

АНАЛІЗ МЕТОДІВ ВИЗНАЧЕННЯ УЗГОДЖЕНОСТІ ДУМКИ ГРУПИ ЕКСПЕРТІВ ПІД ЧАС ОЦІНЮВАННЯ РІВНЯ СФОРМОВАНОСТІ ІНШОМОВНОЇ КОМУНІКАТИВНОЇ КОМПЕТЕНТНОСТІ СЛУХАЧІВ

Застело Ольга В'ячеславівна,

старший викладач кафедри іноземної підготовки, girzov@ukr.net.

Анотація. У статті розглядаються сучасні математичні методи визначення узгодженості думки групи експертів у процесі тестування слухачів з іноземних мов.

Ключові слова: експертне оцінювання, узгодженість групи експертів, коефіцієнт рангової кореляції Спірмена, коефіцієнт конкордації Кендала.



У контексті сучасної світової тенденції розвитку іноземної освіти теорія тестування висуває такі питання, відповідь на які сприятиме її подальшому вдосконаленню. Одним із таких питань є покращення якості тестування з іноземних мов, зокрема підвищення надійності отриманих результатів оцінювання рівня сформованості іноземної комунікативної компетентності (ІКК) слухачів, що, у свою чергу, впливає на вдосконалення процесу навчання іноземних мов.

Сучасним підходом до проведення комплексного тестування з іноземних мов є метод експертного оцінювання. Експертні методи постійно розвиваються й удосконалюються. Основні напрямки цього розвитку визначаються низкою факторів, серед яких можна вказати на прагнення розширити сфери застосування, підвищити ступінь використання математичних методів і електронно-обчислювальної техніки, а також знаходити шляхи усунення недоліків.

Метою оброблення інформації від експертів є отримання узагальненої оцінки рівня сформованості ІКК слухачів. Під час математичного обчислення експертних оцінок важливо також визначити узгодженість інформації від експертів у групі.

Оскільки оцінки експертів можуть розглядатися як випадкові величини, то для аналізу розкиду й узгодженості оцінок використовують подані нижче узагальнені статистичні характеристики.

1. Середнє значення оцінок, що характеризує узагальнену думку експертів.
2. Середньоквадратичне відхилення, що характеризує розкид думок експертів від середнього значення.
3. Коефіцієнт варіації, що характеризує варіабельність оцінок.
4. Показники суперечливості думок експертів, коли розглядається гіпотеза про аномальність думок експертів стосовно узагальненої думки всіх експертів.

Нині під час оцінювання за допомогою методу ранжирування для аналізу узгодженості оцінок, вимірних за порядковою шкалою й отриманих від окремих експертів, доцільно використовувати сучасні математично-статистичні методи визначення узгодженості інформації від групи експертів. Необхідно розглянути, застосування якого з них дозволяє отримати достовірніші результати ступеня узгодженості ранжирувань групи експертів у цілому і для заданих

парних порівнянь експертів, що, у свою чергу, впливає на точність оцінювання рівня сформованості ІКК слухачів.

Аналіз останніх досліджень і публікацій. Питання теорії і практики, пов'язані з типовими стадіями експертного оцінювання, а саме з методами оброблення інформації від експертів і оцінювання компетентності експертів, а також використання методу ранжирування в експертному оцінюванні висвітлені у працях багатьох науковців, зокрема С. Д. Бешелева [1], В. С. Кіма [2], О. І. Орлова [5; 6], О. І. Мухіна [3], А. І. Субетто [8], С. П. Хабарова [9] та інших.

Математичні методи визначення узгодженості думки групи експертів висвітлені в працях багатьох вітчизняних і зарубіжних дослідників, зокрема, Д. Бонета, Т. Райта, [12] Дж. Дуркіна [13], В. П. Новосада [4], А. Н. Павлова [7], А. І. Чегодаєва [10], В. С. Черепанова [11] та інших. Проте доцільність застосування того чи іншого математичного методу визначення узгодженості думки групи експертів під час оцінювання рівня сформованості ІКК слухачів ще недостатньо досліджено.

Метою цієї статті є обґрунтування застосування певного математичного методу визначення узгодженості думки групи експертів під час експертного оцінювання рівня сформованості ІКК слухачів. Задля досягнення цієї мети слід оцінити надійність і переваги певного математичного методу визначення думки групи експертів порівняно з іншими методами.

У практиці проведення тестування з іноземних мов, зазвичай, працюють три і більше експертів (екзамінаторів). Отже, виникає необхідність визначити узгодженість думки всієї групи експертів, тобто зробити висновок про одну з найсуттєвіших характеристик якості результатів групового експертного оцінювання.

Для оцінювання узгодженості одержуваних результатів застосовуються такі **статистичні показники:**

- коефіцієнт конкордації;
- коефіцієнт згоди за парного порівнянні;
- дисперсійний коефіцієнт (для ранжируваних рядів).

Розглянемо ранжирування, які уклав кожний з експертів у групі: $\rho_{i1}, \rho_{i2}, \dots, \rho_{in}$. Кожна з двох сукупностей балів, що виставили експерти, розташовується у вигляді варіаційного ряду з наданням кожному члену ря-

ду відповідного порядкового номера (рангу), вираженого натуральним числом. Порівнювані ознаки можна ранжувати в будь-якому напрямку: як у бік погіршення якості (ранг 1 отримує слухач, який під час тестування найкраще виконав завдання), так і навпаки. Головне, щоб обидві змінні були проранжировані однаково.

Наприклад, 4 експерти проранжировали результати виконання одним слухачем 5-ти різних тестів, що відображено в табличній формі (табл. 1).

Таблиця 1

Приклад ранжировання групою експертів результатів виконання одним слухачем 5-ти різних тестів

Експерти	Тести				
	Тест 1	Тест 2	Тест 3	Тест 4	Тест 5
Експерт 1	1	2	3	4	5
Експерт 2	3	4	2	1	5
Експерт 3	2	1	3	5	4
Експерт 4	4	3	2	1	5

Найбільш грубий підхід заснований на обчисленні коефіцієнтів асоціації за Устюжаніновим, за допомогою яких враховуються лише кількість збіжних і незбіжних відповідей, але не враховується їх послідовність.

Коефіцієнти асоціації за Устюжаніновим використовуються для оцінювання міри схожості думок кожної пари експертів. Інформаційна міра схожості відповідей двох експертів визначається за формулою:

$$S_{ij} = \frac{2m_{ij}}{t_i \log_2(1 + \frac{t_j}{t_i}) + t_j \log_2(1 + \frac{t_i}{t_j})}, \quad (1)$$

де m_{ij} — кількість ознак однаково оцінених i -м та j -м експертом;

t_i — кількість ознак, оцінених i -м експертом;

t_j — кількість ознак, оцінених j -м експертом.

На рис. 1 наведено вікно статистичного програмного засобу «OPINION» з результатами оцінювання коефіцієнтів асоціації за Устюжаніновим під час оброблення думок експертів за даними з таблиці 1.

Коли необхідно визначити узгодженість у ранжированні групи, де більше двох експертів, тоді обчислюється коефіцієнт конкордації Кендала — загальний коефіцієнт рангової кореляції для групи, що складається з m експертів. Коефіцієнт конкордації часто використовується в експертному оцінюванні, на-

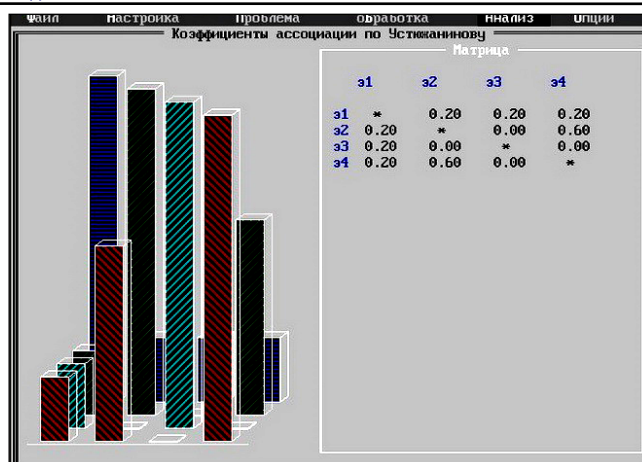


Рис. 1. Екранне вікно програми «OPINION» з результатами обчислення коефіцієнтів асоціації за Устюжаніновим

приклад, для того, щоб визначити ступінь узгодженості думок експертів щодо важливості того чи іншого оцінюваного показника.

Коефіцієнт конкордації W дозволяє оцінити, наскільки узгодженими між собою є ранжировання n об'єктів, побудовані групою m експертів $\|r_{ij}\|$ ($j=1, \dots, m$; $i=1, \dots, n$), де r_{ij} — ранг, що надається j -м експертом i -му об'єкту [1, 2, 7, 9, 10, 11]. Він визначається як відношення D , що описує розкид між ранжированнями до величини D_{max} , що є максимально можливим розкидом й обчислюється за формулою:

$$W = \frac{D}{D_{max}} = \frac{12 \cdot S}{m^2(n^3 - n)}, \quad (2)$$

де m — кількість експертів,

n — кількість слухачів.

Для обчислення суми квадратів відхилень рангів від середнього значення S :

$$S = \sum_{i=1}^n \left(\sum_{j=1}^m r_{ij} - r \right)^2 \quad (3)$$

яка входить до співвідношення (2), додатково будують спеціальну розрахункову таблицю. Потрібні для обчислення коефіцієнта конкордації дані подано у табл. 2.

Коефіцієнт W вимірюється в діапазоні від 0 до 1 і групова оцінка вважається достатньо надійною, коли

Таблиця 2

Розрахунок квадратів відхилень рангів кожного тесту — об'єкта експертизи від середньої суми рангів

Тести	Експерти				S	d_i	d_i^2
	Експерт 1	Експерт 2	Експерт 3	Експерт 4			
Тест 1	1	2	3	4	10	2	4
Тест 2	2	4	1	3	10	2	4
Тест 3	3	2	3	2	10	2	4
Тест 4	4	1	5	1	11	1	1
Тест 5	5	5	4	5	19	-7	49
$\sum d_i^2$							62

S — сумарний ранг об'єкта; d_i — відхилення суми рангів для i -го об'єкта від середньої суми рангів;
 d_i^2 — квадрат відхилення; $\sum d_i^2$ — сума квадратів відхилень.

$W \geq 0,7$. Якщо коефіцієнт конкордації дорівнює нулю, то між думками експертів наявна абсолютна розбіжність. Якщо коефіцієнт конкордації дорівнює одиниці, тоді має місце повна узгодженість думок експертів стосовно результатів виконання слухачем кількох тестів з іноземної мови. Підставивши S у формулу (2), отримуємо $W = 0,3975$ — коефіцієнт конкордації (узгодженості думок експертів). Така величина коефіцієнта свідчить про наявні розбіжності в інформації від експертів, але зовсім не дозволяє охарактеризувати суть цих розбіжностей.

За наявності неузгодженості думок експертів доцільно продовжити аналіз для виявлення причин їх неоднорідності, для чого проводиться перевірка на несуперечність думок, у ході якої виявляються експерти, думки яких істотно відрізняються від загальної думки групи.

Для полегшення обчислення коефіцієнта конкордації Кендала рекомендуємо використовувати статистичний програмний засіб «AtteStat». Приклад результатів обчислення цією програмою коефіцієнта конкордації наведено на рис. 2.

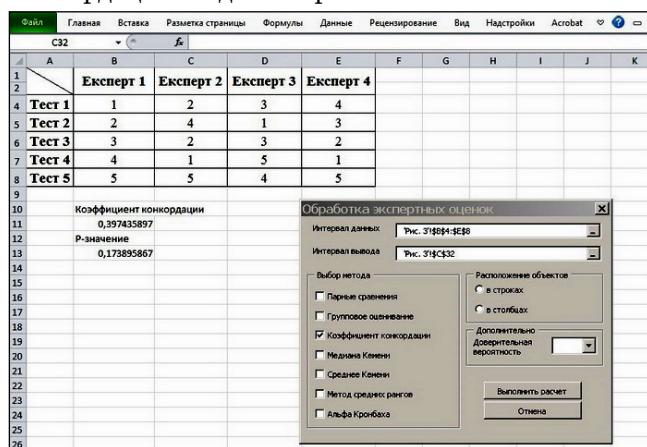


Рис. 2. Екранне вікно програми «AtteStat» з результатами обчислення коефіцієнта конкордації

Встановлено, що аналіз узгодженості думок експертів за допомогою обчислення коефіцієнта конкордації є досить неповним, оскільки цей коефіцієнт орієнтований на перевірку однотайності суджень [1, 2, 9, 10]. Якщо судження експертів концентруються навколо кількох точок, тобто експерти розділяються на кілька різних груп з однаковими поглядами всередині групи, то коефіцієнт конкордації не дозволяє це виявити.

Більш точну перевірку узгодженості за трьома методами експертного оцінювання можна провести, використовуючи методи рангової кореляції Кендала і Спірмена. Коефіцієнт рангової кореляції Спірмена обчислюється за формулою:

$$\rho = 1 - \frac{6}{n^3 - n} \sum_{i=1}^n (p_{i1} - p_{i2})^2, \quad (4)$$

де p_{i1}, p_{i2} — рангова оцінка першого і другого експерта з i -ою альтернативою,
 n — кількість порівнюваних об'єктів.

Коефіцієнт кореляції Спірмена змінюється від -1 до $+1$. Коефіцієнт дорівнює 1 , якщо думки експертів щодо ранжирування результатів виконання слухачем тестів повністю співпадають. Коефіцієнт дорівнює

-1 , якщо думки експертів повністю розбіжні, тобто ранжирування є протилежними. Якщо коефіцієнт кореляції дорівнює нулю, ранжирування вважаються лінійно незалежними.

У таблиці 3 наведено результати обчислення коефіцієнта рангової кореляції Спірмена за балами, що отримав слухач від експертів 1 і 2 за виконання 5 різних тестів.

Таблиця 3

Приклад розрахунку коефіцієнта кореляції Спірмена

Експерти	Тести					S
	Тест 1	Тест 2	Тест 3	Тест 4	Тест 5	
Експерт 1	1	2	3	4	5	
Експерт 2	3	4	2	1	5	
$p_{i1} - p_{i2}$	-2	-2	1	3	0	
$(p_{i1} - p_{i2})^2$	4	4	1	9	0	18

Підставивши дані з таблиці 3 у формулу (4), отримуємо коефіцієнт рангової кореляції $\rho = 0,1$. Оскільки коефіцієнт рангової кореляції Спірмена змінюється від -1 до $+1$, то таке значення $\rho = 0,1$ свідчить про низьку узгодженість думок експертів 1 і 2.

Таким способом обчислюють узгодженість думок будь-яких двох експертів з цієї групи.

На рис. 3 наведено вікно програми «OPINION» з результатами оцінювання коефіцієнтів рангової кореляції Спірмена під час оброблення думок експертів за даними з таблиці 3.

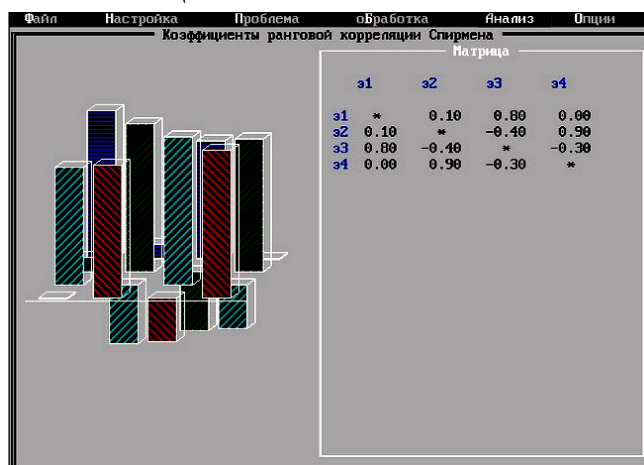


Рис. 3. Екранне вікно програми «OPINION» з результатами обчислення коефіцієнтів рангової кореляції Спірмена

У практиці педагогічного оцінювання часто трапляються випадки, коли експерти виставляють слухачам за виконання певних тестів однакові бали, і, відповідно, тоді їм треба приписувати однакові ранги. Кілька однакових балів поспіль утворюють зв'язку і називаються зв'язними. Кожен зі зв'язних балів отримує ранг, що дорівнює середньому арифметичному рангів, які мали б елементи зв'язки, якби вони були різні. Однакові ранги називаються зв'язними рангами й обчислюються за такою формулою [1, 2, 7, 9, 10, 11]:

$$\tilde{\rho} = \frac{\rho + T_1 + T_2}{\sqrt{(1 - T_1)(1 - T_2)}}, \quad (5)$$

де ρ — оцінка коефіцієнта рангової кореляції, а величини T_1 і T_2 обчислюються за формулами:

$$T_1 = \frac{3}{n^3 - n} \sum_i k_{1i} (k_{1i} - 1), \quad (6)$$

$$T_2 = \frac{3}{n^3 - n} \sum_i k_{2i} (k_{2i} - 1), \quad (7)$$

де k_{1i}, k_{2i} — кількість повторювальних рангів у i -й послідовності відповідних ранжирувань, отриманих від експертів 1 і 2.

Дані для обчислення скорегованого коефіцієнта рангової кореляції наведено у таблиці 4.

Таблиця 4

Результат обчислення скорегованого коефіцієнта рангової кореляції Спірмена для зв'язних рангів

Експерти	Тести					S
	Тест 1	Тест 2	Тест 3	Тест 4	Тест 5	
Експерт 1	2	2	3	4	5	
Експерт 2	3	4	3	3	5	
ρ_{r1-r2}	-1	-2	0	1	0	
$(\rho_{r1-r2})^2$	1	4	0	1	0	6

Підставивши дані з таблиці 4 у формули (4), (6), (7) і (5) відповідно, ми отримуємо такі результати: $\rho = 0,7$; $T_1 = 0,05$; $T_2 = 0,15$; $\hat{\rho} = 1$.

Значення рангового коефіцієнта кореляції Спірмена для зв'язних рангів дорівнює 1 у розглянутому прикладі. На підставі цього значення можна стверджувати, що існує щільна узгодженість у послідовності ранжирувань експертів 1 і 2.

Серед переваг рангового коефіцієнта кореляції Спірмена слід назвати те, що з його допомогою можна вивчати і вимірювати зв'язок не тільки між кількісними, але і між якісними (описовими) ознаками, які проранжировано певним способом.

Однак, за допомогою коефіцієнта рангової кореляції встановлюється щільність зв'язку між двома рядами, що ранжируються, тобто за цим значенням рангового коефіцієнта Спірмена можна зробити висновок про узгодженість думок лише будь-яких двох експертів у групі, а не групи в цілому.

Коефіцієнт рангової кореляції Кендалла за відсутності зв'язаних рангів визначається за формулою:

$$\tau = \frac{2}{n(n-1)} \sum_{i,j=1}^n \text{sign}[(r_{1i} - r_{1j})(r_{2i} - r_{2j})], \quad (8)$$

де n — кількість об'єктів, r_{ij} — ранги об'єктів, $\text{sign } x$ — функція, значення якої дорівнює

$$\text{sign } x = \begin{cases} 1 \text{ при } x > 0, \\ -1 \text{ при } x < 0, \\ 0 \text{ при } x = 0. \end{cases} \quad (9)$$

На рис. 4 наведено вікно програми «Opinion» з результатами оцінювання коефіцієнтів рангової кореляції Кендалла під час оброблення думок експертів за даними з таблиці 1.

Зазначимо, що метод обчислення коефіцієнта рангової кореляції, запропонований Спірменом, є більш простим і вимагає меншого часу обчислень. Слід зауважити також на те, що коефіцієнт кореляції Спірмена і коефіцієнт конкордації Кендалла ніяк не пов'язані між собою.

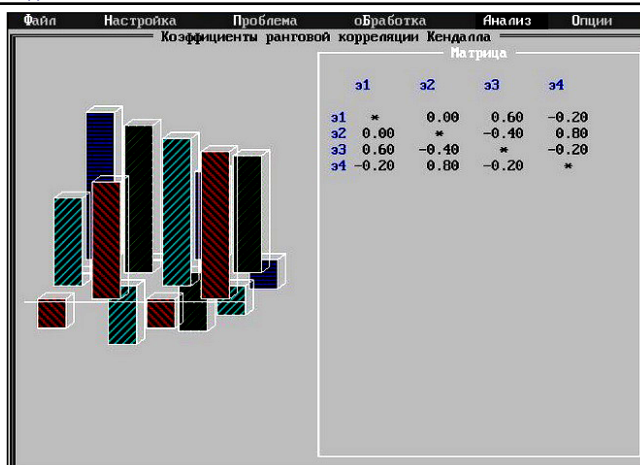


Рис. 4. Екранне вікно програми «OPINION» з результатами оцінювання коефіцієнтів рангової кореляції Кендалла

Отже, під час проведення тестування з іноземних мов врахування впливу згаданих чинників експертного оцінювання може забезпечити отримання надійніших результатів оцінювання рівня сформованості ІКК слухачів.

Висновки. Проведений аналіз обчислення показника узгодженості думки групи експертів за такими поширеними методами як коефіцієнт рангової кореляції Спірмена і коефіцієнт конкордації Кендалла демонструє, що під час експертного оцінювання рівня сформованості ІКК слухачів доцільно використовувати коефіцієнт конкордації Кендалла, оскільки він є єдиним, що відображає узгодженість думки групи експертів у цілому, тоді як коефіцієнт рангової кореляції Спірмена відображає узгодженість думки експертів лише попарно. Обраний математичний метод дозволяє більш точно визначити узгодженість думки групи експертів, а, отже, і більш точно оцінити рівень сформованості ІКК слухачів.

Незважаючи на успіхи, досягнуті в останні роки в розробленні і практичному використанні методу експертних оцінок, існує низка проблем і завдань, що вимагають подальших методологічних досліджень і практичної перевірки. Необхідно вдосконалювати систему відбору експертів, підвищувати надійність характеристик групової думки, розробляти методи перевірки обґрунтованості оцінок, досліджувати приховані причини, що знижують достовірність експертних оцінок. Однак, уже нині експертні оцінки в поєднанні з іншими математично-статистичними методами є важливим інструментом удосконалення управління процесом навчання іноземних мов на всіх рівнях середньої і вищої освіти.

* * *

Застело О. В. Анализ методов определения согласованности мнения группы экспертов при оценивании уровня сформированности иноязычной коммуникативной компетентности слушателей

Анотація. В статье рассматриваются современные математические методы определения согласованности мнения группы экспертов при тестировании слушателей по иностранным языкам.

Ключевые слова: экспертное оценивание, согласованность группы экспертов, коэффициент ранговой корреляции Спирмена, коэффициент конкордации Кендалла.

* * *

Zastelo O. V. The analysis of methods for determining the consistency of the expert group judgement when estimating the student's level of foreign language communicative competence

Annotation. The article deals with modern mathematical methods for determining the consistency of the expert group judgement when testing students in foreign languages.

Keywords: expert evaluation, the consistency of the expert group, Spearman's rank correlation coefficient, Kendall's concordance coefficient.

Література

1. Бешелев С. Д., Гурвич Ф. Г. Математико-статистические методы экспертных оценок. — М.: Статистика, 1980. — 263 с.
2. Ким В. С. Тестирование учебных достижений: монография / В. С. Ким. — Уссурийск, 2007. — 214 с.
3. Мухин О. И. Моделирование систем [Электронный ресурс] / О. И. Мухин. — Режим доступа: <http://stratum.ac.ru/textbooks/modelir/lection36.html>.
4. Новосад В. П., Селиверстов Р. Г. Методологія експертного оцінювання: Конспект лекцій для використання в навчальному процесі в системі підвищення кваліфікації кадрів / Уклад. В. П. Новосад, Р. Г. Селиверстов. — К.: Вид-во НАДУ, 2007. — 56 с.

5. Орлов А. И. Теория принятия решений: учебное пособие. — М.: Изд-во «Март», 2004. — 656 с.
6. Орлов А. И. Экспертные оценки: учеб. пос. / А. И. Орлов. — М.: Экзамен, 2002. — 31 с.
7. Павлов А. Н., Соколов Б. В. Методы обработки экспертной информации: учебно-метод. пособие / А. Н. Павлов, Б. В. Соколов. — ГУАП. СПб., 2005. — 42 с.
8. Субетто А. И. Квалитология образования / А. И. Субетто. — СПб.; М.: Исследовательский центр проблем качества подготовки специалистов, 2000. — 220 с.
9. Хабаров С. П. Экспертные системы [Электронный ресурс] / С. П. Хабаров. — Режим доступа: http://www.habarov.spb.ru/new_es/index.htm.
10. Чегодаев А. И. Математические методы анализа экспертных оценок / А. И. Чегодаев // Вестник Самарского государственного экономического университета. — 2010. — №2. — С. 130–135.
11. Черепанов В. С. Экспертные оценки в педагогических исследованиях / В. С. Черепанов. — М.: Педагогика, 1989. — 152 с.
12. Bonett D. G., Wright T. A. Sample size requirements for Pearson, Kendall, and Spearman correlations [online] / D. G. Bonett, T. A. Wright. — Springer International Publishing AG: Psychometrika, 2000. — Vol. 65, No.1. — Pp. 23–28. Available from: <http://link.springer.com/journal/11336/65/1>.
13. Durkin J. Expert Systems, Design and Development / J. Durkin. — USA, New York: Macmillan Publishing Company, 1994. — 800 p.

★ ★ ★

Вимоги до статей

Останнім часом до редакції надходить багато статей, оформлених за однаковою структурою. У статтях виділяються жирним шрифтом такі складові: Постановка проблеми, Аналіз останніх досягнень, Мета статті, Виклад основного матеріалу тощо. Дотримання авторами такої обов'язкової структури часто призводить до зниження її науковості й творчості й фактично до шаблонності.

Нині основними нормативними документами, у яких наводяться вимоги до наукових статей, є такі:

- Наказ МОН України від 17.10.2012 р. №1111 «Про затвердження Порядку формування Переліку наукових фахових видань України»;
- Постанова Президії ВАК України від 15.01.2003 р. №7–05/1 «Про підвищення вимог до фахових видань, внесених до переліків ВАК України»;
- ДСТУ ГОСТ 7.9:2009 (ИСО 214–76) «Система стандартів по інформації, бібліотечному и издательскому делу. Реферат и аннотация. Общие требования» (ГОСТ 7.9–95 (ИСО 214–76), IDT).

У зазначеній Постанові ВАК України говориться, що наукові статті повинні мати «...такі необхідні елементи: постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями; аналіз останніх досліджень і публікацій, у яких започатковано розв'язання даної проблеми і на які спирається автор, виділення невирішених раніше частин загальної проблеми, котрим присвячується означена стаття; формулювання цілей статті (постановка завдання); виклад основного матеріалу дослідження з повним обґрунтуванням отриманих наукових результатів; висновки з цього дослідження і перспективи подальших розробок у даному напрямку».

Отже, у постанові ВАК йдеться мова про наявність відповідних елементів, а не про структуру наукової статті. Це означає, що кожний автор має право самостійно визначати її структуру. Головне, щоб наукова стаття була творчою, мала наукову і практичну значущість.

Виходячи з вимог наведених документів, редакція журналу просить дотримуватися таких правил оформлення наукових статей.

- На початку статті у лівому верхньому куті ставиться індекс УДК. Далі наводяться назва статті прописними буквами напівжирним шрифтом, прізвище автора (авторів), ім'я, по батькові, посада, повна назва організації, науковий ступінь і наукове звання, e-mail, анотація (3–5 рядків) і ключові слова.
- Текст статті.
- Українською, російською і англійською мовами: назва статті, анотація і ключові слова; прізвище, ім'я, по батькові автора (авторів); посада, повна назва організації, науковий ступінь і наукове звання.
- Література (у порядку посилання на неї у тексті).

Стаття має бути набрана у текстовому редакторі (Word), шрифт Times New Roman, 12 pt, інтервал — 1,5. Параметри сторінки: верхнє і нижнє поле — 2 см, лівє — 2,5 см, правє — 1,5 см.

Рисунки, таблиці і фото розміщуються у тексті статті з обов'язковим посиланням на них. Крім того, якісне фото автора, кожний рисунок і екранні копії додаються в одному з форматів — tiff, png, jpg та інш. окремими файлами.