

ДИАГНОСТИКА НА ЕКОЛОГИЧНАТА КУЛТУРА НА УЧЕНИЦИТЕ ПРИ ИЗУЧАВАНЕ НА VI-A ГРУПА НА ПЕРИОДИЧНАТА ТАБЛИЦА В 8. КЛАС

Антоанета Ангелачева
Пловдивски университет „Паисий Хилендарски“

Резюме. В статията са представени резултатите от статистическото изследване на качествата на критериален тест „VI-A група на Периодичната таблица и опазване на околната среда“. Анализът на тестовите задачи е осъществен в следните аспекти: (а) априорен анализ – реализиран чрез метода на експертната оценка; (б) апостериорен анализ – чрез статистическа обработка на резултатите от апробирането на теста в представителна извадка от ученици в 8. клас. Емпиричните данни показват, че тестът притежава добри статистически характеристики и може да служи за ефективна диагностика на екологичната култура на учениците. Тестът е използван като основно диагностично средство в организираното експериментално обучение, което има за цел да провери целесъобразността на разработената дидактическа технология за обогатяване на екологичната култура на учениците при изучаване на VI-A група на Периодичната таблица в 8. клас.

Keywords: ecological culture, didactical tools, chemistry education, environmental education

Увод

В предишна публикация (Angelacheva & Gergova, 2013) са представени познавателните цели, структурата и съдържанието на критериален тест „VI-A група на Периодичната таблица и опазване на околната среда“, съставен върху общозадължителното учебно съдържание по „Химия и опазване на околната среда“ в 8. клас. За да може конструираният тест да служи за ефективна диагностика на екологичната култура на учениците, той трябва да притежава добри статистически характеристики.

Настоящата статия представя резултатите от статистическия анализ при изследване на качествата на разработения тест и при неговото приложение в процеса на обучение по химия в 8. клас.

Методология на изследването

Целта на статистическото изследване е да се получат достоверни данни за качествата на критериален тест „VI-A група на Периодичната система и опазване на околната среда“. Въз основа на получените резултати да се достигне до обосновано решение за използването на теста като надежден инструмент за диагностика на продуктивността на учебната дейност, свързана с екологичните знания и ценностните отношения на учениците към околната среда в 8. клас.

Дефинираната цел ориентира към решаването на следните задачи: (а) осъществяване на *априорен анализ* чрез метода на експертната оценка; (б) реализиране на *апостериорен анализ* чрез статистическа обработка на резултатите от изпробването на теста в представителна извадка от ученици в 8. клас; (в) използване на теста за диагностика на екологичната култура на учениците в специално организираното експериментално обучение.

Резултати и обсъждане

Априорният анализ е извършен от 10 експерти – учители по химия с висока педагогическа квалификация. Експертите оценяват качествата на теста по показатели, необходими за определяне на неговата съдържателна валидност, надеждността на предложената от авторите система за оценяване на резултатите от теста и стандарта за успешност на теста.

Показател за *съдържателната валидност* на теста е съответствието между всяка от задачите в него и познавателните цели, чието постигане той е предназначен да диагностицира (Бижков, 1996). Това съответствие се оценява от експертите по следната скала: +1 – задачата съответства на дадена цел/цели; 0 – не може да се прецени категорично дали задачата съответства на дадена учебна цел; –1 – задачата не отговаря на нито една цел (познавателните цели на теста са представени от Agelacheva & Gergova (2013)). Количествен израз на разглеждания показател е *коэффициентът на съответствие* CVR [формула (1), Бижков (1996)].

$$CVR = \frac{n_e - N/2}{N/2} \quad (1)$$

където n_e е броят на експертите, които оценяват положително задачата от теста, а N е общият брой на експертите.

След заместване на данните от експертните оценки във формула (1) получените числови стойности на коефициента CVR показват, че съдържателната валидност на отделните задачи и на теста като цяло е добра (таблица 1 в Приложение).

Надеждността на предложената от авторите система, чрез която се оценяват постиженията на учениците и се формира общият тестови бал, е определена от експертите по познатата скала: +1 – отразява съгласието на всеки от експертите с оценката на автора; 0 – колебание дали оценката на автора е коректна; –1 – несъгласие с оценката на автора. При отговор –1 е необходимо експертът да предложи оценка, която счита за коректна. Отговорите на експертите са отразени в таблица 2, Приложение.

Обсъжданият показател се представя количествено с *коефициента на съгласие* RA (percentage agreement), който показва какъв процент от експертите поддържат сходна оценка относно дадена категория спрямо общия брой оценки [формула (2); Стоянова (1996)]:

$$RA = \frac{\text{брой съгласия}}{\text{брой съгласия} + \text{брой несъгласия}} \quad (2)$$

където съгласието означава оценка (+1) на експерта, а несъгласието – оценка (0) или (–1).

Границите, в които варира коефициентът RA, са между 70% и 100%. След заместване на данните от експертните оценки във формула (2) получените числови стойности на RA (таблица 2, Приложение) дават основание да се твърди, че надеждността на предложената от авторите оценъчна система е добра.

Чрез експертна оценка е определен и *стандартът (критерият) за успешност* U, който дава отговор на въпроса: колко от задачите в субтеста и в теста като цяло трябва да бъдат решени, за да се счита, че са постигнати целите на субтеста (респективно на теста).

За да се прецени какъв трябва да бъде минималният брой ($N_{\min}$) вярно решени задачи от ученика при тест с дължина (N), експертите използват данни от таблицата на К. Клауер (Бижков, 1996), според която: (а) при $N = 6$ минимумът вярно решени задачи трябва да е $N_{\min} = 4$; (б) при $N = 12$ минимумът вярно решени задачи трябва да е $N_{\min} = 9$ (при вероятност за грешка $p = 0,95$ и равнище на достоверност $\alpha = 0,01$).

Експертите сравняват предложените от К. Клауер стандарти за успешност U с конкретния тест. Според 9 от общо 10 експерти, участвали в оценката на теста, за да се приеме, че неговите цели са постигнати, трябва да бъдат решени най-малко 9 от общо 12 задачи, т.е. $U(\text{тест}) = 9$; $U(\text{субтест 1}) = 4$; $U(\text{субтест 2}) = 4$.

Апостериорният анализ на теста е осъществен след неговото апробиране в извадка от 60 ученици в 8. клас. Получените данни и тяхната статистическа обработка са необходими за: определяне на *трудността* и на *дискриминативната сила* на задачите от теста; анализ на *дистракторите* (неверните отговори); установяване на *надеждността* и на *критериалната валидност* на теста.

Трудността на задачите от критериално ориентиран тест е пряка функция на целите, за измерването на които той е конструиран. Трудността на тестовите задачи се представя чрез *индекса на трудност* P , който показва процентната част на тестираните лица, които са постигнали поставената цел(и) при решаването на съответната задача, и се изчислява по формула (3) (Бижков, 1996):

$$P = \frac{N_R}{N} \cdot 100 \quad (3)$$

където N_R е броят на учениците от изследваната група, които вярно са решили задачата и така са постигнали поставената цел, а N е общият брой на учениците от групата.

Тест, конструиран от задачи с избран отговор (1 от 5), където вероятността за налучкване на верния отговор е малка, трябва да има средна трудност около 70% (Бижков, 1996). Следователно за тест, който се състои от 12 задачи, средният брой вярно решени задачи от ученик трябва да е 8 – 9 задачи.

След заместване на емпиричните данни във формула (3) получените стойности на индекса на трудност P (вж. таблица 3, Приложение) са в допустимите граници (около 70%), т.е. тестовите задачи са с добри показатели по отношение на изследваното качество.

Дискриминативната сила показва възможността дадена задача да разграничи учениците, които са постигнали учебната цел, от тези, които не са успели. Определя се чрез метода на контрастните групи („силна“ и „слаба“ група), образувани въз основа на стандарта за успешност на теста. „Силната“ група включва ученици, които са постигнали целите на теста; „слабата“ група – ученици, които са решили по-малко от 9 задачи. Дискриминативната сила се изчислява по формула (4) (Бижков, 1996):

$$BI = \frac{V_1}{N_1} - \frac{V_2}{N_2} \quad (4)$$

Където V_1 , V_2 – брой на учениците, съответно от „силната“ и „слабата“ група, с вярно решение на задачата, а N_1 , N_2 – общият брой ученици в тези групи ($N_1=N_2=60$).

Според Р. Ебел приемливият от практическа гледна точка индекс на дискриминативната сила е между 0,40 и 0,60 (Бижков, 1996). Данните в таблица 3 (Приложение) показват, че индексите на дискриминативната сила на тестовите задачи са над 0,40. Следователно задачите от конструирания тест позволяват да се различат учениците, които са постигнали целите на теста, от тези, които не са ги постигнали.

При задачи с избран отговор анализът на *дистракторите* е необходим, за да се установи дали и в каква степен те са приемливи за учениците и доколко позволяват да се разграничат „силните“ от „слабите“ ученици. Данните от тестирането

показват, че тестовите задачи отговарят на определените от Р. Брек основни изисквания при анализа на дистракторите (Бижков, 1996): (а) всеки от дистракторите е посочен от повече „слаби“ ученици, отколкото „силни“; (б) всеки дистрактор е посочен също така от поне няколко от „силните“ ученици; (в) в групата на „силните“ ученици нито един от дистракторите не е посочен от повече ученици, отколкото верния отговор.

Надеждността на критериалния тест се определя от трайността на резултатите от две последователни тестиране (тест Т / ретест РТ). За анализ на надеждността се използват алтернативни признаци, тъй като от експертната оценка на теста е изведен конкретен стандарт за успешност (9 вярно решени задачи). По него учениците могат да се разграничат на „постигнали“ и „непостигнали“ целите на теста при неговото повторение.

Количествено надеждността се представя с *коефициента на корелация* ϕ , който показва дали съществува зависимост между резултатите от двукратното тестиране и описва силата на тази зависимост. Той приема стойности от 0 до 1 и колкото по-високи са тези стойности, толкова по-надежден е конкретният тест. *Коефициентът* ϕ се изчислява по формула (5) (Бижков, 1996):

$$\phi = \sqrt{\frac{\chi^2}{n}} \quad (5)$$

Където n е броят на учениците в извадката, а χ^2 е критерий, който показва дали съществува статистически значима корелация между резултатите от първото и от второто тестиране. Този критерий се изчислява по формула (6) (Бижков, 1996):

$$\chi^2 = \frac{(ad - bc)^2 \cdot n}{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)} \quad (6)$$

Във формула (6) a и d отразяват броя на случаите, притежаващи или непритежаващи изследваните признаци (постигнали, респективно непостигнали целите при теста и при ретеста), b и c – броя на случаите, при които се притежава единият, но не се притежава другият признак (постигнали целите на теста, но непостигнали целите на ретеста или обратно).

За да се определи *коефициентът на корелация* ϕ , е необходимо резултатите от двукратното тестиране да се представят в таблица с четири полета (Таблица 1).

Таблица 1. Трайност на резултатите от теста

Тест (Т) – Ретест (РТ)	Постигнали целите	Непостигнали целите	Общо (ретест)
Постигнали целите	a = 48	b = 12	a + b = 60
Непостигнали целите	c = 6	d = 54	c + d = 60
Общо (тест)	a + c = 54	b + d = 66	n = 120
$\chi^2 = 59,39 > \chi^2_{0,01/1} = 6,64$ $\phi = 0,7$			

От данните в таблица 1 може да се твърди, че надеждността на съставения тест е приемлива, тъй като съществува статистически значима корелация между резултатите от теста и ретеста (изчислената стойност на коефициента χ^2 е по-голяма от критичната стойност $\chi^2_{0,01/1} = 6,64$), а стойността на коефициента ϕ е висока ($\phi = 0,7$).

Критериалната валидност на теста (степената на съответствие на теста и предварително определен критерий) е анализирана по отношение на: (а) *вътрешната критериална валидност* (сравняване на резултатите от двата субтеста и резултатите от целия тест); (б) *външната критериална валидност* (сравняване на резултатите от теста „VI-A група на Периодичната таблица и опазване на околната среда“ с резултатите от друг критериален тест, който измерва същата зависима променлива).

За определяне на критериалната валидност в откритите два аспекта са използвани *коефициентът на корелация ϕ* [формула (5)] и *вътрешно специфичният коефициент на асоциация $\phi_{\text{кор.}}$* (формула (7); Клаус & Ебнер, 1971), чрез който се оценява степента на зависимост между променливите величини.

$$\phi_{\text{кор.}} = \frac{bc - ad}{n \cdot \min(b, c) - (bc - ad)} \quad (7)$$

където n – брой на учениците в извадката; a, d – отразяват броя на случаите, които притежават или не притежават изследваните признаци (постигнали или непостигнали целите на теста); b, c – отразяват броя на случаите, които притежават единия признак, но не притежават другия.

Прилагането на коефициентите ϕ , $\phi_{\text{кор.}}$ и критерия χ^2 при изследване на *вътрешната критериална валидност* на теста изисква резултатите от субтест 1, от субтест 2 и от теста да се представят в таблици с четири полета (таблицы 2, и 3).

Таблица 2. Корелация между резултатите от субтест 1 и от теста като цяло

Резултати от субтест 1	Резултати от теста		Общо
	<i>Постигнали</i>	<i>Непостигнали</i>	
<i>Постигнали</i>	a = 47	b = 12	a + b = 59
<i>Непостигнали</i>	c = 7	d = 54	c + d = 61
Общо	a + c = 54	b + d = 66	n = 120
$\chi^2 = 56,34 > \chi^2_{0,01/1} = 6,64$ $\phi = 0,69$ $ \phi_{\text{кор.}} = 0,74$			

Таблица 3. Корелация между резултатите от субтест 2 и от теста като цяло

Резултати от субтест 2	Резултати от теста		Общо
	<i>Постигнали</i>	<i>Непостигнали</i>	
<i>Постигнали</i>	a = 49	b = 12	a + b = 61
<i>Непостигнали</i>	c = 6	d = 53	c + d = 59
Общо	a + c = 55	b + d = 65	n = 120
$\chi^2 = 59,46 > \chi^2_{0,01/1} = 6,64$ $\phi = 0,7$ $ \phi_{\text{кор.}} = 0,79$			

Статистическата обработка на резултатите, отразени в таблици 2 и 3, е необходима за проверка на статистическите хипотези:

H_0 : $p_1 = p_2$ не съществува зависимост между резултатите от двата субтеста и общия резултат от теста.

H_1 : $p_1 \neq p_2$ съществува зависимост между резултатите от двата субтеста и общия резултат от теста.

Приложените статистически процедури и получените от тях данни (таблици 2 и 3) показват, че стойностите на критерия χ^2 са по-високи от критичната стойност $\chi^2_{0,01/1} = 6,64$, а абсолютните стойности на коефициента $\phi_{\text{кор.}}$ са по-високи от 0,5. Следователно може да се приеме алтернативната хипотеза H_1 , т.е. вътрешната критериална валидност на теста е емпирично доказана.

За установяване на *външната критериална валидност* резултатите от теста „VI-A група на Периодичната система и опазване на околната среда“ (тест 1) се сравняват с резултатите от тест „V-A група на Периодичната таблица и опазване на околната среда“ (тест 2) (Ангелачева & Гергова, 2003). Използват се същите статистически величини и процедури както при оценката на вътрешната критериална валидност (таблица 4).

Таблица 4. Корелация между резултатите от тест 1 и от тест 2

Продуктивност на ученето (тест 1)	Продуктивност на ученето (тест 2)		Общо
	Постигнали	Непостигнали	
Постигнали	a = 45	b = 18	a + b = 63
Непостигнали	c = 8	d = 49	c + d = 57
Общо	a + c = 53	b + d = 67	n = 120
$\chi^2 = 39,97 > \chi^2_{0,01/1} = 6,64$ $\phi = 0,58$ $ \phi_{\text{кор.}} = 0,68$			

Данните от таблица 4 са необходими за проверка на статистическите хипотези:

H_0 : $p_1 = p_2$ не съществува зависимост между резултатите от двата теста в двете извадки.

H_1 : $p_1 \neq p_2$ съществува зависимост между резултатите от двата теста в двете извадки.

Тъй като емпиричните стойности на критерия χ^2 са по-високи от стойността на $\chi^2_{0,01/1} = 6,64$, а абсолютните стойности на $\phi_{\text{кор.}}$ са по-високи от 0,5, се приема алтернативната хипотеза H_1 , т.е. съществува значителна корелация между резултатите от тест 1 и тест 2. Следователно може да се твърди, че външната критериална валидност на теста е доказана.

Резултатите от априорния и от апостериорния анализ показват, че качествата на теста удовлетворяват описаните в литературата основни изисквания към критериалните тестове (Иванов, 1999; Наследов, 2012; Стойчев, 2007). Следователно той може ефективно да се използва за диагностика на екологичната култура на учениците при изучаване на VI-A група в 8. клас.

Създаденият тест е използван като основен диагностичен инструмент за измерване на резултатите от организираното експериментално обучение. То има за цел да провери целесъобразността на разработената дидактическа технология за обогатяване на екологичната култура на учениците при изучаване на посоченото учебно съдържание. Обучението е осъществено с две групи ученици: експериментална група (обучава се по създадения технологичен вариант, който ще бъде представен в следваща публикация) и контролна група (обучава се по общоприетата учебна програма). Групите са изравнени по признака успеваемост по химия чрез резултатите от входящ тест „Химични елементи от II-A група“. На случаен принцип от двете групи са избрани по 30 ученици с добър и 30 ученици с много добър успех, които са постигнали на входящия тест. След обучението в раздела за VI-A група е проведено изходящо тестиране с цел диагностика на екологичната култура на учениците (използван е критериалният тест „VI-A група на Перио-

дичната система и опазване на околната среда⁶). Резултатите от тестирането са представени в таблица 5.

Таблица 5. Резултати от теста в контролната група (КГ) и в експерименталната група (ЕГ)

Бал (брой решени задачи)	КГ (Субтест 1), честота	КГ (Субтест 2), честота	ЕГ (Субтест 1), честота	ЕГ (Субтест 2), честота
2	4	5	-	2
3	16	14	9	8
4	25	22	10	11
5	12	13	30	29
6	3	6	11	10

Непрекъснатият характер на белега (продуктивност на ученето по химия с акцент върху екологичната култура на учениците) мотивира избора на статистическите параметри: *средноаритметична $\bar{x}$, дисперсия s^2 , стандартно отклонение s* (Генкова & Найденова, 2003; Димитрова, 2006; Лакюрски, 1999) (таблица 6).

Таблица 6. Статистически величини, които характеризират продуктивността на ученето в изследваните групи

Характеристики	КГ, субтест 1	КГ, субтест 2	ЕГ, субтест 1	ЕГ, субтест 2
Брой изследвани лица (n)	60	60	60	60
Средноаритметична ($\bar{x}$)	3,9	4,02	4,72	4,62
Дисперсия (s^2)	0,94	1,2	0,88	1,05
Стандартно отклонение (s)	0,97	1,1	0,94	1,03

Съществен за изследването е проблемът: различават ли се разпределенията на случайните величини X и Y , характеризиращи продуктивността на ученето по химия, в експерименталната и в контролната група?

H_0 : Разликата между разпределенията на случайните величини X и Y в изследваните групи е несъществена и се дължи само на случайни фактори.

H_1 : Между разпределенията на случайните величини X и Y в изследваните групи съществува значима разлика, която изключва влиянието на случайни фактори.

Статистическите хипотези се изследват с непараметрични методи, които не са свързани с формата на разпределение на измерените величини (Иванов, 1999). Приложен е *еднофакторен дисперсионен анализ*, тъй като се проследява въздействието на един комплексен фактор (конструираната дидактическа технология за обогатяване на екологичната култура на учениците) върху резултативната променлива (продуктивността на ученето по химия с акцент върху екологичната култура на учениците). За да се установи дали този фактор предизвиква систематична вариация в разпределението на случайните величини X и Y , отразяващи продуктивността на учене в групите, обучавани по различните варианти, е необходимо да се провери наличието или отсъствието на разлика между средните параметри на резултативния признак в извадките. Ако разликата е статистически значима, то вариацията, предизвикана от съвкупното влияние на фактора, има системен, закономерен характер.

Прилагането на еднофакторен дисперсионен анализ изисква експерименталните данни да са получени от независими случайни извадки, принадлежащи към генерални съвкупности с еднаква дисперсия (Клаус & Ебнер, 1971). Дали е изпълнено това условие, може да се провери с *тест на Бартлет*. Статистическите хипотези са:

$H_0: s_1^2 = s_2^2$ няма разлика между дисперсиите в двете извадки;

$H_1: s_1^2 \neq s_2^2$ дисперсиите в двете извадки се различават.

При избрано равнище на достоверност $\alpha = 0,01$ експерименталните данни, необходими за теста на Бартлет, са представени в таблица 7.

Означения: $\bar{x}$ – средноаритметична стойност на извадката; $\sum_{i=1}^{60} x_{ij}^2$ – сума от квадратите на бала на всички членове на извадката; s_i – сума от x_{ij} в извадката; s_i^2 – дисперсия на групата.

Таблица 7. Данни за теста на Бартлет

Характеристики	$\bar{x}$	$\sum_{i=1}^{60} x_{ij}^2$	S_i	$S_i^2 / 60$	s_i^2	$\lg s_i^2$
КГ субтест 1	3,9	968	234	912,6	0,94	- 0,0269
ЕГ субтест 1	4,72	1387	283	1334,82	0,88	- 0,0555
сума (Σ)	-	2355	517	2247,42	-	- 0,0824
КГ субтест 2	4,02	1039	241	968,02	1,2	0,0792
ЕГ субтест 2	4,62	1341	277	1278,82	1,05	0,0212
сума (Σ)	-	2380	518	2246,84	-	0,1004

Дисперсиите s_i^2 и претеглената средна величина от дисперсиите s^2 се изчисляват по формули (8) и (9) (Клаус & Ебнер, 1971):

$$s_i^2 = \frac{1}{n_i - 1} \left(\sum_{j=1}^{60} x_{ij}^2 - \frac{S_i^2}{n_i} \right) \quad (8)$$

където $n_i = 60$ е обемът на извадката.

$$s^2 = \frac{1}{N - k} \sum_{i=1}^k \left(\sum_{j=1}^{60} x_{ij}^2 - \frac{S_i^2}{n_i} \right) \quad (9)$$

където N е обемът на извадките, а k е броят на извадките.

След заместване на данните във формула (9) се получава: за субтест 1 $s^2 = 0,9117 \Rightarrow \lg s^2 = -0,0402$; за субтест 2 $s^2 = 1,1285 \Rightarrow \lg s^2 = 0,0525$.

В теста на Бартлет проверяващата величина е *критерият* χ^2 , който се изчислява по формула (10) (Клаус & Ебнер, 1971):

$$\chi^2 = \frac{2,3026}{c} k (n-1) \left(\lg s^2 - \frac{1}{k} \sum_{i=1}^k \lg s_i^2 \right) \quad (10)$$

където $c = 1 + \frac{k+1}{3k(n-1)} = 1 + \frac{2+1}{3 \cdot 2(60-1)} = 1,0085$.

След заместване във формула (10) се получава: за субтест 1 $\chi^2 = 0,27$; за субтест 2 $\chi^2 = 0,62$.

Получените стойности на χ^2 се сравняват с критичната стойност $\chi_{0,01/1}^2 = 6,64$.

Тъй като и за двата субтеста $\chi^2 \ll \chi_{0,01/1}^2$, може да се приеме нулевата хипотеза H_0 , т.е. няма разлика между дисперсиите на контролната и на експерименталната група.

След като с теста на Бартлет е доказано условието за еднаква дисперсия в двете групи, може да се приложи *еднофакторен дисперсионен анализ* при независими извадки с равен обем ($n_1 = n_2 = 60$) и избрано равнище на достоверност $\alpha = 0,01$ (Клаус & Ебнер, 1971). Статистическите хипотези са:

H_0 : $\mu_1 = \mu_2$ няма разлика между средноаритметичните величини в двете извадки;

H_1 : $\mu_1 \neq \mu_2$ средноаритметичните величини в двете извадки се различават.

Формулите и изчисленията за дисперсионния анализ са представени в таблица 8.

Таблица 8. Таблица на дисперсионния анализ

Вариация	Обща	Между групите	По групи
Сума от квадратите на отклоненията (СКО): субтест 1 субтест 2	$\sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - \frac{S^2}{N}$ 2355 - 2227,41 = 127,59 2380 - 2236,03 = 143,97	$\frac{1}{n} \sum_{i=1}^k S_i^2 - \frac{S^2}{N}$ 2247,22 - 2227,41 = 20,01 2246,84 - 2236,03 = 10,81	$\sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - \frac{1}{n} \sum_{i=1}^k S_i^2$ 2355 - 2247,42 = 107,58 2380 - 2246,84 = 133,16
Брой на степените на свобода	$N - 1 = 120 - 1 = 119$	$k - 1 = 2 - 1 = 1$	$N - k = 120 - 2 = 118$
Среден квадрат (СК): субтест 1 субтест 2	-	$CK_M = \frac{CKO_M}{k - 1}$ 20,1 10,81	$CK_{гр.} = \frac{CKO_{гр.}}{N - k}$ 0,91 1,13

Проверяващата величина F се изчислява по формула (11) (Клаус & Ебнер, 1971):

$$F = \frac{CK_M}{CK_{гр.}} \quad (11)$$

където CK_M е среден квадрат между групите, а $CK_{гр.}$ е среден квадрат по групи.

След заместване на данните във формула (11) се получава: за субтест 1 $F = 21,99$; за субтест 2 $F = 9,57$.

Получените стойности на F се сравняват с критичната стойност $F_{0,01; 1/118} = 6,84$.

Тъй като и за двата субтеста $F > F_{0,01; 1/118}$, нулевата хипотеза H_0 се отхвърля. Следователно може да се приеме алтернативната хипотеза H_1 : средните параметри на на двете изследвани групи се различават.

С отхвърлянето на нулевата хипотеза възниква въпросът за оценка на значимостта на разликата между средните величини на сравняваните извадки. Използва се *тестът на Данкън*, при който статистически критерий е *значимият вариационен интервал* R_p (формула (12); Клаус & Ебнер (1971):

$$R_p = q_\alpha (p; N - k) \cdot S_x \quad (12)$$

където S_x е стандартното отклонение и се изчислява по формула (13); Клаус & Ебнер, (1971):

$$S_x = \sqrt{\frac{CK_{гр.}}{n}} \quad (13)$$

Разликата между средните величини на двете извадки се сравнява с вариационния интервал R_p . При това: (а) ако $\bar{X}_1 - \bar{X}_2 < R_p$, разликата между средните величини не е значима; (б) ако $\bar{X}_1 - \bar{X}_2 > R_p$, разликата между средните величини е значима.

В нашия случай разликата между средните величини е: за субтест 1 $\bar{X}_1 - \bar{X}_2 = 0,82$; за субтест 2 $\bar{X}_1 - \bar{X}_2 = 0,6$.

При $p = 2$ стойността на $q_{0,01}(2; 118)$ е 3,702 (Клаус & Ебнер, 1971).

След заместване на данните във формули (12) и (13) се получава: за субтест 1 $s_x = 0,1232$; $R_p = 0,4561 < 0,82$; за субтест 2 $s_x = 0,1372$; $R_p = 0,5080 < 0,6$.

Тъй като за двата субтеста разликата между средните величини $\bar{X}_1 - \bar{X}_2 > R_p$, нулевата хипотеза H_0 се отхвърля и се приема алтернативната хипотеза H_1 : разликата между средноаритметичните величини на двете групи е значима, т.е. резултатите на експерименталната група превъзхождат тези на контролната група. Следователно тестът не само доказва по-добрия ефект на разработената дидактическа технология в експерименталната група, но той може да служи и за диагностика на продуктивността на ученето по химия с акцент върху екологичната култура на учениците.

Таблица 9. Корелация между резултатите от субтест 1 и от субтест 2 за контролната група

Субтест 1	Субтест 2		Общо
	Постигнали	Непостигнали	
Постигнали	a = 9	b = 2	a + b = 11
Непостигнали	c = 14	d = 35	c + d = 49
Общо	a + c = 23	b + d = 37	n = 60

Таблица 10. Корелация между резултатите от субтест 1 и от субтест 2 за експерименталната група

Субтест 1	Субтест 2		Общо
	Постигнали	Непостигнали	
Постигнали	a = 38	b = 3	a + b = 41
Непостигнали	c = 9	d = 10	c + d = 19
Общо	a + c = 47	b + d = 13	n = 60

Интересен за нас е проблемът за трудността на задачите от двата субтеста съответно за експерименталната и за контролната група. Решението на този проблем може да открие по-силно разликата между двата варианта на обучение.

Критерият на Макнамара е разработен специално за сравняване на резултатите на две зависими извадки, с две категории на измерване на изучаваното свойство, при което двойките на случайните величини X и Y са взаимно независими (Грабарь & Краснянская, 1977; Клаус & Ебнер, 1971).

За да се използва този критерий, е необходимо данните за изследваните групи да се представят във вид на четириклетъчни таблици „2x2“ за две серии наблюдения на променливите X и Y , характеризиращи продуктивността на учене по химия от двата субтеста (таблици 9 и 10).

Статистическите хипотези имат вида:

H_0 : $p(x_i = 0, y_i = 1) = p(x_i = 1, y_i = 0)$ субтест 1 и субтест 2 са с еднаква трудност;

H_1 : $p(x_i = 0, y_i = 1) \neq p(x_i = 1, y_i = 0)$ субтест 1 и субтест 2 не са с еднаква трудност.

При $n = b + c < 20$ като проверяваща величина се използва статистиката на критерия T_2 . Тя се изчислява по формула (14) (Грабарь & Краснянская, 1977): $T_2 = \min(b, c)$.

От таблица за вероятностите p за биномиално разпределение при $p = 0,5$ се намират стойностите на p за експерименталната и за контролната група (Грабарь & Краснянская, 1977): при $\alpha = 0,01$, n (КГ) = 16, T_2 (КГ) = 2 и p (КГ) = 0,002; при $\alpha = 0,01$, n (ЕГ) = 12, T_2 (ЕГ) = 3 и p (ЕГ) = 0,073.

Табличните стойности на p се сравняват със стойността на $\alpha/2$. При това: (а) ако $p < \alpha/2$, нулевата хипотеза H_0 се отхвърля; (б) ако $p > \alpha/2$, нулевата хипотеза H_0 се приема.

За контролната група $p = 0,002 < \alpha/2$. Следователно H_0 се отхвърля и се приема алтернативната хипотеза H_1 за разлика в трудността на задачите от двата субтеста.

За експерименталната група $p = 0,073 > \alpha/2$. Следователно се приема нулевата хипотеза H_0 , т.е. за тази група ученици задачите от двата субтеста се възприемат като задачи с еднаква трудност.

Резултатите от тестирането показват, че в експерименталната група разработената технология създава добри условия за продуктивно учене по критериите приложимост на екологичните знания в различни ситуации и проявено отношение към екологичните проблеми в рамките на учебния материал за VI-A група на Периодичната таблица.

Заклучение

В заключение може да се отбележи: (а) качествата на тестовите задачи и на теста като цяло отговарят на описаните в литературата основни изисквания към

критериалните тестове; (б) конструираният вариант на критериален тест може ефективно да изпълнява ролята на инструментариум за диагностика на екологичната култура на учениците (на системата от екологични знания и ценностни отношения към околната среда) при изучаване на VI-A група на Периодичната таблица в 8. клас; (в) подобни диагностични средства могат да бъдат създадени в рамките и на други теми от учебното съдържание по химия, тъй като целенасоченото формиране и диагностика на екологична култура на учениците, която е неотменима част от общата култура на личността, следва да бъде специална задача на обучението по химия в училище.

ПРИЛОЖЕНИЕ

Таблица 1. Оценка на експертите за съответствието между тестовите задачи и познавателните цели, които тестът се стреми да реализира

Тестови задачи	1	2	3	4	5	6	7	8	9	10	11	12
Цели на теста*	1а	1б	1в	1г	1д	1е	2	3	4	5	6	7
Оценка на експертите (е)	+1/ 10е	+1/ 10е	+1/ 10е	+1/ 9е	+1/ 10е	+1/ 10е	+1/ 8е	+1/ 10е	+1/ 9е	+1/ 10е	+1/ 10е	+1/ 10е
CVR	1	1	1	0,8	1	1	0,6	1	0,8	1	1	1
Съдържателна валидност на теста като цяло: $CVR = 0,93$ $-1 \leq CVR \leq +1$												

* Целите на теста са представени в: Angelacheva & Gergova, 2013

Таблица 2. Оценка на експертите за авторската оценка на отговорите на тестовите задачи

Тестови задачи	1	2	3	4	5	6	7	8	9	10	11	12
Оценка на експертите (е)	+1/ 10е	+1/ 10е	+1/ 9е	+1/ 10е	+1/ 9е	+1/ 10е	+1/ 10е	+1/ 10е	+1/ 9е	+1/ 10е	+1/ 10е	+1/ 10е
PA, %	100	100	90	100	90	100	100	100	90	100	100	100

Таблица 3. Трудност и дискриминативна сила на задачите от теста

Тестови задачи	1	2	3	4	5	6	7	8	9	10	11	12
Трудност P, %	69	70	67	68	71	63	65	66	64	70	68	69
Индекс на дискриминативната сила VI	0,6	0,6	0,5	0,6	0,6	0,6	0,5	0,7	0,6	0,5	0,6	0,5

ЛИТЕРАТУРА

- Ангелачева, А. & Гергова, Е. (2003). Критериален тест „V-A група на Периодичната система и опазване на околната среда“. *Научни трудове Пловдивски университет „Паисий Хилендарски“*, 40(2), 81 – 90.
- Бижков, Г. (1996). *Теория и методика на дидактическите тестове*. София: Просвета.
- Генкова, Л. & Найденова, В. (2003). *Опитно-приложната и диагностично-изследователската дейност на учителя*. София: Дъга – ВК.
- Грабарь, М. & Краснянская, К. (1977). *Применение математической статистики в педагогических исследованиях*. Москва: Педагогика
- Димитрова, В. (2006). *Диагностика в университетското обучение по химия*. Благоевград: Унив. изд. „Неофит Рилски“.
- Иванов, И. (1999). *Основи на педагогическата диагностика*. София: Аксиос.
- Клаус, Г. & Ебнер, Х. (1971). *Основи на статистиката за психолози, педагози и социолози*. София: Наука и изкуство.
- Лакюрски, А. (1999). *Математико-статистически методи в психолого-педагогическите изследвания*. София: Софттрейд.
- Наследов, А. (2012). *Математические методы психологического исследования: анализ и интерпретация данных*. Санкт-Петербург: Речь.
- Стойчев, Д. (2007). *Количествени методи за моделиране и диагностика в експерименталната педагогика, психология и социология*. Пловдив: Унив. изд. „Паисий Хилендарски“.
- Стойанова, Ф. (1996). *Тестология за учители*. София: Атика.
- Angelacheva, A. & Gergova, E. (2013). Didactical tools for developing an environmental culture of pupils in chemistry in the secondary school. *Chemistry*, 22, 532 – 543 [In Bulgarian].

DIAGNOSTICS OF THE ENVIRONMENTAL CULTURE OF PUPILS: A CASE STUDY – TEACHING OF THE VI-A GROUP OF THE PERIODIC TABLE IN 8TH GRADE

Abstract. The educational research is realized on the basis of a criterion test “VI-A Group of the Periodic Table”. The item-analysis is carried out following the methods: (a) expert evaluation of content validity and standard of successfulness of the test; (b) analysis the results of the test trial in a representative sample of pupils in the 8th grade. The empirical data show that the test has good statistical characteristics and can be used for effective diagnostics of the environmental culture of students.

✉ **Dr. Antoaneta Angelacheva**
Plovdiv University „Paisii Hilendarski”
24, Tsar Assen Str.,
4000 Plovdiv, Bulgaria
E-mail: angel@uniplovdiv.bg