saati method to automate the process of choosing a method for developing a dynamic management model of the personnel security system. The prospect of the study is the development of a dynamic decision-making model with the usage of fuzzy logic tools for the personnel security intelligent management system.

personnel security, dynamic decision making model, economic security, qualitative indicators, multicriterion models, intellectual management, fuzzy logic, Saati method

Одержано (Received) 16.05.2018

Прорецензовано (Reviewed) 22.05.2018

Прийнято до друку (Approved) 28.05.2018

УДК 371.26

JEL Classification C1; C4

DOI: 10.32515/2413-340X.2018.33.189-198

Л.М. Макаренко

Центральноукраїнський національний технічний університет, м. Кропивницький, Україна

Методика шкалювання результатів автоматизованого тестування респондентів соціодинамічних систем

У статті розглянуто питання шкалювання первинних оцінок результатів автоматизованого тестування респондентів соціодинамічних систем щодо лінійних шкал з фіксованим числом діапазонів, а також показано застосування методики шкалювання при поданні результатів тестувань респондентів соціодинамічних систем. Розроблена методика шкалювання первинних тестових оцінок дозволяє одержувати результуючі оцінки в будь-якій лінійній шкалі оцінювання безпосередньо при перегляді результатів. Порівняння отриманих результатів шкалювання зі статистичними даними має практичне співпадіння результатів, що підтверджує достовірність запропонованої методики.

автоматизоване тестування, респонденти, соціодинамічна система, шкалювання нормалізованих оцінок, методика, статистичні дані

Л.М. Макаренко

Центральноукраїнський національний технічний університет, г. Кропивницький, Україна

Представление результатов автоматизированного тестирования респондентов социодинамических систем

В статье рассмотрены вопросы шкалирования первичных оценок результатов автоматизированного тестирования респондентов социодинамических систем по линейным шкалам с фиксированным числом диапазонов, а также показано применение методики шкалирования при представлении результатов тестирования респондентов социодинамических систем. Разработанная методика шкалирования первичных тестовых оценок позволяет получать результирующие оценки в любой линейной шкале оценивания непосредственно при просмотре результатов. Сравнение полученных результатов шкалирования статистических данных имеет практическое совпадение результатов, подтверждающего достоверность предложенной методики.

автоматизированное тестирование, респонденты, социодинамическая система, шкалирование нормализованных оценок, методика, статистические данные

Постановка проблеми. В останні роки автоматизоване комп'ютерне тестування респондентів соціодинамічних систем (РСС) все ширше використовується в Україні. Але інтерпретація результатів тестування та їх використання щодо подальшого відбору РСС без шкалювання результатів не є можливим.

Результати тестування РСС щодо їх професійно важливих якостей, як правило, виражаються в деяких умовних одиницях (оцінках). Правила формування оцінок визначають систему оцінювання, яка може бути шкаловою або нешкаловою. У шкалованій системі результат у вигляді так званих «первинних» тестових оцінок

ставиться в один з діапазонів шкали оцінювання, а в нешкалованих системах оцінювання розраховується безпосередньо за сумою набраних балів з попереднім визначенням умов рейтингування щодо відбору на цю професію. «Первинні» тестові оцінки обчислюються звичайно як частка (відсоток) правильних відповідей серед усіх відповідей на тестові завдання, які давалися респондентові.

Одним з головних факторів, що впливає на спосіб шкалювання, є форма розподілу «первинних» оцінок по оцінюваному діапазону. У тестуванні прийнято вважати, що більшість розподілів «первинних» оцінок найкраще апроксимується на криву нормального закону розподілу, що дає можливість описувати та обробляти результати тестування з використанням загальноприйнятих процедур математичної статистики [2].

Аналіз останніх досліджень і публікацій. Слід зазначити, що на теперішній час достатньо добре опрацьовані питання автоматизованого тестування респондентів з використанням батареї різних тестів, але набагато менше дослідники приділяють увагу багатовимірному шкалюванню (БШ). Так, А.Ю. Терьохіна [13, с. 76] вказує, що “задача БШ полягає у тому, щоб виявити структуру множини стимулів, що вивчаються. Під виявленням структури розуміється виділення набору основних факторів (параметрів), за якими різняться стимули, і опис кожного стимулу у термінах цих факторів (параметрів)”. С.В. Дронов [6, с. 179] вказує, що БШ “...орієнтоване на надання дослідницьким даним наочної структури. Задача БШ вважається вирішеною, якщо вдається зобразити усі дані у вигляді точок простору відносно невеликої розмірності”. А.Д. Наследов [11, с. 299] виділяє мету і завдання БШ. Мета – виявлення структури множини об’єктів. Завдання – “реконструкція психологічного простору, заданого невеликою кількістю вимірів-шкал, і розташування у ньому точок-стимулів так, щоб відстані між ними найкращим чином відповідали вихідним суб’єктивним відмінностям” [11, с. 299]. Методу БШ присвячені також теоретико-методичні роботи М. Дейвісона [5], С.В. Дронова [6], А.Д. Наследова [11], А.Ю. Терьохіної [13], Р.Н. Шепарда [14] та емпіричні дослідження Г.В. Лосика [8], Є.Н. Соколова, Ч.А. Измайлова, В.Л. Завгородньої [12].

З досліджених наукових джерел витікає, що інтенсивний процес розвитку і практичної реалізації різноманітних підходів тестування професійно важливих якостей респондентів соціодинамічних систем відбувається перманентно [4, 9 та ін.]. Багатьма дослідниками (В.В. Мочалін, 1974; Л.С. Нерсисян, О.Н. Конопкін, 1978; Е.М. Псядло, 1999 та ін.) вирішувались питання діагностики окремих професійно важливих якостей респондентів соціодинамічних систем. При цьому абсолютна більшість методик (тестів), що використовуються, передбачає дихотомічні відповіді ("так – ні", "згоден – не згоден") на відповідні питання. Тоді міра властивості певної досліджуваної якості визначається загальною кількістю відповідей згідно "ключа", котра за умови залучення до обстежень представницької вибірки респондентів слугує в подальшому основою для встановлення певної статистичної "норми" [1,3].

Для деяких тестів шкала відповідей має більш деталізований характер і передбачає їх варіювання від абсолютної згоди до абсолютної незгоди з відповідним твердженням. Наведене можна подати, спираючись на методологію нечітких множин як терм-множину (множину термінів, назв) лінгвістичної змінної "ступінь згоди з твердженням". Однак при цьому під час підрахунку результатів тестування звичайно враховуються тільки відповіді у відповідності з "ключем", скажімо, за "повне так" – 2 бали, за "скоріше так, ніж ні" – 1 бал, чому не має пояснення, також як і тому, чому не враховуються інші варіанти відповідей.

Тому представлені у фаховій літературі питання шкалювання первинних оцінок результатів автоматизованого тестування респондентів соціодинамічних систем

потребують подальшого розвитку з метою забезпечення ефективності та результативності методика шкалювання первинних тестових оцінок.

Постановка завдання. Метою даної роботи є розробка методики як наближеного, так і точного лінійного шкалювання нормалізованих «первинних» тестових оцінок і застосування цієї методики при формуванні результатів тестувань в системі РЕМАК.

Виклад основного матеріалу. В системі тестування знань РЕМАК кожен респондент проходить сеанс тестування, що складається з фіксованої і заздалегідь певної кількості N питань різних типів. Кожному питанню ставиться у відповідність максимальний бал за правильну відповідь B_i^{\max} і фактично набраний бал за відповідь B_i . Фактично набраний бал може бути менше максимального. Оцінка за сеанс тестування визначається за формулою (1):

$$B = \frac{\sum_{i=1}^N B_i}{\sum_{i=1}^N B_i^{\max}} * 100 \%, \quad (1)$$

де N – кількість питань в сеансі.

Оцінка B таким чином визначає відсоток (частку) правильних відповідей під час тестування. Для простоти подальшого викладу матеріалу будемо вважати, що в «первинній» оцінці вже усунена частка вгадування правильних відповідей у питаннях закритого типу [10], і аналіз впливу цього фактора не є предметом розгляду в даній роботі.

Недоліком методики перекладу «первинних» тестових оцінок в лінійну шкалу, яка застосовувалася в системі РЕМАК, було прийняття виду розподілу «первинних» оцінок рівномірним і, як наслідок, відсутність їх нормалізації. Для невеликих груп респондентів (до 30 осіб) похибки, що виникають при цьому, не надавали істотного впливу на межі діапазонів шкали. З ростом кількості респондентів, що використовують систему РЕМАК виникає необхідність нормалізації «первинних» тестових оцінок для більш точного обчислення меж діапазонів лінійних шкал оцінювання.

Отримати точні оцінки меж діапазонів лінійної шкали можна тільки по завершенню тестування, коли відомі значення M і σ . Але досить часто респонденти бажають дізнатися свій результат в діапазоні лінійної шкали оцінювання безпосередньо після проходження сеансу тестування. Для поточного перегляду оцінок можна скористатися наближеним методом переведу «первинних» оцінок респондентів в лінійну шкалу на основі ідеалізованої кривої нормального розподілу.

Для визначення способу подання й перетворення первинних тестових оцінок B_i зазвичай розглядають гістограми їх розподілу. Вони дозволяють виявляти ліво- і правосторонню асиметрію, позитивний або негативний ексцес та інші відхилення від нормальності. Для аналізу тестових показників застосовують різні типи лінійного і нелінійного перетворення «первинних» тестових оцінок.

Під стандартною формою розуміють лінійне перетворення нормальної (або штучно нормалізованої) «первинної» тестової оцінки такого вигляду (2):

$$Z_i = \frac{B_i - M}{\sigma}, \quad (2)$$

де Z_i – стандартна тестова оцінка i -го респондента;

B_i – нормальна оцінка i -го респондента;

M і σ – середнє арифметичне значення (медіана) і середнє квадратичне відхилення результату B_i відповідно.

Нормалізованими є стандартні показники, виражені в одиницях розподілу, яке було перетворено з метою його приведення до виду нормальної кривої. Такі показники можна розраховувати за допомогою таблиць, в яких наводиться відсоток випадків, що припадає на ділянки, які відрізняються від середнього значення кривої нормального розподілу на певне число одиниць σ .

Одним з видів нормалізації тестових показників є квантильна стандартизація, при якій кожній оцінці «первинної» шкали B_i присвоюється нове значення її квантильного рангу, який розбиває емпіричний розподіл тестових оцінок на відповідну кількість рівних частин. Квантиль є загальним поняттям, окремими випадками якого є кuartилі, квінтелі, ділили і процентилі. Три кuartильні відмітки (Q_1, Q_2, Q_3) розбивають розподіл тестових оцінок на 4 частини (кuartи) таким чином, що 25% респондентів розташовуються нижче Q_1 , 50% – нижче Q_2 і 75% – нижче Q_3 . Чотири квінтеля (K_1, K_2, K_3, K_4) ділять вибірку аналогічним чином на 5 частин із кроком 20% і дев'ять децилєй (D_1, \dots, D_9) розбивають вибірку на десять частин з кроком 10%, а 99 процентилів – на 100 частин по 1%. Квантильні показники виражаються в одиницях відсотку осіб, що становлять вибірку стандартизації, результат яких нижче встановленої «первинної» оцінки. Головний недолік квантилєв (процентилів) пов'язаний з нерівністю їх як одиниць виміру, особливо на краях розподілу.

Номер відповідного квантилю використовується в якості нової перетвореної тестової оцінки. Квантильна шкала відрізняється тим, що її побудова ніяк не пов'язана з видом розподілу первинних тестових оцінок, яке може бути нормальним або мати будь-яку іншу форму. Єдиною умовою для її побудови є можливість ранжирування респондентів за величиною B_i . Квантильні ранги мають прямокутний розподіл, тобто в кожному інтервалі квантильної шкали міститься однакова частка респондентів. Стандартизація тестових оцінок шляхом їх переведення до квантильної шкали стирає відмінності в особливостях розподілу тестових показників, оскільки зводить будь-який розподіл до прямокутного. Тому з позиції теорії вимірювань квантильні шкали відносяться до шкал порядку: вони дають інформацію, у кого з респондентів сильніше виражена певна властивість, але нічого не дозволяють сказати про те, наскільки або у скільки разів сильніше [7].

Для вирішення поставленої задачі застосовується квантильна стандартизація розбиття вибірки респондентів на D рівних частин, де D – кількість діапазону (бальність) шкали оцінювання. Прийемо, що вибірка респондентів розподілена нормально в довірчому інтервалі 95%.

Виходячи з формулі (1) $B_i = M + Z_i \cdot \sigma$ і розбивши Z – оцінку ідеалізованої кривої нормального розподілу на D частин, можемо отримати усереднені наближені границі діапазонів заданої шкали оцінювання.

Для розбиття Z – шкали на рівні частини скористаємося площею кривої нормального розподілення, яка розраховується з використанням функції Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz.$$

Згідно з визначенням площа нормальної кривої в діапазоні

$-3\sigma \leq x \leq 3\sigma$ приймається рівною 1, тобто $2\Phi(3\sigma) \cong 1$ або $\Phi(3\sigma) = 0,5$. Таким чином, якщо приймати весь діапазон ідеальної кривої нормального розподілу за 100%, то $\sigma_i = 16,7\%$.

Будується ідеалізована крива нормального розподілу з показниками $M_{ідеал} = 50\%$ і $\sigma_{ідеал} = 16,7$ для умовної групи із 100 респондентів з довірчим інтервалом 95%, для якого надійність оцінки $\gamma = 0,95$. Для обчислення меж діапазонів використовується функція Лапласа, яка в довірчому інтервалі буде $\Phi(t) = \frac{\gamma}{2}$, а $t = 1,96$. Таким чином, вся шкала оцінювання, яка знаходиться в довірчому інтервалі від $-1,96\sigma$ до $+1,96\sigma$, приймається за 100%.

1. Відповідно з заданою кількістю діапазонів шкали оцінювання D вся площа ідеалізованої кривої нормального розподілу розбивається на $(D-1)$ ділянок рівної площі. Кожна з цих ділянок містить рівну кількість квантилів.

2. Для розрахунку меж діапазонів шкали оцінювання використовується таблиця значень функції Лапласа в інтервалі від 0 до 0,5. Межі діапазонів визначає зворотна функція Лапласа $x = \Phi^{-1}(\delta)$, де δ – значення функції Лапласа для верхньої межі відповідного діапазону. Вказаний інтервал значень функції Лапласа від 0 до 0,5 розбивається на $\left(\frac{D}{2}-1\right)$ рівних відрізків. Таким чином, границя i -го діапазону в

термінах площини нормальної кривої буде $\delta = 0,5 \frac{0,5}{D} * i = 0,5 - \frac{i}{D}$. Зауважимо, що δ і

σ мають протилежні знаки і дорівнюють 0 при $x_i = M_{ідеал} = 50\%$. Отже

$$x_i = \begin{cases} -\Phi^{-1}(\delta_i) & \text{при } \delta_i > 0 \\ \Phi^{-1}(-\delta_i) & \text{при } \delta_i < 0 \end{cases},$$

при цьому область значень x_i буде $-1,96 \leq x_i \leq 1,96$.

3. Розраховуються верхні межі діапазону на лінійній шкалі оцінювання у відсотках оцінки Z – шкали, виходячи з пропорції:

$$\begin{cases} 2 * t * \sigma_{ідеал} & - & 100\% \\ x_i & - & z_i\% \end{cases}.$$

Отже

$$z_i = \frac{x_i * 100}{2 * t * \sigma_{ідеал}} = \frac{x_i * 100}{2 * 1,96 * \frac{100}{6}} = \frac{x_i * 3}{1,96} = x_i * 1,53.$$

Таким чином, на основі $B_i = M + Z_i * \sigma$ верхні межі діапазонів у відсотках лінійної шкали будуть (3):

$$c_i = M + x_i * 1,53 * \sigma, \quad (3)$$

При цьому необхідно враховувати, що верхня межа останнього діапазону дорівнює 100%, а нижня межа першого діапазону – 1%.

За результатами поточного тестування у даному тесті розраховуються показники реального нормалізованого розподілу респондентів M і σ відносно 100-бальної чи 100% шкали.

Якщо значення M і σ невідомі або не розраховані, то вони приймаються рівними відповідним значенням для ідеалізованої кривої нормального розподілу, а саме, $M = 50\%$, $\sigma = 16,7\%$, і верхні межі діапазонів при цьому будуть (4):

$$c_i = M + x_i * 1,53 * \sigma = 50\% + x_i * 1,53 * 16,7\% = 50\% + x_i * 25,55\%. \quad (4)$$

Для ідеалізованих кривих нормального розподілу з $M=50\%$, $\sigma=16,7\%$ межі діапазонів при $D=10$ наведені в табл. 1, а при $D=12$ в табл. 2.

Таблиця 1 – Діапазони ідеалізованих кривих нормального розподілу

Діапазон (i)	1	2	3	4	5	6	7	8	9
$\frac{i}{D} * 100$	10	20	30	40	50	60	70	80	90
δ	0,4	0,3	0,2	0,1	0	-0,1	-0,2	-0,3	-0,4
x_i	-1,28	-0,84	-0,53	-0,25	0	0,25	0,53	0,84	1,28
c_i	17	28	36	44	50	56	64	72	83

Джерело: розроблено автором.

Таблиця 2 – Діапазони ідеалізованих кривих нормального розподілу

Діапазон (i)	1	2	3	4	5	6	7	8	9	10	11
$\frac{i}{D} * 100$	8,3	16,7	25	33,3	41,6	50	58,4	66,7	75	83,3	91,7
δ	0,42	0,33	0,25	0,16	0,08	0	-0,08	-0,16	-0,25	-0,33	-0,42
x_i	-1,41	-0,97	-0,67	-0,43	-0,21	0	0,21	0,43	0,67	0,97	1,41
c_i	14	25	33	39	45	50	55	61	67	75	86

Джерело: розроблено автором.

Розглянемо, як співвідноситься запропонована методика розрахунку меж діапазонів лінійної шкали оцінювання з результатами реального тестування, отриманого в ході проведення тестування серед респондентів.

У табл. 3 наведені межі діапазонів 12-бальної шкали оцінювання для визначення професійно важливих якостей респондентів соціодинамічних систем таких як схильність до ризику та самовпевненість з параметрами розподілу респондентів $M=48\%$, $\sigma=14\%$ і $M_2=39\%$, $\sigma_2=14\%$.

Таблиця 3 – Межі діапазонів 12-бальної шкали оцінювання для визначення професійно важливих якостей

Діапазон		1	2	3	4	5	6	7	8	9	10	11
1	Схильність до ризику	7	12	18	23	29	35	42	50	57	62	68
2	Самовпевненість	13	22	30	38	43	48	53	58	64	73	83

Джерело: розроблено автором.

Для зазначених параметрів реальних розподілу респондентів ($M_1=39\%$, $\sigma_1=14\%$, і $M_2=48\%$, $\sigma_2=14\%$) за умови ідеалізованого нормалізованого розподілу респондентів верхні межі діапазонів 12-бальної шкали приведені в табл. 4.

Таблиця 4 – Розподіл респондентів по верхніх межах діапазонів 12-бальної шкали

Діапазон	1	2	3	4	5	6	7	8	9	10	11
x_i	-1,41	-0,97	-0,67	-0,43	-0,21	0	0,21	0,43	0,67	0,97	1,41
$c_i(M_1, \sigma_1)$	8	16	25	30	34	39	44	50	53	60	70
$c_i(M_2, \sigma_2)$	17	27	34	39	43	48	53	59	62	69	79

Джерело: розроблено автором.

Перший розподіл (другий рядок табл. 3) має лівосторонню (позитивну) асиметрію, тому межі діапазонів зміщені в область менших значень в порівнянні з ідеалізованим розподілом (до 7% розбіжностей для 3-го діапазону). Другий розподіл (третій рядок табл. 3) ближче до нормального, але має негативний ексцес (середня важкість завдань з тесту досить висока – 33,6%), тому початкові діапазони (з 1-го по 3-й) і останні (з 10-го по 12-й) дещо розширені (розбіжність до 5%).

Таким чином, проведені порівняння запропонованої методики з результатами реального тестування дозволяють зробити висновок, що розрахунок усереднених меж діапазонів шкал оцінювання на основі ідеалізованих нормальних розподілів дає цілком прийнятні результати з похибкою, що не перевищує в середньому 5%. Разом з тим запропонована методика шкалювання нормалізованих оцінок має систематичну погрішність у районі двох верхніх і двох нижніх процентільних діапазонів, що обумовлено прийнятим довірчим інтервалом нормального розподілу в 95%. Але ця похибка істотно не впливає на результат шкалювання, так як в ці діапазони потрапляє не більше 0,1% від загальної кількості респондентів в реальній вибірці.

В системі РЕМАК для реалізації описаної методики шкалювання оцінок в базу даних закладена таблиця значень x_i для 99 верхніх меж діапазонів 100-бальної шкали оцінювання в порядку зростання від 1 до 99.

Для віднесення «первинної» оцінки до відповідного діапазону шкали оцінювання виконуються наступні перетворення.

1. Виконується стандартизація «первинної» оцінки:

$$x_i = \frac{B_i - M}{1,53 * \sigma}.$$

Як згадувалося вище, якщо значення M і σ невідомі або не обчислені, то вони приймаються $M = 50\%$, $\sigma = 16,7\%$.

2. По таблиці значень сотих часток x_i базі даних РЕМАК знаходиться відсоткове значення оцінки в лінійній шкалі β .

3. Оцінка у відповідній лінійній шкалі оцінювання буде визначатися номером діапазону, який обчислюється як найближче найбільше ціле від i , тобто

$$i = \left\lceil \frac{\beta}{\frac{100}{D}} \right\rceil = \left\lceil \frac{\beta * D}{100} \right\rceil$$

В системі РЕМАК лінійне (квантильне) шкалювання оцінок може виконуватися безпосередньо після проходження сеансу тестування при отриманні екранної форми перегляду результатів і після проходження тесту всіма респондентами при аналізі статистики по тесту в цілому.

В першому випадку шкалювання здійснюється виходячи з припущення, що розподіл респондентів близький до нормального з параметрами $M = 50\%$, $\sigma = 16,7\%$. З

одного боку, таке припущення не має під собою достатнього обґрунтування, але, з іншого боку, дозволяє хоча б наближено представити результати тестування в термінах звичної шкали оцінювання.

В екранній формі перегляду результатів системи РЕМАК ключовими параметрами є часовий інтервал (рік, місяць, число), коли проводилося тестування, ім'я тесту, назва групи (в «випадаючих» списках у верхній частині екрану) і бальність шкали оцінювання.

У формі для перегляду вказується список респондентів, час початку сеансу тестування, тривалість сеансу тестування в хвилинах, ім'я (логін) викладача, який проводив тестування, число питань у сеансі тестування, бал, набраний в ході сеансу кожним студентом (кількість правильних відповідей), відсоток правильних відповідей і оцінка в обраній шкалою оцінювання, розрахована за описаною методикою з урахуванням нормалізації вибірки респондентів. Дана форма дозволяє виконувати впорядкування респондентів за зростанням (спаданням) у будь-якому з стовпців форми. Крім того, існує можливість отримати тверду (друковану) копію відомості результатів тестування).

У другому випадку, після завершення всіх тестувань з даного тесту, при аналізі статистики по тесту в цілому на підставі матриці результатів тестування обчислюються параметри M і σ , що характеризують розподіл респондентів за шкалою оцінювання. Після цього може бути сформована таблиця меж діапазонів вибраної шкали оцінювання та список респондентів з оцінками в цій шкалі.

Висновки та перспективи подальших досліджень. Проведене дослідження дозволяє зробити такі висновки:

1. Розроблена методика шкалювання первинних тестових оцінок дозволяє одержувати результуючі оцінки в будь-якій лінійній шкалі оцінювання безпосередньо при перегляді результатів.

2. Результуюча оцінка може бути отримана безпосередньо після проходження сеансу тестування кожним респондентом, виходячи з припущення про нормальний розподіл 95% респондентів, або за результатами реального розподілу респондентів після проходження всього тесту.

3. Порівняння отриманих результатів шкалювання зі статистичними даними показало практичне співпадіння результатів, що підтверджує достовірність запропонованої методики.

4. Розроблена комп'ютерна система тестування знань РЕМАК, яка підтримує основні типи питань закритого типу (типу «вибір одного з декількох», «відповідність»), протягом року функціонує у Центральноукраїнському національному технічному університеті на базі кафедри економічної теорії, маркетингу та економічної кібернетики.

Враховуючи практичну спрямованість даного дослідження, подальші розробки в цьому напрямі мають концентруватися на вдосконаленні методики шкалювання первинних тестових оцінок.

Список літератури

1. Анастаси А. Психологическое тестирование. В 2-х кн.: пер. с англ. [Текст] / А. Анастаси.; под ред. К.М. Гуревича. В.И. Лубовского. – М.: Педагогика. 1982. – Кн. 1. – 320 с., Кн. 2. – 336 с.
2. Анастаси А. Психологическое тестирование [Текст] / А. Анастаси, С. Урбина; 7-е международное издание. – СПб: Питер, 2003. – 688с.
3. Гуревич К.М. Статистическая норма или психологический норматив? [Текст] / К.М. Гуревич, М.К. Акимова, В.Т. Козлова // Психологический журнал. – 1986. – Т.7, № 3. – С.136–142.

4. Гласс Дж. Статистические методы в педагогике и психологии [Текст] / Дж. Гласс, Дж. Стенли; Общ.ред. Ю.П. Адлера; Пер. с англ. Л.И. Харусовой. – М. : Прогресс, 1976. – 496 с.
5. Дэйвисон М. Многомерное шкалирование. Методы наглядного представления данных [Текст] / Пер. с англ. В.С. Каменского. – М.: Финансы и статистика, 1988. – 254 с.
6. Дронов С.В. Многомерный статистический анализ [Текст] / С.В.Дронов. – Барнаул: Изд-во Алтайского гос. ун-та, 2003. – 213 с.
7. Дюк В.А. Компьютерная психодиагностика [Текст] / В.А. Дюк. – СПб: Братство, 1994. – 364с.
8. Лосик Г.В. Исследование восприятия гласных методом многомерного шкалирования психодиагностика [Текст] / Г.В. Лосик // Психологический журнал. – 1992. – Т.13, №2. – С. 67–75.
9. Макаренко Н.В. Теоретические основы и методики профессионального психофизиологического отбора военных специалистов [Текст] / Н.В. Макаренко. – К.: НИИ проблем военной медицины, 1996. – 336 с.
10. Напрасник С.В., Цимбалюк Е.С., Шкиль А.С, Гаркуша Е.В. Технология построения вопросов закрытого типа с нулевой вероятностью угадывания в системе тестирования OpenTEST2 [Текст] // Образование и виртуальность – 2007. Сборник трудов 11 международной конф. УАДО, Харьков-Ялта, 2007. – Харьков: ХНУРЭ, – 2007. – С. 346–353.
11. Наследов А.Д. Математические методы психологического исследования [Текст] / А.Д. Наследов. – СПб.: Речь, 2004. – 392 с.
12. Соколов Е.Н. Многомерное шкалирование знаковых конфигураций [Текст] / Е.Н. Соколов, Ч.А.Измайлов, В.Л. Завгородняя // Вопросы психологии. – 1985. – №1. – С. 133–139.
13. Терехина А. Ю. Многомерное шкалирование в психологи [Текст] / А. Ю. Терехина // Психологический журнал. – 1983. – Том 4, №1. – С.76–88.
14. Шепард Р.Н. Многомерное шкалирование и безразмерное представление различий [Текст] / Р.Н. Шепард // Психологический журнал. – 1980. – Т. I, 4. – С. 72–83.

References

1. Anastazi, A. (1982). *Psychological testing*. K.M. Gurevicha, & V.I. Lubovskogo (Ed.). (Vol. 1-2). Moscow: Pedagogika [in Russian].
2. Anastazi, A. & Urbina, S. (2003). *Psychological testing*. (7n ed.). (Trans). SPb: Piter [in Russian].
3. Gurevich, K.M., Akimova, M.K., & Kozlova, V.T. (1986). Statisticheskaja norma ili psihologicheskij normativ? [Statistical standard or psychological standard?]. *Psihologicheskij zhurnal – A psycho. f* .Vol. 7, 3, 136-142 [in Russian].
4. Glass, Dzh., & Stenli, Dzh. (1976). *Statistical methods in pedagogy and psychology* Ju.P. Adlera (Ed.). (L.I. Harusovoj, Trans). Moscow: Progress [in Russian].
5. Dejvison, M. (1988). *Multidimensional Scaling. Methods for visualization of data*. (B.C. Kamenskogo, Trans). Moskov: Finansy i statistika [in Russian].
6. Dronov, S.V. (2003). *Mnogomernyj statisticheskij analiz [Multivariate statistical analysis]*. Barnaul: Izd-vo Altajskogo gos. un-ta [in Russian].
7. Djuk, V.A. (1994). *Komp'juternaja psihodiagnostika [Computer psychodiagnostics]*. SPb: Bratstvo [in Russian].
8. Losik, G.V. (1992). Issledovanie vosprijatija glasnyh metodom mnogomernogo shkalirovanija [The study of the perception of vowels by the method of multidimensional scaling]. *Psihologicheskij zhurnal – Psychological Journal, Vol.13, 2* [in Russian].
9. Makarenko, N.V. (1996). *Teoreticheskie osnovy i metodiki professional'nogo psihofiziologiskogo otbora voennyh specialistov [Theoretical bases and methods of professional psychophysiological selection of military experts]*. Kiev: NII problem voennoj mediciny [in Russian].
10. Naprasnik, S.V., Cimbajuk, E.S., Shkil', A.S., & Garkusha, E.V. (2007). Tehnologija postroenija voprosov zakrytogo tipa s nulevoj veroyatnost'ju ugadyvanija v sisteme testirovanija OpenTEST2 [Closed-type closed-loop questioning technology with zero probability of guessing in OpenTEST2 testing system]. *Education and virtuality: 11 mezhdunarodnoj konf. UADO (2007 hod) – 11 International conferences UADO*. (pp. 346-353). Har'kov: HNURJe [in Russian].
11. Nasledov, A.D. (2004). *Matematicheskie metody psihologicheskogo issledovanija [Mathematical methods of psychological research]*. SPb.: Rech' [in Russian].
12. Sokolov, E.N., Izmajlov Ch.A., & Zavgorodnjaja, V.L. (1985). Mnogomernoe shkalirovanie znakovyh konfiguracij [Multidimensional scaling of sign configurations]. *Voprosy psihologii – Questions of Psychology, 1*, 133-139 [in Russian].
13. Terehina, A. Ju. (1983). Mnogomernoe shkalirovanie v psihologii [Multidimensional scaling in psychologists]. *Psihologicheskij zhurnal – Psychological Journal, Vol. 4, 1*, 76-88 [in Russian].

14. Shepard, R. N. (1980). Mnogomernoe shkalirovanie i bezrazmernoe predstavlenie razlichij [Multidimensional scaling and dimensionless representation of differences]. *Psihologicheskij zhurnal – Psychological Journal*, Vol. I, 4, 72-83 [in Russian].

Lina Makarenko

Central Ukrainian National Technical University, Kropyvnytskyi, Ukraine

Representation of Results of Automated Testing of Respondents of Sociodynamic Systems

The article deals with the problem of scaling primary assessments of the results of automated testing of respondents of social-dynamical systems in relation to linear scales with a fixed number of ranges, and also shows the application of scalping techniques in presenting the results of tests of respondents of sociodynamic systems.

The purpose of this work is to develop a technique for both approximate and accurate linear scaling of normalized "primary" test ratings and the application of this technique in the formation of test results in the REMAK system. In the REMAK system, linear (quantile) scaling of scores can be performed immediately after passing the testing session when receiving an on-screen form for viewing the results and after passing the test by all respondents when analyzing the statistics for the test as a whole. The disadvantage of the methodology for the translation of "primary" test scores into the linear scale used in the REMAK system was to adopt the kind of distribution of "primary" assessments uniform and, as a consequence, the lack of their normalization. For small groups of respondents (up to 30 people), errors occurring at the same time did not have a significant effect on the boundaries of scale ranges. With the growth in the number of respondents using the REMAK system, there is a need for the normalization of "primary" test ratings to more accurately calculate the limits of the ranges of linear scales.

As a result of the study, a method was developed for scaling primary test scores, which allows you to obtain the resulting scores in any linear scale of evaluation directly when viewing the results. When testing respondents of sociodynamic systems, a result evaluation was obtained immediately after passing the testing session by each respondent, based on the assumption of a normal distribution of 95% of respondents, or on the results of the actual distribution of respondents after passing the entire test. It should be noted that the comparison of the results of scaling with the statistical data showed a practical coincidence of the results, which confirms the reliability of the proposed method.

automated testing, respondents, sociodynamic system, scaling of normalized assessments, methodology, statistical data

Одержано (Received) 19.04.2018

Прорецензовано (Reviewed) 22.05.2018

Прийнято до друку (Approved) 28.05.2018