

Нормативният дидактически тест в образованието

Николинка Александрова Атанасова

Normative didactic Test in Education

Nikolinka Aleksandrova Atanasova

Резюме: Нарастващата популярност на теста в България предизвика възникването на редица въпроси по съставянето му, а също и появата на множество тестове със съмнителни качества. По тази причина настоящата публикация цели да разкрие технология за съставяне на нормативен дидактически тест, съдържащ задачи с множествен избор на отговора. Тя е насочена към всички специалисти в образователната сфера, които се интересуват от тази проблематика.

Abstract: The increasing popularity of the test in Bulgaria caused the emergence of a number of issues in its composition, and the emergence of many tests of dubious quality. Therefore, this publication aims to bring technology Quality of normative teaching test tasks containing multiple choice answers. It is aimed at all professionals in education who are interested in this issue.

Ключови думи: тест, обективна трудност на задачата, дискриминативна сила, ефективност на дистракторите, надеждност, валидност, обективност, приложимост

Keywords: test, objective difficulty of the task, discriminative power, efficiency of distractor, reliability, validity, objectivity, utility

В България тестирането в образованието заема все по-значимо място като метод за диагностика. Фактът, че външното оценяване при завършване на определен етап на образование се осъществява чрез дидактически тестове е достатъчно доказателство за актуалността на този проблем. Нарастващата популярност на теста у нас предизвика възникването на редица въпроси по съставянето му, а също и появата на множество тестове със съмнителни качества, които свободно могат да се използват при подготовката на учениците за предстоящо изпитване и оценяване. По тази причина цел на настоящата публикация е разкриването на технология за качествено съставяне на нормативен дидактически тест, съдържащ задачи с множествен избор на отговора.

Първоначално тестирането възниква в областта на психологията. То се свързва с имената на Френсис Галтон, Чарлз Спирмън, Алфред Бине, Теодор Симон и др.

Счита се, че в началото на XX век американският изследовател В. А. Маккол за първи път разделя тестовете на психологически и образователни, а Едуард Торндайк създава първия тест за измерване в областта на обучението и образованието.

В България за първи път се обръща внимание на този проблем през 1932 г., когато излиза първото системно издание по въпросите на тестирането с автори Г. Гиздов, Л. Краев, Ал. Цанев.

Днес в световен мащаб съществуват различни организации, които се занимават с проучване възможностите на дидактическото тестиране, а в България съществува национален тестов център. (Радев, Пл., Ал. Александрова, 2003)

В Германия, Австрия и Швейцария са възприети общо седем групи педагогически тестове. Кои са те?

☞ Училищни тестове за постижения (комплексни, междупредметни).

☞ Комбинирани тестове за училищни постижения и тестове за пригодност (тук се отнасят: тестове за завършване на VIII и следващите класове, тестове за прием в професионални училище и някои скъли за оценка на поведението).

☞ Тестове за развитие и за училищна зрелост (тестове за деца от предучилищна възраст, тестове за училищна зрелост).

☞ Тестове за интелигентност и пригодност (индивидуални и групови, вербални и невербални, общи и специални).

☞ Тестове за концентрация и внимание (диференциални тестове за различни възрастови групи).

☞ Социални тестове (тестове за диагностика на социалните отношения, социометрични тестове, диагностична социограма и др.).

☞ Анкети (за диагностика на отношението към различни области за деца и възрастни, за училищен страх, интереси и др.).

В САЩ тестовете биват:

- тестове върху отделни теми по английски език, изкуства, чужд език, математика, религия, медицина, социални науки;

- тестове по четене за начална и средна училищна възраст;

- тестове по специална педагогика;

- тестове за диагностика на речта, слуха и зрението;

- тестове за оценка на отношението на учителите;

- тестове за професионалните интереси и способности;

- тестове по безопасност на движението и др.

Г. Бижков в критичен план счита, че тези класификации на тестовете не са “чисти”, което означава, че е налице смесица от педагогически и психологически тестове (Бижков, Г., В. Краевски, 2007).

Какво представляват тестовете, ориентирани към количествена норма (нормативни), които са само една от множеството разновидности на теста?

Нормата представлява емпирично изведен показател от резултатите на представителна извадка (група изследвани лица), който е получен на базата на седни величини (средна аритметична, медиана и т. н.) и дисперсии. Количествените норми могат да бъдат: универсални, национални, регионални и локални (вътрешногрупови). Тук се прилагат конкретни математико-статистически процедури.

Дидактическото тестиране е предизвикателство за педагозите, тъй като е свързано със спазване на редица етапи и изисквания. Технологиията за съставянето на нормативен дидактически тест, предложена в настоящата публикация, е реализирана при спазването на следната поетапност:

1. Определяне целите на тестирането – тук се дава отговор на въпроса какво ще се измерва.

2. Съставяне тест-спецификация – отнася се до планиране на теста, като се определят видът и броят на задачите, учебното съдържание.

3. Разработване на тестови задачи – за стандартизираните тестове голяма част от задачите отпадат поради ниско качество и затова обикновено се препоръчва съставянето на повече задачи, като правилото е 3 : 2, т.е. ако в окончателния вариант тестът ще съдържа 100 задачи, то на този етап трябва да се съставят поне 150.

4. Аprobация на теста – отнася се до изпробването на теста в реални условия, за да се види как той “работи” и да се съберат данни за пристъпване към следващия етап. За професионално съставени тестове обикновено извадката обхваща от 300 до 500 души. За учителски тест извадката обхваща между 30 и 100 ученика.

5. Анализ на качествата на отделните задачи – тук се изследват характеристики на задачите като: трудност, дискриминативната сила, ефективността на погрешните отговори (дистракторите) при тестовете с множествен избор на отговори.

На базата на този анализ се съставя окончателният вариант на теста.

6. Втора апробация (ако се налага).

7. Анализ на качествата на теста като цяло – тук се определят надеждността на теста с грешката на измерването и неговата валидност.

8. Определяне на стандартите за оценка – на този етап се определят нормите или критериите за оценка на резултатите от тестирането.

9. Подготовка на теста за масово приложение – отнася се за тестовете, които са предназначени за масова употреба. По-конкретно подготовката се състои в следното:

- описание на качествата на теста и процедурите, които са спазвани при създаването му;

- инструкция за провеждане на теста;

- инструкция за анализ и оценка на резултатите от теста;

- тестов материал, листове за отговори, ключове за оценка, таблици и др. помощни материали.

Този етап се пропуска, ако тестът е предназначен за ограничена употреба (в 1 - 2 класа). (Стоянова, Ф., 1996)

Как се реализира на практика тази поетапност?

Съставяне на стандартизиран, нормативен, тясноориентиран (по един учебен предмет), събирателен дидактически тест със задачи с множествен избор на отговора по Български език за IV клас

✳ Определяне целите на тестирането

В случая целта е измерване постиженията на учениците по български език в края на учебната година. Като “.....под понятието *учебни постижения* разбираме конкретните учебни резултати по определен предмет или пък свързаните с тях умения и подходи от по-общ характер.” (Бижков, Г, 1999).

Дидактическият тест за установяване на резултатите от обучението по български език е съставен на базата на когнитивната таксономия на Б. Блум. Шестте познавателни равнища на таксономията позволяват детерминирането на критериите за диагностика на постиженията на учениците. Към всеки критерий са диференцирани по три показателя, като при формулирането им са използвани подходящи активни глаголи. Съдържанието на показателите е конструирано и на базата на учебни цели на ниво теми и раздели от учебната програма по български език за IV клас. По-конкретно за теста са използвани по шест учебни цели, описани в бланките за експертна оценка при определяне съдържателната валидност на тестовете. Тук всяка от тях се подразделя на три показателя, които са включени в три различни равнища по таксономията. Така е оформена следната система от критерии и показатели за диагностика:

Тест по български език – изходно ниво

Знания

1. Способност за възпроизвеждане на факти

1.1. умение за разграничаване на бъдеще време от сегашно, минало свършено и минало несвършено време на глагола;

1.2. умение за определяне начина на писане на числителното име в мъжки род, единствено число, когато е подлог в изречението;

1.3. умение за възпроизвеждане на въпроса, с който се открива сказуемото като главна част в простото изречение.

Разбиране

2. Способност за интерпретация на информация

- 2.1. умение за илюстриране с пример на правилната употреба на наречие за време;
- 2.2. умение за обясняване значението на антонимите за изменение смисъла на даден текст;
- 2.3. умение за обясняване причината за необходимостта от редактирането на даден текст, като се проследява правописът на съществителните и числителните имена в него.

Приложение

3. Способност за прилагане на знания и умения в конкретни ситуации

- 3.1. умение за проверяване количеството на глаголите от минало свършено време в даден текст;
- 3.2. умение за установяване на подходящото наречие за място по даден звуков модел;
- 3.3. умение за прилагане на подходящ начин за свързване на прости изречения в сложно.

Анализ

4. Способност за разграничаване на елементи от едно цяло

- 4.1. умение за разпознаване на глаголи във второ лице, множествено число, минало несвършено време;
- 4.2. умение за откриване синоним на дадена дума сред думите в няколко изречения;
- 4.3. умение за разграничаване на подлога, съставен от съществително име, от подлюзите, съставени от прилагателно име или местоимение, в няколко прости изречения.

Синтез

5. Способност за свързване на елементи в едно цяло

- 5.1. умение за свързване на синоними по двойки;
- 5.2. умение за съставяне на сложно изречение чрез комбинирание на определен набор от думи;
- 5.3. умение за свързване на числителните имена с техния правопис.

Оценка

6. Способност за изразяване на мнение

- 6.1. умение за заключаване относно възможностите за запазване смисъла на изречение след степенуване на наречието като част на речта;
- 6.2. умение за преценяване на доминиращия вид изречения по състав в конкретна комуникативна ситуация;
- 6.3. умение за доказване значението на наречието за преобразуване на различните изречения по състав от един вид в друг.

За апробацията на теста са отчитани броят на вярно решените задачи по всички показатели от шестте равнища.

✳ Съставяне на таблици на спецификациите и разработване на тестови задачи

На този етап първоначално е съставена таблица на спецификациите, чрез която са определени: учебното съдържание, видът и броят на тестови задачи и дължината на теста като цяло.

Таблица 1

Когнитивни равнища Учебно съдържание по теми	Знания	Разбиране	Приложение	Анализ	Синтез	Оценка	Общ брой задачи
	брой задачи	брой задачи	брой задачи	брой задачи	брой задачи	брой задачи	
№ 20, 21	1		1	1			3
№ 25, 26, 27, 30		1	1			1	3
№ 10, 11, 12, 15, 18		1		1	1		3
№ 7, 8, 9			1		1	1	3
№ 5, 6, 7, 22, 25	1			1		1	3
№ 22, 23, 24	1	1			1		3
Брой задачи по равнища	3	3	3	3	3	3	18

Във връзка с този етап са взети предвид следните основни факти: за да има задоволителна надеждност, един дидактически тест трябва да съдържа повече от десет задачи (Стоянова, Ф., 1996), поради което тестът е с 18 задачи; "... винаги, когато целите на теста и съответното учебно съдържание позволяват, да се използват задачи с избран отговор, а не с конструируем, тъй като за тяхното решаване е необходимо по-малко време." (Бижков, Г., В. Краевски, 2007). Поради тази причина всички задачи са множествен избор на отговорите като всяка задача има само по един верен отговор от три възможни.

При съставянето на задачите са взети предвид следните правила:

- Въпросите и задачите да са ясно написани.
- Отговорите на въпросите и задачите да изглеждат еднакво приемливи, но само един да е верен.
- Тестовите въпроси и задачи да са независими едни от други и отговорът на предната да не подпомага отговора на следващата.
- Да се избягва употребата на негативни фрази.
- Възможните избори да се подредят в произволен числов или азбучен ред.
- Алтернативните отговори да не са синоними (Радев, Пл., Ал. Александрова, 2003).

✳ Аprobация на теста

На този етап целта е да се установи как на практика "работи" тестът и да се получи информация за качествата на тестовите задачи и на теста като цяло, за да се завърши окончателно и да може да се използва като инструмент за диагностика на постиженията на учениците. Имайки предвид факта, че за аprobацията на дидактически тестове е необходима представителна извадка, включваща от 300 до 1500 лица (Бижков, Г., В. Краевски, 2007), в конкретния случай извадката включва 305 ученика от четвърти клас в градовете Пловдив и Пазарджик. Аprobацията е осъществена за една учебна година. В този период е спазена следната поэтапност на изпробването:

1. Проведен е предварителен разговор с класните ръководители на четвъртите класове. Тук е уточнено времето за решаване на теста (за един учебен час - в края на учебната година), целите на аprobацията, учебното съдържание, денят и часът на провеждане на тестирането.

2. Тестиране. То е осъществено под ръководството на автора на теста и в присъствието на съответния класен ръководител. Във всеки клас началото се поставя с представяне на изследователя, който съобщава целта на изпробването, а именно за създаване на тест, който ще измерва реално и справедливо постиженията на учениците. Обръща се внимание на учениците, че те трябва да решават тестовите задачи внимателно, спокойно и съсредоточено, без да се притесняват за оценките, тъй

като в този случай е важен броят вярно решени задачи, което няма да повлияе върху годишната им оценка по български език.

3. Следва съобщаване на инструкцията, с която учениците трябва да се съобразяват при решаване на задачите. Сред правилата са следните:

- могат да се задават въпроси по процедурата на провеждане на тестирането;
- по време на тестирането да се работи напълно самостоятелно, без да се разговаря или подсказва;
- задачите могат да се решават последователно или в друг ред, избран от всеки ученик;
- към всяка задача има само по един верен отговор;
- от тестовите бланки само се четат задачите, за да се прецени кой е верният отговор, но не се пише върху тях;
- отговорите се отбелязват, като се оградят с кръгче в съответната бланка, с която разполага всеки ученик.

Таблица 2
БЛАНКА ЗА НАНАСЯНЕ НА ОТГОВОРИТЕ ОТ УЧЕНИЦИТЕ

УЧИЛИЩЕ	ГРАД.....
ИМЕ, ФАМИЛИЯ НА УЧЕНИКА	КЛАС
ПОЛ: М Ж	
език, на който най-често се говори в семейството:	Б Т Р Д български – турски – ромски – друг

Оградете с кръгче верния отговор!

Задача №	Отговори			Задача №	Отговори		
1.	а	б	в	10.	а	б	в
2.	а	б	в	11.	а	б	в
3.	а	б	в	12.	а	б	в
4.	а	б	в	13.	а	б	в
5.	а	б	в	14.	а	б	в
6.	а	б	в	15.	а	б	в
7.	а	б	в	16.	а	б	в
8.	а	б	в	17.	а	б	в
9.	а	б	в	18.	а	б	в

4. Пристъпва се към решаване на теста от учениците, като в този момент се използват прякото и описателното наблюдение като помощен изследователски метод спрямо дидактическото тестиране. Наблюдението се реализира, за да се гарантира обективното спазване на диагностичните процедури.

✳ Анализ на качествата на отделните задачи

На този етап от подготовката на дидактическите тестове се пристъпва към формиране на контрастните (екстремалните) групи – силна и слаба, т. е. групите от ученици, които се различават на основата на общия тестов бал. Тук се спазват следните основни стъпки:

1. Пресмята се тестовият бал на всеки ученик, който се получи от сбора на вярно решените от него задачи. Максималният тестов бал е 18.

Таблица 3
КЛЮЧ С ВЕРНИТЕ ОТГОВОРИ

Задача №	Отговор			Задача №	Отговор		
	а	б	в		а	б	в
1.	а	б	✓	10.	✓	б	в
2.	а	✓	в	11.	а	б	✓
3.	✓	б	в	12.	а	✓	в
4.	а	✓	в	13.	✓	б	в
5.	а	б	✓	14.	а	✓	в
6.	а	б	✓	15.	а	б	✓
7.	а	✓	в	16.	✓	б	в
8.	а	б	✓	17.	а	б	✓
9.	а	б	✓	18.	а	✓	в

2. Резултатите се подреждат по реда на нарастване на тестовия бал. Когато има ученици с еднакъв тестов бал те се подреждат последователно на случаен принцип, както е в конкретния случай.

3. Отделят се десетте работи, които са с най-висок тестов бал. Така се оформя силната група, решили вярно между 70 и 100% от тестовите задачи.

4. Отделят се десетте работи, които са с най-нисък тестов бал. По този начин се оформя слабата група, решили вярно под 50% от тестовите задачи.

Работата продължава с анализ на тестовите задачи, като се използват резултатите от двете екстремални групи. По-конкретно се установяват трудността, дискриминативната сила и ефективността на дистракторите за всяка задача. Как става това?

5. Определяне обективната трудност на тестовите задачи.

Известно е, че трудността на всяка тестова задача за постижения е нейна характеристика, показваща статистическото равнище на решаемостта ѝ в дадена извадка. Тя се изчислява в проценти по следната формула:

$$P=100 \cdot \frac{N_R}{N}$$

P – индекс за трудност на тестовата задача;

NR - брой на лицата, решили вярно задачата;

N – общ брой на всички лица, решавали задачата (Бижков, Г., В. Краевски, 2007).

Тук се взема предвид фактът, че критерий за “добра” задача в това отношение при нормативните тестове е $P \in [20\%; 80\%]$ (Стоянова, Ф., 1996).

Трудността на всяка тестова задача се изчислява в проценти по съответната формула, като тук се получават следните резултати: задача №1 - $P = 70\%$, задача №2 - $P = 50\%$, задача №3 - $P = 70\%$, задача №4 - $P = 50\%$, задача №5 - $P = 60\%$, задача №6 - $P = 50\%$, задача №7 - $P = 60\%$, задача №8 - $P = 65\%$, задача №9 - $P = 60\%$, задача №10 - $P = 50\%$, задача №11 - $P = 55\%$, задача №12 - $P = 60\%$, задача №13 - $P = 60\%$, задача №14 - $P = 55\%$, задача №15 - $P = 65\%$, задача №16 - $P = 60\%$, задача №17 - $P = 55\%$, задача №18 - $P = 55\%$.

При сравняване на получените резултати с горния критерий е установено, че тестовите задачи са с трудност, която е в границите на допустимото, т. е. те варират между 50 и 70%.

6. Установяване на дискриминативната сила на тестовите задачи

Дискриминативната сила на тестовата задача нейна характеристика, разграничаваща изследваните лица на “знаещи” и “незнаещи”. Тя се изчислява по формулата:

$$DP = \frac{R_u - R_L}{\frac{1}{2}T}$$

DP – индекс за дискриминативната сила;

RU – брой на учениците от силната група, решили вярно задачата;

RL – брой на учениците от слабата група, решили вярно задачата;

T – общ брой на учениците от двете екстремални групи (Бижков, Г.,

В. Краевски, 2007).

Този момент от анализа е съобразен с факта, че приемливият индекс на дискриминативна сила е между 0,40 и 0,60. Ако задачата е с индекс под 0,40 трябва да се промени, но ако е с индекс под 0,20 трябва да отпадне. С други думи, задачите, които могат да останат в теста или да се променят в известна степен, трябва да са с индекс на дискриминативна сила $0,80 \geq DP \geq 0,20$ (Бижков, Г., В. Краевски, 2007).

На тази основа са получени следните резултати: задача №1 – $DP = 0,60$, задача №2 – $DP = 0,60$, задача №3 – $DP = 0,60$, задача №4 – $DP = 0,60$, задача №5 – $DP = 0,40$, задача №6 – $DP = 0,40$, задача №7 – $DP = 0,60$, задача №8 – $DP = 0,50$, задача №9 – $DP = 0,40$, задача №10 – $DP = 0,60$, задача №11 – $DP = 0,50$, задача №12 – $DP = 0,60$, задача №13 – $DP = 0,60$, задача №14 – $DP = 0,70$, задача №15 – $DP = 0,50$, задача №16 – $DP = 0,60$, задача №17 – $DP = 0,50$, задача №18 – $DP = 0,70$.

Сравнението между получените индекси за дискриминативна сила на задачите от теста по български език за края на учебната година и описаните изисквания показват, че те се вменват в рамките на допустимото, тъй като варират в границите между 0,40 и 0,70. Тези данни дават основание да се запазят тестовите задачи в този си вид.

7. Изчисляване ефективността на дистракторите

Ефективността на дистракторите е необходима при тестовете с множествен избор на отговорите, за да се установи дали и в каква степен дистракторите са приемливи за учениците и доколко те позволяват да се разграничават силните от слабите ученици. В случая тя се изчислява по следната формула:

$$D = \frac{R_u}{N_u} - \frac{R_D}{N_D}$$

D – индекс за трудност на задачата;

RU – брой на учениците от силната група, избрали съответния

дистрактор;

RD – брой на учениците от слабата група, избрали съответния

дистрактор;

NU – общ брой на учениците в силната група;

ND – общ брой на учениците в слабата група (Стоянова, Ф., 1996).

Тук се вземат предвид следните факти:

1. “За разлика от верния отговор един дистрактор е добър, когато относителният дял на учениците от слабата група, които са го предпочели, е по-голям от относителния дял на избралите го ученици от силната група.” (Стоянова, Ф., 1996).

2. Дискриминативната сила на дистракторите трябва да е отрицателна (Стоянова, Ф., 1996).

3. “Един дистрактор е ефективен, ако неговата дискриминативна сила е в интервала $[-0,20; -1,00]$.” (Стоянова, Ф., 1996).

Изчисляването на ефективността на дистракторите се осъществява по съответната формула, като за целта предварително се съставя таблица, включваща

боя на учениците от силната и слабата група, които са избрали съответния отговор по всяка задача. Таблицата може да има следният вид.

Таблица 4

Задача №	1.			2.			3.			4.			5.			6.			7.			8.			9.		
Отговори	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в
	Екстремални групи																										
Силна група	0	0	1	1	8	1	0	1	0	2	8	0	1	1	8	2	1	7	0	9	1	1	0	9	2	0	8
Слаба група	2	4	4	4	2	4	3	4	3	5	2	3	3	3	4	4	3	3	3	3	4	4	2	4	4	2	4

Задача №	10.			11.			12.			13.			14.			15.			16.			17.			18.		
Отговори	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в	а	б	в
	Екстремални групи																										
Силна група	8	1	1	1	1	8	1	9	0	9	0	1	0	9	1	0	1	9	9	1	0	2	0	8	1	9	0
Слаба група	2	5	3	4	3	3	3	3	4	3	4	3	4	2	4	3	3	4	3	4	3	4	3	4	3	3	5

Резултатите са представени в следващата таблица.

Таблица 5

ЕФЕКТИВНОСТ НА ДИСТРАКТОРИТЕ

Задача №	Отговори			Задача №	Отговори		
	а)	б)	в)		а)	б)	в)
1	- 0,20	- 0,40	0,60	10	0,60	- 0,40	- 0,20
2	- 0,30	0,60	- 0,30	11	- 0,30	- 0,20	0,50
3	- 0,30	0,60	- 0,30	12	- 0,20	0,60	- 0,40
4	- 0,30	0,60	- 0,30	13	0,60	- 0,40	- 0,20
5	- 0,20	- 0,20	0,40	14	- 0,40	0,70	- 0,30
6	- 0,20	- 0,20	0,40	15	- 0,30	- 0,20	0,50
7	- 0,30	0,60	- 0,30	16	0,60	- 0,30	- 0,30
8	- 0,30	- 0,20	0,50	17	- 0,20	- 0,30	0,50
9	- 0,20	- 0,20	0,40	18	- 0,40	0,70	- 0,30

Данните показват, че в този тест дистракторите са със стойности между – 0,20 и –0,40, което е основание да се смята, че те са ефективни, имайки предвид описаните по-горе изисквания.

След като са установени качествата на задачите се пристъпва към анализ на качествата на теста като цяло.

✳ Анализ на качествата на теста като цяло

Тук се проверяват надеждност, валидност, обективност и приложимост на съставения тест.

1. Установяване на надеждността на теста

Известно е, че надеждността на един тест показва точността на този диагностичен инструмент и устойчивостта на получените чрез него резултати от към действието на странични и случайни фактори. Като се взема предвид това, а и съществуващите подходи за анализ на надеждността на един тест, разработени в класическата теория на тестовете, тук се използва анализът на вътрешната съгласуваност. При този подход за изчисляване на надеждността е необходимо да се

ползват данните от еднократното изпробване на теста, като се излиза от допускането, че съответният тест може да се раздели на толкова части, колкото задачи съдържа. На тази база се определя и коефициентът на корелация между резултатите от всяка отделна задача и общият резултат от целия тест, който всъщност показва надеждността на дидактическият тест (Бижков, Г., В. Краевски, 2007). Като се има предвид, че резултатите от този дидактически тест се измерват на базата на нормативния подход, се избират и подходящите процедури и формули за пресмятане коефициента на надеждност. Тя е изчислена по една от най-често използваните формули за определянето ѝ в смисъл на вътрешна съгласуваност при тестове със задачи с изборен отговор, а именно това е формулата на Кюдер и Ричардсон, имаща следното съдържание:

$$KR_{21} = \frac{K}{K-1} \left[1 - \frac{M \cdot (K - M)}{K \cdot s^2} \right]$$

KR_{21} – коефициентът за надеждност на теста;

K – общият брой тестови задачи, т. е. дължината на теста;

M – средна аритметична стойност на тестовия бал;

s – стандартното отклонение на тестовия бал (Стоянова, Ф., 1996).

Работата продължава със съставянето на следващата таблица, която включва степента на трудност за всяка задача както и стойността на нейната дискриминативна сила.

Таблица 6

Задача №	P	DP
1	60	0,60
2	65	0,50
3	55	0,50
4	60	0,40
5	55	0,50
6	65	0,50
7	50	0,40
8	50	0,40
9	65	0,50
10	60	0,60
11	70	0,60
12	65	0,50
13	60	0,60
14	65	0,50
15	60	0,60
16	55	0,50
17	65	0,70
18	45	0,50
Σ	1070	9,40
Σ : 18 = средна стойност	1070:18 =59,44	9,40:18 =0,52

Тези данни са необходими за осъществяване на останалите процедури, които са следните:

➤ Събират се индексите за трудност на всички тестови задачи, след което полученото число се разделя на 100, за да се открие числото, съвпадащо със средната аритметична стойност на тестовия бал, която в случая е 10,50.

➤ Събират се индексите за дискриминативна сила на всички тестови задачи, след което се пресмята стандартното отклонение по формулата:

$$s = \frac{\sum D_i}{\sqrt{6}}$$

D_i – показателят за дискриминативна сила на задача – i (Стоянова, Ф., 1996). В конкретния тест се установи, че $S = 4,08$.

➤ След като са станали ясни стойностите на основните характеристики на честотното разпределение на тестовия бал (M и S) се пристъпва към изчисляване коефициента на надеждност по формулата на Кюдер и Ричардсон.

Тук се имат предвид и следните факти: коефициентът на надеждност може да приема стойности между 0 и 1; надеждността, която е изразена чрез съответния индекс не бива да е по-малка от 0,65 (Стоянова, Ф., 1996).

Резултатът от тези процедури показва, че коефициентът на надеждност тук е със стойност $KR_{21} = 0,78$.

Като се имат предвид, получения резултат, а също и фактите, че коефициентът на надеждност може да приема стойности между 0 и 1 и че надеждността, изразена чрез съответния индекс, не бива да е по-малка от 0,65 се стига до заключението, че този тест е с достатъчно висока надеждност, което показва, че това качество на теста е налице.

➤ Работата продължава с изчисляване стандартната грешка на измерването, което се налага, тъй като всеизвестно е, че резултатите от всяко измерване повече или по-малко се различават от истинската стойност. По своята същност тя представлява стандартното отклонение на грешките, които могат да се получат при различни тестове от дадена област. Тук стандартната грешка на измерването се изчислява по следната формула:

$$SE = s \cdot \sqrt{1 - R}$$

SE – стандартната грешка на измерването;

S – стандартното отклонение на тестовия бал;

R – коефициентът на надеждност на теста (Стоянова, Ф., 1996).

Резултатът от прилагането на формулата показва, че $SE = 1,92$, което означава, че за 2/3 от изследваните ученици полученият тестов бал се различава с по-малко от 1,92 единици от истинската стойност на измерването и само за 5% от учениците разликата между техния тестов бал и истинската стойност е по-голяма от 3,84, което се получава от следното произведение: $2 \cdot SE$. Тези стойности показват, че стандартната грешка на измерването е в рамките на допустимото, т.е. тя е по-малка от две единици, което е характерно предимно за тестове с коефициент на надеждност по-голям от 0,85.

2. Установяване валидността на теста

На този етап от анализа на качествата на тестовете се пристъпва към установяване на тяхната съдържателна валидност, която, както е известно, показва степента на съответствие между съдържанието на включените в един тест задачи и учебното съдържание, чието усвояване той е предназначен да измерва (Стоянова, Ф., 1996). Като се има предвид, че “..... съдържателната валидност на теста, независимо от това дали той е нормативен или критериален, се прави от компетентни лица (експерти).” (Бижков, Г., В. Краевски, 2007), то за конкретния случай участват двадесет експерта. На всеки експерт е дадена бланка за експертна оценка на тестовите задачи.

Съдържателната валидност е установена при спазване на следните процедури:

1. Съставена е групата от експерти: двама експерти по начално образование, двама директори на начални училища и 16 начални учители.

2. Раздадени са бланките за експертна оценка, във всяка от които експертите са отбелязали в каква степен задачите отговарят на съответните цели от учебната програма по български език за IV клас. Степента е определена по следната скала: 0 – не съответства, 1 – съответства отчасти, 2 – съответства изцяло.

3. След предаване на бланките е установен броят на експертите, дали най-висока оценка.

4. Премахва се към изчисляване коефициента на съдържателна валидност на всяка тестова задача, като се прилага методът на Лоуш. Формулата в този случай е следната:

$$V_c = \frac{n_a - \frac{N}{2}}{\frac{N}{2}}$$

V_c – коефициент за съдържателна валидност на тестовата задача;

n_a – броят експерти, които са посочили максималната оценка по скалата;

N – общият брой експерти (Стоянова, Ф., 1996).

Резултатите за всяка задача, са следните: задача №1 – $V_c = 0,60$, задача №2 – $V_c = 0,70$, задача №3 – $V_c = 0,80$, задача №4 – $V_c = 1$, задача №5 – $V_c = 0,60$, задача №6 – $V_c = 0,60$, задача №7 – $V_c = 1$, задача №8 – $V_c = 0,70$, задача №9 – $V_c = 1$, задача №10 – $V_c = 0,80$, задача №11 – $V_c = 1$, задача №12 – $V_c = 1$, задача №13 – $V_c = 0,60$, задача №14 – $V_c = 0,60$, задача №15 – $V_c = 0,90$, задача №16 – $V_c = 0,80$, задача №17 – $V_c = 1$, задача №18 – $V_c = 0,90$.

5. Изчислява се коефициентът за съдържателна валидност на целия тест. Той се получава като сборът от коефициентите за съдържателна валидност на всяка задача се разделя на броя на задачите в теста и по този начин средноаритметичната стойност на всички задачи показва коефициента за съдържателна валидност на съответния тест. В случая съдържателната валидност е 0,81.

На този етап се вземат предвид следните факти: ако броят на експертите е по-малък от 5, не могат да се направят статистически достоверни изводи относно съдържателната валидност на теста; една задача или цял тест има достатъчно висока съдържателна валидност, ако получената стойност за V_c е по-голяма от минималната стойност, показана в таблица, разкриваща критичните стойности на коефициента на Лоуш (по Стоянова, Ф., 1996).

На тази база е установено, че тестът има достатъчно висок коефициент на съдържателна валидност.

Спазването на всички описани етапи, процедури и правила при съставянето на дидактическия тест дават достатъчно основания да се смята, че резултатите от него няма да бъдат повлияни от субективни предпочитания, т. е. критерият *обективност* е налице.

Освен това, чрез наблюдение на процеса на апробация на дидактическия тест в конкретна училищна среда, е доказано че критерият *приложимост* също е удовлетворен, тъй като те са лесни за използване и лесни за отчитане на резултатите.

В заключение може да се каже, че описаната технология по съставянето на нормативен дидактически тест е приложима не само от педагози в училище, но и от преподаватели и студенти във висшите учебни заведения. Тя би била особено полезна при съставяне на кандидат-студентски тестове, както и при провеждане на научни изследвания в областта на образованието.

ЛИТЕРАТУРА

1. Бижков, Г., Педагогическа диагностика, Унив. Издателство “Св. Климент Охридски”, София, 1999
2. Бижков, Г., В. Краевски, Методология и методи на педагогическите изследвания,
3. Унив. Издателство “Св. Климент Охридски”, София, 2007
4. Радев, Пл., А. Александрова, Дидактическа тестология или тестиране в образованието, Пловдивско университетско издателство, Пловдив, 2003
5. Стоянова, Ф., Тестология за учители, Издателство Атика, София, 1996

Гр. Пловдив
Бул. България 236
Пловдивски университет “Паисий Хилендарски”
Педагогически факултет
Катедра “НУП”
Гл.ас. д-р Николинка Александрова Атанасова
e-mail: al_nina@abv.bg
GSM: 0886382707