

УНИВЕРЗИТЕТ У БЕОГРАДУ

ФИЛОЗОФСКИ ФАКУЛТЕТ

Ивана М. Луковић

ТЕСТОВИ ЗНАЊА У ШКОЛСКОЈ ПРАКСИ

докторска дисертација

Београд, 2016

UNIVERSITY OF BELGRADE

FACULTY OF PHILOSOPHY

Ivana M. Luković

ACHIEVEMENT TESTS IN SCHOOL PRACTICE

Doctoral Dissertation

Belgrade, 2016

Ментор:

др Наташа Матовић, ванредни професор, Универзитет у Београду,
Филозофски факултет

Чланови комисије:

1. др Наташа Вујисић-Живковић, редовни професор, Универзитет у Београду, Филозофски факултет
2. др Биљана Бодрошки-Спариосу, доцент, Универзитет у Београду, Филозофски факултет
3. др Саша Дубљанин, доцент, Универзитет у Београду, Филозофски факултет

Датум одбране: _____

ТЕСТОВИ ЗНАЊА У ШКОЛСКОЈ ПРАКСИ

САЖЕТАК

У раду се разматрају могућности примене тестова знања у школској пракси, карактеристике примењених тестова и начини на које се могу користити резултати добијени тестовима знања. Иако тестови знања нису нов истраживачки проблем, последњих деценија дошло је до промена на плану теорија тестирања и приступа конструкцији теста. Намера истраживача је била да се постојећа сазнања у вези са тестовима допуне испитивањем методолошких карактеристика тестова знања који се користе у нашој школској пракси. У истраживању је коришћена дескриптивно-аналитичка метода, технике анкетања, скалирања, анализе садржаја и интервјуисања. Испитивање је реализовано на узорку од 362 наставника предметне наставе. Истраживањем је утврђено да су могућности за примену тестова релативно повољне, али се могу унапређивати, изменама у сфери иницијалног професионалног образовања и стручног усавршавања наставника. Резултати тестова знања користе се превасходно у сврхе сумативне евалуације, уз извесно препознавање потенцијала који примена тестова може имати и у домену формативне евалуације. Када је реч о могућностима примене тестова знања и начинима коришћења резултата добијених тестирањем, године радног стажа у већој мери објашњавају разлику у одговорима испитаника у односу на предмет који наставници предају. Анализом чији циљ је био да се опишу методолошке карактеристике тестова обухваћено је 80 тестова знања израђених за потребе наставе српског језика и математике. Овог пута, регистроване су значајне разлике у појединим карактеристикама тестова. Методолошка знања наставника разликују се у доменима препознавања фаза конструкције теста, карактеристика различитих типова задатака, нивоа знања који се тестом могу мерити и стандардизације услова тестирања.

Кључне речи: тест знања, конструкција теста знања, методолошке карактеристике тестова знања, примена тестова знања у школској пракси.

Научна област: Педагогија

Ужа научна област: Општа педагогија са методологијом и историја педагогије

УДК: 37.012:371.263

ACHIEVEMENT TESTS IN SCHOOL PRACTICE

ABSTRACT

This paper concerns potentials for applying achievement tests in school practice, characteristics of applied tests and ways in which testing results can be used. Although achievement tests are not a new research topic, over the past decades changes have occurred in testing theory and construction of achievement tests. Our intention was to complement relevant scientific findings with new findings about the characteristics of tests applied in our school practice. The research was implemented using the descriptive method, techniques of survey, scaling, content analysis and interviews. Field research was conducted on a sample of 362 elementary school teachers. Results show that potentials for applying achievement tests are good, but they can be promoted according to changes in initial professional education and vocational training of teachers. Testing results are commonly used for the purpose of summative evaluation, with some tendency toward using results for the purpose of formative evaluation. Referring to potentials for applying tests and ways in which testing results are used, differences between teachers' opinions are better explained considering years of teachers' practice service than considering subject domains that they teach. Aiming to describe methodological characteristics of applied tests, we analyzed 80 achievement tests applied in courses of Serbian language and mathematics. This time, significant differences were found between the characteristics of the tests. Methodological knowledge of teachers differs in domains of identifying phases in test construction, characteristics of some types of questions, levels of knowledge that can be measured by a test and standardization of testing requirements.

Key words: the achievement test, construction of achievement tests, methodological characteristics of achievement tests, application of achievement tests in school practice.

Scientific field: Pedagogy

Scientific subfield: General Pedagogy with Methodology and History of pedagogy

UDC: 37.012:371.263

САДРЖАЈ

УВОД	1
ТЕОРИЈСКИ ПРИСТУП ИСТРАЖИВАЊУ	4
ТЕСТ ЗНАЊА КАО МЕРНИ ИНСТРУМЕНТ	5
Одређење теста знања	5
Врсте тестова знања	11
Две теорије тестирања: класична теорија и теорија ставског одговора	21
КОНСТРУКЦИЈА ТЕСТА ЗНАЊА.....	29
Фазе у конструкцији теста знања	29
Приступи разради садржаја теста знања	36
Задатак – основна јединица теста знања.....	47
<i>Терминолошке дилеме у вези са одређењем задатка</i>	<i>47</i>
<i>Врсте задатака и њихове карактеристике</i>	<i>51</i>
<i>Критеријуми за утврђивање и бодовање тачних одговора.....</i>	<i>73</i>
Метријске карактеристике теста знања	77
<i>Провера метријских карактеристика теста знања.....</i>	<i>77</i>
<i>Обликовање коначне верзије теста знања.....</i>	<i>93</i>
ПРИМЕНА ТЕСТА ЗНАЊА У ШКОЛСКОЈ ПРАКСИ.....	98
Специфичности процеса конструкције тестова знања које наставници самостално израђују	98
Особености тестова знања који се користе у настави за праћење и вредновање рада ученика	109

Примена стандардизованих тестова знања у испитивању образовних постигнућа ученика	116
Припрема и реализација тестирања.....	121
Могућности коришћења резултата добијених применом теста знања за унапређивање васпитно-образовног рада школе.....	126
МЕТОДОЛОГИЈА ИСТРАЖИВАЊА.....	140
ПРЕДМЕТ ИСТРАЖИВАЊА И ДЕФИНИСАЊЕ ОСНОВНИХ ПОЈМОВА	141
ЦИЉ И ЗАДАЦИ ИСТРАЖИВАЊА.....	142
ХИПОТЕЗЕ У ИСТРАЖИВАЊУ.....	143
ВАРИЈАБЛЕ У ИСТРАЖИВАЊУ	145
УЗОРАК ИСТРАЖИВАЊА	147
МЕТОДЕ, ТЕХНИКЕ И ИНСТРУМЕНТИ ИСТРАЖИВАЊА.....	153
НИВО И НАЧИН ОБРАДЕ ПОДАТАКА.....	155
ОРГАНИЗАЦИЈА И ТОК ИСТРАЖИВАЊА	155
АНАЛИЗА И ИНТЕРПРЕТАЦИЈА РЕЗУЛТАТА ИСТРАЖИВАЊА.....	156
МОГУЋНОСТИ ПРИМЕНЕ ТЕСТА ЗНАЊА У НАСТАВИ.....	157
Предности и ограничења теста знања у односу на друге облике праћења и вредновања рада ученика.....	157
Припремљеност наставника за израду теста знања	168
<i>Допринос иницијалног професионалног образовања припреми наставника за израду теста знања</i>	<i>170</i>

<i>Допринос стручног усавршавања припреми наставника за израду теста знања</i>	<i>174</i>
Тешкоће наставника у процесу израде теста знања у настави и начини њиховог превазилажења	183
Учесталост и динамика коришћења теста знања у настави	186
Предмет који наставници предају и могућности примене теста знања у настави	188
Године радног стажа наставника у просвети и могућности примене теста знања у настави	194
МЕТОДОЛОШКЕ КАРАКТЕРИСТИКЕ ТЕСТОВА ЗНАЊА У НАСТАВИ СРПСКОГ ЈЕЗИКА И МАТЕМАТИКЕ.....	204
Фазе конструкције теста знања	204
Садржај теста знања	208
Структура теста знања	221
Начин утврђивања постигнућа на тесту знања	231
Провера квалитета теста знања.....	238
Припрема за тестирање и реализација тестирања у настави.....	244
Сличности и разлике између методолошких карактеристика тестова знања у настави српског језика и математике	250
КОРИШЋЕЊЕ РЕЗУЛТАТА ДОБИЈЕНИХ ПРИМЕНОМ ТЕСТА ЗНАЊА ЗА УНАПРЕЂИВАЊЕ ВАСПИТНО-ОБРАЗОВНОГ РАДА ШКОЛЕ.....	262
ЗАКЉУЧАК.....	272
ЛИТЕРАТУРА.....	285
ПРИЛОЗИ	296

УВОД

Реализација педагошког истраживања обухвата више етапа. Етапа прикупљања података представља једну од најсложенијих и може се остварити применом различитих истраживачких техника. Једна од техника која се користе у циљу прикупљања података јесте техника тестирања. Реч је о техници која се ослања на примену тестова којима се могу испитивати различита својства – знања, вештине, способности, црте личности, темперамент и слично. За домен педагошке праксе од посебног значаја је примена тестова знања. Тестови знања су врста инструмента чија примена може да обезбеди високу валидност, поузданост, дискриминативност и објективност података добијених у вези са знањем као предметом мерења. Одликује их сложена конструкција, специфични услови примене, као и избор оквира интерпретације резултата добијених тестирањем. Тестови се могу применити како би се проверила или продубила сазнања стечена применом других инструмената, али исто тако, резултати тестирања могу покренути низ додатних испитивања применом других инструмената или нових тестова знања.

Бројна питања у вези са применом тестова знања, разматрана су у педагошкој литератури ранијих периода. Међутим, са појавом нових теорија тестирања, модела мерења, другачијих интерпретација знања као својства мерења, на стара питања било је потребно понудити нове одговоре. Начин на који се данас приступа мерењу знања другачији је него у дугом раздобљу истраживања тестова током XX века. Упркос томе, не може се очекивати да савремени приступ пружа коначне одговоре у вези са конструкцијом и применом тестова знања.

Развоју тестова у савремено доба доприноси и информатичка технологија. Све се више говори о електронским тестовима као новој врсти тестова знања. Конципирање тестова знања уз помоћ рачунара израз је тежње да се конструишу тестови који ће имати повољнија метријска својства у односу на класичне тестове знања. Међутим, на тренутном стадијуму развоја науке,

конструкција електронских тестова није у целини ослобођена појединих методолошких дилема које прате и конструкцију класичних тестова знања.

Традиционално, истраживања проблематике тестова знања у највећој мери су оријентисана на конструкцију тестова који ће се користити у научном истраживању. Тестови које наставници самостално израђују углавном су били предмет изучавања у области дидактике. Врло су ретка испитивања која проблему приступају са становишта методологије педагошких истраживања. Недостају нам сазнања о томе какве су методолошке карактеристике тестова знања које наставници самостално конструишу, шта одликује процес примене тестова знања у школи, како наставници процењују своје компетенције у домену израде и примене тестова. Таква сазнања су нам неопходна да бисмо могли потпуније да сагледамо процес конструкције и примене тестова знања у школској пракси, и како би се могло приступити осмишљавању стратегија за унапређивање релевантних стручних компетенција наставника.

Иако је неоспорно да тестови знања налазе широку примену у области наставе, још увек се дискутује о питању домета њихове примене. Одређени аутори наглашавају предности тестова знања у односу на друге начине праћења и вредновања постигнућа ученика, пре свега са становишта метријских својстава која поседују. Истовремено, други аутори наглашавају ограничења тестова с обзиром на усмереност на резултат (учења) и занемаривање процеса којим се стиже до резултата.

Поред израде тестова за потребе редовне наставне праксе, тестови знања се све чешће примењују и у различитим евалуативним испитивањима у националним и међународним оквирима. То, заправо, значи да се тестови данас не користе само у циљу праћења и вредновања ученичких постигнућа, већ и знатно шире. Имајући то у виду, важно је истражити стварни потенцијал тестова у новим доменима примене, пре свега у домену креирања образовних политика.

У оквиру теоријског дела рада разматрају се три питања: одређење теста као мерног инструмента, конструкција теста знања, примена тестова знања које наставници самостално израђују и стандардизованих тестова знања.

Теоријски приступ истраживању прати приказ методологије и резултата испитивања. Резултати истраживања груписани су у три целине. У оквиру прве целине, приказују се и дискутују могућности примене теста знања у настави са становишта: предности и ограничења теста у односу на друге облике праћења и вредновања рада ученика, припремљености наставника за израду теста знања, тешкоћа са којима се суочавају у процесу конструкције теста, учесталости и динамике коришћења теста знања. Такође, у првој целини приказују се и анализирају резултати у вези са испитивањем разлика у сагледавању могућности примене теста знања од стране наставника, с обзиром на предмет који предају и године радног стажа. У оквиру друге целине, приказују се и дискутују резултати у вези са методолошким карактеристикама тестова знања који се користе у настави српског језика и математике. Трећом целином обухваћен је приказ и анализа резултата који се односе на коришћење резултата добијених применом теста знања за унапређивање васпитно-образовног рада школе. У закључку рада дат је осврт на најважније резултате истраживања и указано на нека питања у вези са применом тестова знања која би се могла даље истраживати.

ТЕОРИЈСКИ ПРИСТУП ИСТРАЖИВАЊУ

ТЕСТ ЗНАЊА КАО МЕРНИ ИНСТРУМЕНТ

Одређење теста знања

Претечом савремених тестова знања сматрају се испити за највише државне функције, конципирани у VII веку у Кини. На испитима, кандидати су имали задатак да напишу есеј на задату тему (Mathews, 2006). Вековима након тога, израда есеја као форма писане провере знања заједно са усменим проверама, представљаће окоснице испитивања ученичких знања. Од средине до краја XIX века било је више покушаја увођења објективног испитивања ученичког знања (Бобавац и сар, 1972; Gallagher, 2003). Ипак, препознавање потенцијала тестова знања у домену наставне праксе, догодиће се тек након експанзије психолошких мерења интелигенције, почетком XX века. Подстицај развоју тестова знања пружиће Торндајкова разматрања проблематике тестова знања (1904; 1910), појава Бине-Симоновог теста интелигенције (1905), Келијево конципирање задатака са вишеструким понуђеним одговорима (1914) и примена Отис-Јеркесовог *Army Alfa/Beta* теста (1914) у раду са групама испитаника (Sax, 1989; Clarke *et al.*, 2000; Popham, 2001). Уследиће деценије обележене интензивном продукцијом различитих тестова знања и њихово имплементирање у наставну праксу, најпре у Америци, а касније и у другим земљама света (Мужић, 1968).

Реч „тест“ потиче од латинског глагола „testor“ који се преводи као „посведочити“ или „доказати“ (Жепић, 1941). У енглеском језику под тестирањем се подразумева испитивање, проверавање, процедура праћења и проверавања појединих својстава (Мужић, 2004). Тако, Америчка асоцијација психолога тест одређује као стандардизовани поступак мерења, са циљем процењивања одговарајућих својстава испитаника (Turner *et al.*, 2001).

Сакс (Sax, 1989) наглашава да је суштинско одређење теста постојање задатака чијем решавању испитаник приступа. При том, задаци могу да мере различита својства – знања, вештине или одређене способности. Код тестова је

од пресудне важности да се одговори испитаника могу довести у везу са својством које се тестом мери. Поменута релација наглашена је и код других страних аутора (*Testing in American Schools: Asking the Right Questions*, 1992; Postlethwaite, 2005).

Америчка асоцијација психолога прецизира да се на основу одговора испитаника на задатке у тесту добија, заправо, узорак понашања испитаника чије се вредновање и бодовање остварује стандардизованом процедуром (Turner *et al.*, 2001). Мерено својство могуће је проценити на основу добијеног узорка понашања. Поједини аутори иду и корак даље, и тест изједначавају са узорком понашања (на пример: Vansickle, 2008).

У страниј литератури, тестови се често дефинишу са становишта функције коју имају. Тестови постигнућа, на пример, одређују се као тестови којима се процењује колико је знања или вештина ученик стекао током процеса учења (Ebel, 1979; Cizek, 1998). У погледу структуре, тестови постигнућа садрже низ задатака од којих сваки има тачан одговор (Ebel & Frisbie, 1991).

У домаћој литератури, тест се употребљава најчешће у оном значењу које има у психологији. Истраживачи у области педагогије (на пример: Мужих, 1986; Гојков, 2012) често се ослањају на дефиницију теста психолога Зорана Бујаса. Према становишту Бујаса (1959), тест је стандардизовани поступак којим се изазива одређена активност и чији учинак се вреднује упоређивањем индивидуалног резултата са резултатима других субјеката у истоветној ситуацији, или упоређивањем са једнозначно постављеним критеријумом. Слично, Буквић (1996) одређује тест као устаљени поступак за изазивање одређене активности на основу које се испитаници могу поредити. Дата активност се испољава као резултат деловања одговарајућих подстицаја (на пример, постављања одређених питања испитанику). На готово идентичан начин, тест дефинише и Фајгељ (2004) истичући да је реч о стандардизованом поступку за испитивање одређеног понашања помоћу проверених дражи на основу којих се испитаници пореде по особини која је предмет мерења.

Наведене дефиниције садрже структурално истоветне елементе. Разлике које се могу уочити у појединачним дефиницијама углавном су терминолошке природе. Тестирање се сагледава као поступак који се припрема, реализује и чији се резултати вреднују према унапред одређеним правилима и принципима. Резултати тестирања омогућавају да се испитаници међусобно упореде у погледу оствареног постигнућа, или да се постигнуће испитаника упореди са постављеним критеријумом. Узорак понашања који омогућава да се остваре неопходне процене резултат је испитаникових одговора на постављене захтеве у тесту.

Све приказане дефиниције одликује поистовећивање теста као инструмента са тестирањем као техником. Повремено се тест означава као „инструмент или процедура“ (на пример: Postlethwaite, 2005: 36). Иако границу између технике и инструмента није увек лако одредити, постоји тенденција да се у методолошкој литератури појам инструмента означава ужим од појма технике (Банђур и Поткоњак, 1999). У случају тестова та разлика није наглашена. Образложење се може потражити у високом степену стандардизованости понашања субјеката који су укључени у процес тестирања (истраживача/наставника и испитаника) и просторно-временских услова у којима се тестирање дешава. Тест се схвата као „нешто више од инструмента“.

Због чега је наглашена стандардизација поступка тестирања? Устаљеност или стандардизованост поступка тестирања неопходан је услов међусобног поређења постигнућа које испитаници остварују, као и процењивања постигнућа испитаника у односу на постављени критеријум. Стандардизованост поступка тестирања у ширем смислу значи да сваки тест мора поседовати одређене метријске карактеристике да би се могао применити. У ужем смислу, стандардизованост подразумева дефинисани начин задавања теста, оцењивања активности испитаника на тесту и тумачења коначних резултата тестирања (Буквић, 1996; Vansickle, 2008).

У којој мери је стандардизација тестирања важна, може се закључити на основу све заступљенијег поистовећивања тестова са стандардизованим тестовима (Hambleton & Jirka, 2004). Иако тестови често јесу стандардизовани,

они могу бити и нестандардизовани (неформални). Категорији нестандардизованих тестова припадали би и тестови које наставници самостално израђују (Rudner & Schafer, 2002). За нестандардизоване тестове карактеристично је да им нису испитане метријске карактеристике. У домаћој литератури такви тестови се повремено означавају као *низови задатака објективног типа* (Бјекић и Папић, 2005). Њихова примена у наставне сврхе и у домену квалитативних педагошких истраживања сасвим је оправдана (Мужић, 2004). Међутим, трагање за правим тестовима знања обично је ограничено на домен квантитативних истраживања у којима се сложеним поступцима провере и баждарења добијају тестови високе ваљаности, поузданости, објективности и дискриминативности.

У дефиницији тестова Америчке асоцијације психолога (Turner *et al.*, 2001), наглашено је да се тестирање разуме као поступак мерења. Самим тим, тест би се могао означити мерним инструментом. Означавање теста као мерног инструмената проистиче и из чињенице да су тестови инструменти који су развијени у области психометрије. Као што и сам назив научне дисциплине каже, психометрија се бави „мерењем у психологији“ (Фајгел, 2005: 21). Фајгел (2004: 295) иде корак даље и мерење означава „суштином постојања и примењивања теста“. Иако се тестом могу испитивати и различита квалитативна својства (темперамент, екстровертност-интровертност личности и слично) тест увек одликује тежња да се утврди и количина мереног својства (Матовић, 2007). Наведену специфичност је важно нагласити, јер она није својствена свим инструментима који налазе примену у педагошким истраживањима (на пример, анкетама и различитим врстама протокола интервјуа).

Применом тестова, мерење знања се остварује посредно, а поступак мерења има одлике изведеног мерења. Изведена мерења подразумевају да се одређено својство мери на основу ефеката или последица које изазива на другим својствима или објектима. На пример, јачина земљотреса према Меркалијевој скали одређује се на основу оштећења грађевинских објеката и промена у изгледу земљишта на подручју земљотреса. Истраживања у домену

нуклеарне физике такође се ослањају на посредне облике мерења. Најситније честице материје ниједан научник није видео, али су након мерења промена које се дешавају у језгру атома потврђене хипотезе да постоје честице ситније од атома. На сличан начин реализују су мерења и у области педагогије и психологије (Буквић, 2007).

Примена теста подразумева мерење својства које није могуће регистровати чулима. Таква својства се обично означавају као латентна или скривена својства (Hambelton & Cook, 1977; Nowak, 1977; Traub & Wolfe, 1981). Латентна својства имају своју манифестацију која се може регистровати чулима. Заправо, испитивањем манифестација латентног својства стичемо сазнања о самом латентном својству. Примењено у области конструкције тестова знања, то би значило да су истраживачу потребни показатељи знања који су очигледни или манифестни, како би могао да „измери“ знање. Манифестна својства знања могу бити решења задатака у тесту, усмени одговори или практичне радње у одговарајућој испитној ситуацији. Дакле, кључни задатак истраживача који саставља тест је да конструише довољно добре задатке, односно задатке који су заиста мера знања у одговарајућој области. Колико је то тежак посао, сликовито може да представи ситуација два наставника истог предмета код којих ученици добијају различите тестове знања и на којима постигнуће истих ученика може да се разликује. Два различита наставника за тест знања из исте области одабраће задатке које лично сматрају најрелевантнијим. При том, њихове процене могу бити потпуно различите. Избор задатака ће зависити од стручних и методичких компетенција наставника, компетенција за израду теста знања, њиховог односа према раду и познавања ученика.

Знање је комплексно својство чије мерење одликују и друге специфичности. Наиме, поседовање „знања“ није довољно одредити само преко присуства одговарајућег показатеља. Присуство показатеља допуњује се додатним критеријумима као што су учесталост показатеља, њихов квалитет, количина, ширина, јачина, брзина јављања и трајање (Буквић, 2007). За тестове знања посебно су релевантни критеријуми количине и квалитета (Буквић,

2007). У вези са критеријумом количине, очекује се да испитаник који успешно реши више задатака у тесту поседује знање у „вишем степену“ у односу на испитаника који успешно реши мањи број задатака. Такође, очекује се да испитаник који успешно реши више тешких задатака поседује више знања од испитаника који успешно реши мањи број тешких задатака (критеријум квалитета).

Са појавом теорије латентних својстава и након њене успешне примене у области психологије, појавила су се очекивања да теорија латентних својстава може допринети и развоју тестова знања. Утемељење тестова знања на моделима теорије латентног својства пружало би тестовима сигурну теоријску основу (Traub & Wolfe, 1981) за којом се дуго трагало. Међутим, веза индикатора знања и знања као мереног својства, показала се сложенијом него што се очекивало. Различити аутори су заступали становиште да се знање не може посматрати као једнодимензионални конструкт (Traub & Wolfe, 1981), чиме је доведена у питање могућност да се знање означи латентном варијаблом (Фајгељ, 2005). Самим тим, конципирање теста на моделима теорије латентног својства није се сматрало оправданим.

Како би се надоместио недостатак теоријског оквира који омогућава да се боље разуме суштина знања, школско знање се у новијој литератури означава као емергентна варијабла (Фајгељ, 2005). Код емергентних варијабли, скуп индикатора представља списак дескрипција које, према правилу, дефинишу неко стање. Конкретан низ дескрипција је само један од могућих низова дескрипција и истраживач самостално одабира низ које ће означити релевантним. Релација индикатора и емергентних варијабли је другачија него у случају латентних варијабли. Латентна варијабла индукује понашање које се мери, док скуп индикатора операционално дефинише емергентне варијабле.

Иако би се могло очекивати да је нови теоријски приступ понудио одговоре у вези са разумевањем знања као својства које се мери, интересовање за моделе теорије латентних својстава поново је актуелизовано. Напретком компјутерске технологије и статистике омогућено је конструисање модела мерења који су мултидимензионални, што значи да једнодимензионалност

својства више није неопходан услов примене модела теорије латентних својстава. Конципирање тестова знања данас је могуће у оквирима теорије латентних својстава (He & Tumm, 2005).

Врсте тестова знања

Класификације тестова које се разматрају у литератури, утемељене су на различитим критеријума. Уз извесна преклапања, класификације се међусобно разликују у врсти и броју критеријума који се препознају као основа поделе. Критеријуми који се најчешће користе ослоњени су на: предмет мерења, намену теста, димензију брзина—снага, стандардизованост, начин решавања задатака у тесту и број испитаника (Баковљев, 1997).

Према предмету мерења, односно према ономе што се тестом мери, тестови се деле на (Баковљев, 1997):

- *тестове могућности* (називају се и тестови у ужем смислу, тестови знања, умења и способности) и
- *остале тестове* (називају се и тестови у ширем смислу или тестови личности).

Тестови могућности могу се поделити на:

- *тестове остварења* (тестови постигнућа, тестови знања и умења)
- *тестове способности*: сензорних способности, менталних способности (тестови опште менталне способности или тестови интелигенције и тестови посебних специјалних менталних способности) и тестове моторних способности.

Нагласимо да тестове остварења није увек лако разграничити од тестова могућности, пошто многе могућности зависе, како од наслеђа и општег животног искуства, тако и непосредно од учења (Anastasi, 1984; Баковљев, 1997). У основи, када се прави разграничење, тестови остварења реферишу на реализовано учење, а тестови способности на предуслове за успех у учењу (Баковљев, 1997).

Када је реч о тестовима личности, међу њима разликујемо: социометријски тест, инвентар личности и пројективне тестове. Тестови личности су само условно тестови. Они се обично не састоје од задатака, већ од низа питања на које је потребно да испитаник одговори што искреније. Њихове метријске карактеристике нису тако повољне као метријске карактеристике тестова постигнућа. Ипак, тестовима личности мере се својства која је тешко поуздано испитати неком другом техником. Таква својства су, на пример, емоције, црте карактера и темперамент (Мужић, 1986; Баковљев, 1997; Мужич, 2004).

Тестови знања се могу поделити и на *тестове нивоа* и на *тестове брзине* (Баковљев, 1997). Тестовима нивоа се настоји установити ниво мереног својства, па је у првом плану тачност решавања тестовских задатака. Код тестова брзине главно питање је колико брзо испитаник може да реши одговарајући задатак (тест брзине читања, писања, рачунања, стенографисања итд.). При тестирању тестом брзине, време решавања теста је ограничено. Што се, пак, тиче тестирања тестом нивоа, теоријски је сваком испитанику остављено неограничено време за решавање, али се у пракси поступа другачије. Наиме, када 90-95% испитаника преда тест, замоле се и остали испитаници да то учине (Мужић, 1986; Баковљев, 1997; Мужич, 2004).

Тестови се деле и на *нормативне* и *критеријумске*. Код нормативних тестова, вредност појединачног резултата се утврђује у односу на норму популације којој је тест намењен. Код критеријумских тестова тестирање се спроводи са циљем утврђивања да ли је испитаник остварио одређени образовни циљ (Van der Linden, 1982; Vansickle, 2008).

Резултати нормативног теста омогућавају да се изведе закључак као што је следећи: „Јелена се налази на 93. перцентилу стандардизованог теста знања“. Резултати критеријумског теста омогућавају да се изведе закључак налик следећем: „Петров резултат говори о томе да је он савладао 85% од укупно 500 страних речи“ (Porham, 2003). Критеријумски тестови у односу на нормативне обезбеђују податке који се непосредније могу применити у

ситуацији доношења различитих одлука у настави и омогућавају прецизније мерење ученичког знања (Rudner & Schafer, 2002).

Табела 1. илуструје разлику између критеријумских и нормативних тестова знања када је реч о циљевима примене, одређивању метријских карактеристика теста и приказу резултата мерења.

Табела 1. Поређење нормативних и критеријумских тестова знања¹

	Нормативни тестови знања	Критеријумски тестови знања
Шта се испитује тестом?	<i>колико је ученик научио у односу на друге ученике</i>	<i>које наставне циљеве и у којој мери су ученици достигли</i>
Каква је природа критеријума?	<i>критеријуми су релативни</i>	<i>критеријуми су апсолутни</i>
Када се одређују критеријуми постигнућа?	<i>после пробног испитивања</i>	<i>пре израде и употребе теста</i>
Каква је дискриминативност теста?	<i>дискриминативност се односи на утврђивање индивидуалних разлика у постигнућу ученика: дискриминативни задаци праве разлику између ученика који су овладали градивом у односу на оне који то нису</i>	<i>дискриминативност се односи на савладаност градива: најдискриминативнији су задаци које пре обраде градива нико није решавао успешно, а после обраде решавају сви ученици</i>
Каква дистрибуција знања се очекује?	<i>претпоставља се нормална дистрибуција знања</i>	<i>полази се од очекивања да ће већина ученика постићи одређене васпитно-образовне циљеве</i>
Како се утврђује валидност теста?	<i>поређењем са оценама које даје наставник, али и поређењем са садржајем предмета</i>	<i>испитује се колико су задаци на тесту репрезентативни за градиво и колико одговарају образовним циљевима</i>
Како се исказују резултати?	<i>квантитативно, израчунавањем оствареног скорa на тесту</i>	<i>квантитативно, изражава се проценат градива које је ученик савладао.</i>

¹ Адаптирано према: Бјекић и Папић, 2006

Полазна разлика између нормативних и критеријумских тестова, као што смо видели, јесте њихова намена. Сходно намени, критеријумски тестови су краћи и обухватају мањи обим градива (Taylor, Greer & Mussio, 1978). Додатно, код нормативних тестова тежина и дискриминативност ајтема играју важну улогу у избору ајтема за коначну верзију теста. Код критеријумских тестова, тежина и дискриминативност ајтема немају такав значај. Из финалне верзије критеријумског теста искључују се само ајтеми који озбиљно нарушавају спецификацију садржаја. Недовољно дискриминативни ајтеми могу остати саставни део критеријумског теста уколико се односе на важан део градива (Gipps, 2003).

Предуслов израде критеријумских тестова јесу унапред формулисани васпитно-образовни циљеви наставе или одређеног наставног предмета. Циљевима се прецизира до ког нивоа је потребно да ученици савладају градиво, како би им се пружила повратна информација о остварености дефинисаног циља (Бјекић и Папић, 2006). Поред формулисања васпитно-образовних циљева, изради критеријумских тестова претходи дефинисање граничних поена на основу којих се процењује да ли је испитаник „положио“ или не. Одређивање граничних поена, у основи озбиљно методолошко питање, најчешће је резултат консензуса међу истраживачима и практичарима. Граница успешности се обично поставља на 70-80% тачних одговора испитаника (Вучић, 1979).

Заједничке карактеристике нормативних и критеријумских тестова знања су да се обе врсте тестова могу стандардизовати, користе исте врсте задатака и слична упутства за реализацију (Гојков, 2012).

Поједине класификације тестова знања полазе од специфичности примене тестова у пракси. Тако, у школској пракси се често указује на разлику између *сумативних* и *формативних тестова*. Њихово разликовање своди се на комбинацију више критеријума: циља примене, времена задавања и форме саопштавања резулата (*National Testing of Pupils in Europe: Objectives, Organisation and Use of Results*, 2009). У том смислу, сумативним тестовима се сумирају образовна постигнућа ученика на крају одређеног периода учења или циклуса

школовања. Резултат сумативног теста се изражава квантитативно, најчешће оценом (*Guide to Assessment, 2015*), и омогућава наставнику да постигнуће појединачног ученика посматра у односу на постигнуће одељења или генерације ученика у школи. Резултати сумативних тестова могу имати пресудан значај и за наставак школовања ученика (*Vansickle, 2008*). У датим околностима, резултат који ученик оствари на тесту посматра се у односу на дефинисане норме постигнућа за одговарајућу популацију ученика.

Циљ задавања формативних тестова, јесте унапређивање процеса учења код ученика. Могу се задавати више пута у току школске године, пре или у току процеса учења. Резултат остварен на формативном тесту обично се изражава квалитативно (*National Testing of Pupils in Europe: Objectives, Organisation and Use of Results, 2009*). За разлику од сумативних тестова, формативни тестови омогућавају да се постигнуће ученика континуирано прати и да се идентификују специфичне потребе ученика (*Vansickle, 2008*). Формативни тестови као мерило постављају оствареност одређеног образовног циља.

Према намени теста знања, Бјекић и Папић (2006), консултујући класификацију Андриловића (1988) разликују:

- *дијагностичке тестове* којима се утврђује присутност мереног својства код испитаника
- *прогностичке тестове* који предвиђају будуће понашање и успешност појединца у обављању одређене активности и
- *истраживачке тестове* које примењују истраживачи у циљу проучавања одређених карактеристика испитаника и њихове повезаности са другим карактеристикама (на пример, повезаност знања и појединих личних карактеристика).

У извесној мери, у датој класификацији нејасна је позиција истраживачких тестова, јер се они могу свести на дијагностичке или прогностичке чији се резултати могу касније доводити у вези са другим својствима испитаника.

Гојков (2012) наведену класификацију тестова знања према намени, допуњује *ревизионим и тестовима инвентара знања*. Ревизиони тестови

проверавају знање ученика стечено у дужем временском периоду, а инвентарни проверавају знање стечено у краћем временском периоду.

У зависности од тога да ли су прошли поступак стандардизације, тестови знања се могу поделити на *стандардизоване* и *нестандардизоване* (Rudner & Schafer, 2002; *Assessment in the primary school curriculum: guidelines for schools*, 2007). Стандардизовани тест се добија баждарењем у ужем смислу те речи, то јест применом теста на репрезентативном узорку ради долажења до скала које омогућају упоређивање резултата које испитаници постижу с утврђеним стандардима. Небаждарени тест је нестандардизован. Под *формалним* тестом се подразумева тест који је баждарен у ширем смилу. То значи да су му проверене све метријске карактеристике (валидност, релијабилност, објективност и дискриминативност). Будући да је објективност једина метријска карактеристика којом *неформални* тестови располажу, сматрају се мање погодним за научни рад (Баковљев, 1997). Неформалним тестовима се најчешће означавају тестови које праве наставници (Џордан, 1966).

Када је реч о начину решавања задатака у тесту, тестови знања се деле на: *усмене тестове*, *тестове типа „папир и оловка“*, *тестове чина* (Rudner & Schafer, 2002). Усмени тест подразумева да испитивач питања поставља усмено, и да испитаник саопштава своје одговоре усменим путем, тако да се ова форма тестирања понекад означава формом „строго везаног интервјуа“ (Муџић, 1986: 298). Усмено задавање задатака се врши у одређеним околностима: у раду са испитаницима који не умеју да читају (Рајстон, Џастман и Робинс, 1966) или онда када се инсистира да сви испитаници приступе решавању сваког задатка (да се не би догодило да од решавања одређених задатака одустану услед пада мотивације) (Муџић, 1986). Тестови типа „папир и оловка“ подразумевају да испитаници читају задатке са папира, а своје одговоре уписују на исти лист или на папир за одговоре. Тестови чина подразумевају проверу способности извођења одређених радњи или активности (Баковљев, 1997).

У модерно доба, класификацији тестова знања према начину решавања задатака придружују се и различити тестови који се решавају посредством рачунара (Бјекић и Папић, 2006). У даљем тексту, сви тестови који се ослањају

на примену компјутерске технологије у процесу примене и оцењивања теста биће означени као „електронски“ тестови. Користићемо скраћенице СВТ и САТ, али је важно нагласити да није реч о синонимима. Линеаран рачунарски тест, или СВТ, обично представља само рачунарску верзију папирног теста (Маравић-Чисар, 2011), док је САТ поступак тестирања утемељен на принципима теорије ставског одговора.

Полазна разлика између теста „папир и оловка“ (у даљем тексту: *класичних* тестова) и електронског теста је та што решавање електронског теста претпоставља да се задатак чита са екрана уместо са папира, и да се одговори уписују коришћењем тастатуре и миша уместо папира и оловке (Ashton *et al.*, 2005; Weiss, 2011). Суштинска разлика између тестова „папир и оловка“ и електронских тестова је концепцијске природе, и односи се на начин креирања теста и реализације тестирања.

Наиме, класични тестови садрже низ задатака који су истоветни за све испитанике. „Добар“ класичан тест састављен је превасходно од задатака који обезбеђују максималну поузданост инструмента, а то су најчешће задаци средње тежине ($p = 0.50$) (Weiss, 2004). Решавање тако конципираног теста, превише је тешко за исподпросечне ученике, али је и превише лако за натпросечне ученике. Истовремено, решавање класичног теста увек је временски ограничено и свим ученицима је на располагању исти временски период за рад. Последично, код ученика који су несигурни и споро дају одговоре или су, напротив, брзи и склони грешкама, отежано је објективно процењивање знања.

За разлику од класичног концепта тестирања, САТ омогућава да се за сваког испитаника одабере низ задатака који најпрецизније мере жељено својство и да се испитанику прилагоди време решавања теста. Улогу испитивача овог пута преузима рачунар, који бира задатке тежином прилагођене успешности ученика у одговарању на неколико полазних задатака (Фајгел, 2004).

У основи електронских тестова налазе се сложени софтверски системи. Софтвер омогућава испитивачу да утврди карактеристике појединачних

задатака (на пример, индекс тежине и дискриминативности) и да састави тест са жељеним карактеристикама ајтема/задатака. Пре коначне примене теста, истраживач може да провери статистичке параметре за тест у целини. У модерно доба, претраживање банке ајтема у просеку траје неколико секунди, у зависности од величине базе. За неколико тренутака добија се запис свих карактеристика теста у табеларној или графичкој форми. Уколико су корекције потребне, истраживач може да их реализује додавањем нових ајтема или искључивањем постојећих (Weiss, 2011).

Електронска тестирања доносе низ погодности које могу наћи своје место и у наставном процесу. Примењени у школским околностима, компјутерски системи за израду тестова знања омогућавају састављање ајтема, формирање банке ајтема, дизајнирање теста, примену и анализу резултата добијених тестом. Додатно, примена система као што је CADATS (*computer assisted design, analysis and testing system*), омогућава поређење резултата које је конкретан ученик остварио са постигнућем одељења, школе или националне норме. Систем омогућава наставнику да за конкретног ученика добије податак о успешности решавања сваког појединачног задатка у тесту, као и да сазна конкретан одговор ученика на постављени задатак (He & Tumm, 2005). Решавање електронског теста омогућава да се утврди тачно време које је испитанику било потребно за решавање одговарајућег задатка. То се постиже мерењем времена протеклог од испитаниковог првог приступа задатку до давања коначног одговора (He & Tumm, 2005; Weiss, 2011).

Електронски тестови омогућили су мерење одређених варијабли које није било могуће мерити класичним тестовима. На пример, данас електронски тест може да обухвати задатке са графичким, аудио и видео записима, као и симулације проблемских ситуација (Weiss, 2011) чије решавање тестира примену одговарајућих знања и вештина.

У експерименталном истраживању Маравић-Чисар (2011) утврђено је да су испитаници у просеку успешније решавали рачунарски адаптивни тест знања него испитаници који су радили класичан тест знања. Решавање САТ временски краће траје, мада ауторка то објашњава немогућношћу испитаника

да се врате уназад и евентуално поправе решења. Две трећине испитаника је изјавило да полагање теста на рачунару представља пријатније искуство у односу на класичан тест, јер је мање стресно и задаци су прилагођени по тежини (искључено је задржавање на превише тешким задацима, односно околност да су испитанику сви задаци превише лаки).

Сумарно, предности компјутерски адаптивног тестирања у односу на традиционално конципирано тестирање огледају се у следећем (Фајгел, 2003):

- не захтева се штампање тестова
- резултати остварени на тесту су одмах доступни
- не захтева се штампање тестова
- испитаник решава тест својим темпом, чиме се смањује почетна анксиозност
- у односу на класичне тестове, САТ може да скрати време тестирања за 50%, јер је потребно значајно мање задатака да се постигне прихватљива тачност
- САТ пружа поуздане скорове за већи распон способности испитаника, док су класични тестови најтачнији код просечних испитаника
- калибрисање базе задатака може се вршити непрестано, након сваког појединачног тестирања.

САТ има и одређене недостатке у односу на традиционално конципирано тестирање који су утврђени низом емпиријских студија (Фајгел, 2003). Недостаци су следећи:

- испитаник не види задатке који следе, што отежава планирање времена за решавање одговарајућих задатака и утврђивање редоследа решавања
- у решавању математичких задатака за сада нема могућности за приказивањем поступка решавања
- САТ не омогућава враћање уназад и исправљање датих решења
- испитаници који су вешти у коришћењу рачунара у предности су у односу на испитанике са мање искуства, без обзира што се софтвер за САТ конструише тако да буде максимално једноставан за коришћење

- у примени САТ испитаници различитог пола и старости у различитој мери су успешни – испитанице и старији испитаници генерално су мање успешни у САТ тестирању
- неке вештине не могу се испитати применом компјутерски адаптивног тестирања
- у зависности од хардвера и капацитета радних станица, графички и тродимензионални прикази у задацима нису увек довољно јасни да би могли да се користе у процесу тестирања
- постоје хардверска, просторна и материјална ограничења за тестирање већих група испитаника.

Интернет омогућава ширење и развијање мреже електронских тестова и њихово решавање *online*. Међутим, тестирање знања посредством интернета доноси са собом одређене ризике, пре свега по стандардизацију поступка тестирања. Наиме, тестовима знања који су доступни путем интернета испитаници могу приступити у различитим ситуацијама (када нису сами, када пред собом имају помоћна средства за решавање теста), а приказ садржаја теста може бити измењен у зависности од резолуције екрана на коме испитаник ради (Weiss, 2011). Такође, брзину интернета није могуће уједначити у свим околностима, тако да процес тестирања може различито да траје за различите испитанике, што ствара погрешну слику о брзини одговарања испитаника на задатке у тесту (Weiss, 2011).

Према броју испитаника, тестови знања се могу поделити на *индивидуалне* и *групне* (Баковљев, 1997). Усмени тестови и тестови радњи се обично реализују индивидуално, док тестови „папир и оловка“ подразумевају истовремени рад групе испитаника на истом тесту. Реализацију тестова подржаних компјутерском технологијом могуће је остварити индивидуално и на нивоу групе (Баковљев, 1997).

Две теорије тестирања: класична теорија и теорија ставског одговора

У психометрији постоје различите теорије тестирања. Период до половине XX века обележила је доминација *класичне теорије тестирања* (*Classical Test Theory*) која се скраћено означава као „СТТ“. Класична теорија тестирања подржава стандардизоване тестове типа „папир и оловка“, намењене дефинисаној популацији испитаника. Основе класичне теорије тестирања поставио је Спирман у првој декади XX века (Weiss & Davison, 1981; Фајгељ 2004). На прелазу са педесетих на шездесете године XX века конституише се нова теорија тестирања, позната под називом *теорија ставског одговора* (*Item Response Theory*), или скраћено „IRT²“. Теорија ставског одговора утемељена је радом Терстона, а касније радом Мосијеа, Гутмана и Лазарсфелда (Mosier, Guttman, Lazarsfeld). Научници најзаслужнији за популаризацију IRT у свету су Лавли, Лорд и Рајт (Lawley, Lord, Wright). Сматра се да је Лорд заслужан за популаризацију IRT у области образовања (Hambelton & Cook, 1977). Рајт је допринео промоцији најутицајнијег IRT модела у изради тестова постигнућа – Рашовог модела. Данас, IRT моделе примењује преко 500 организација за развијање тестова знања широм света (Natarajan, 2009).

И класична теорија и IRT налазе примену у домену мерења латентних својстава. Ипак, теорије полазе од различитих претпоставки када је у питању природа својства које се мери, као и начин на који испитаник одговора на ајтеме/задатке у тесту. Класична теорија посматра скор који испитаник освари на тесту као „прави“ скор коме се додаје одговарајућа вредност грешке мерења. Грешка мерења варира у свакој ситуацији тестирања, у зависности од узорка испитаника, а самим тим варира и скор који испитаник остварује у поновљеним мерењима. Стварни скор се сматра теоријски непроменљивим у свакој ситуацији тестирања истим тестом. Када се исто својство мери неким

² Фајгељ (2003; 2004) скраћеницу IRT преводи као ТСО у значењу „теорија ставског одговора“, али наглашава да се епитет „ставски“ не односи на став него на ставку. У домаћој литератури може се срести и скраћеница ТАО, у значењу „теорија ајтемског одговора“, али се готово све три скраћенице још увек равноправно користе (Фајгељ, 2003).

другим тестом, не очекује се исти стварни скор, већ другачији. IRT полази од претпоставке да сваки испитаник има свој стварни скор („тета“) на континууму латентног својства и да тај скор неће значајно варирати који год тест да се примени за мерење одговарајућег својства. IRT пружа „процену својства“ на основу обрасца који се читава у начину одговарања испитаника на сетове ајтема. Дакле, СТТ скор је у потпуности одређен конкретним сетом ајтема које испитаник решава, док је IRT скор од њих потпуно независан. Самим тим, интерпретација СТТ скорa је увек одређена контекстом – ајтемима и узорком испитаника. Интерпретација IRT скорa, пак, није зависна од контекста (Di Nisio, 2010; Sharkness & DeAngelo, 2011). У случају класичне теорије, стандардна грешка мерења је константна за све испитанике, док у случају IRT модела мерења она може да варира при одговору испитаника на сваки појединачни ајтем (Hambelton & Cook, 1977; Sharkness & DeAngelo, 2011). Такође, у случају класичне теорије, тежина ајтема се одређује на основу процента испитаника који су тачно одговорили на задати ајтем, чинећи је детерминисаном нивоом постигнућа узорка испитаника. У случају IRT, тежина ајтема се израчунава унапред, независно од испитаника и њихових одговора на дати ајтем. Класична теорија тестирања искључује погађање тачних одговора као чинилац који значајно одређује постигнуће, док IRT могућност погађања израчунава у готово свакој ситуацији мерења (Rivera, 2007). У моделима мерења који одговарају класичној теорији тестирања, однос између квалитета дистрактора у задацима бирања одговора и резултата на тесту углавном се не испитује, за разлику од IRT модела у којима се квалитет дистрактора сматра важним за постигнуће испитаника на тесту (Thissen, Steinberg & Fitzpatrick, 1989; Rivera, 2007).

Класична теорија тестирања подржава конструисање базе ајтема, али не инсистира на њеном развијању док то не постане неопходно. Уколико ајтеми у тесту који мери одређену способност одражавају све оно што се о тој способности зна на одређеном стадијуму развоја науке, и уколико ајтеми имају задовољавајуће статистичке параметре за популацију испитаника којима су намењени – нема потребе да буду замењени. Тек када се појави нова теорија

дате способности коју тест мери, могло би се говорити о потреби да се и ајтеми у тесту измене. Без обзира на извесну „нефлексибилост“ класичне теорије, она ипак представља кохерентан систем, добро развијен и статистички коректан. У ситуацији када је потребно конструисати јединствен тест намењен одређеној популацији испитаника и када је довољно утврдити распон скорова на нивоу популације, дизајнирање теста у духу класичне теорије било би сасвим оправдан избор (Van der Linden, 2005). Међутим, за прецизнију процену присутности одговарајућег својства код испитаника примена IRT модела мерења представља бољи избор (Natarajan, 2009).

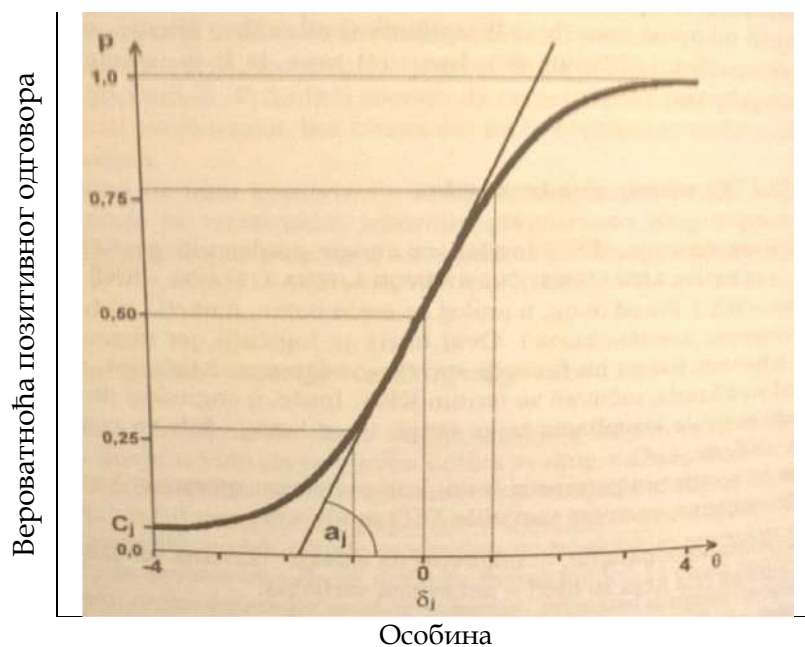
Ван дер Линден (2005) указује на четири начина бирања ајтема у IRT моделима мерења. Први начин избора ајтема подразумева да се ајтеми насумично узимају из банке ајтема како би се добио тест одређене дужине. Што је тест дужи, то се резултати сматрају поузданијим. Други начин подразумева израчунавање грешке мерења после сваког испитаниковог одговора. Бирање ајтема из банке се зауставља онда када се постигне одређени, унапред дефинисани ниво грешке мерења. Трећи начин се назива оптималним дизајном теста и подразумева да се унапред дефинише циљ мерења, а потом се применом одговарајућег софтвера из банке бирају ајтеми одређене садржинске спецификације (која, заправо, одговара циљу мерења). Четврти начин је дизајнирање адаптивног теста. Идеја је да се процена испитаниковог мереног својства, на пример одређене способности, остварује након сваког датог одговора. Пошто се процена направи, бира се следећи ајтем који одговара нешто вишем нивоу способности у односу на онај који је претходно утврђен. Уколико испитаник на изабрани ајтем одговори погрешно, бира се следећи ајтем који ће одговарати нешто нижем нивоу способности. Како тестирање одмиче, процена нивоа испитаникове способности све је прецизнија (Van der Linden, 2005). На описани начин, поступак тестирања постаје ефикаснији – потребно је 40% мање ајтема него у ситуацији када група испитаника решава јединствени тест, да би се постигао исти ниво прецизности мерења (Hattie, Jaeger & Bond, 1999).

Теорија ставског одговора убрзала је развој и имплементацију адаптивног тестирања у пракси. Реч је о начину тестирања који се развијао независно од теорије мерења (Weiss & Davison, 1981), али су IRT и високо софистицирана технологија израде рачунара допринели да адаптивно тестирање постане део наше свакодневице. Данас је компјутерски адаптивно тестирање (CAT) имплементирано у многе програме тестирања широм света (Van der Linden, 2005). Неке од основних предности компјутерски адаптивног тестирања су: ефикаснија процена мереног својства, брза повратна информација о постигнућу на тесту, могућност да се више својстава мери паралелно, могућност да се у тест уврсте и задаци испуњавања налога и друге алтернативне форме ајтема (задаци сређивања, рад са текстом, аудитивни задаци и тако даље), реализација тестирања у различито време, приступ тестовима путем различитих сајтова на интернету (Hattie, Jaeger & Bond, 1999). Највећи недостатак компјутерски адаптивног тестирања је тај што је за његову успешну реализацију неопходно направити велику банку ајтема. Практично, ниједан испитаник не би требало да решава исти сет ајтема, како тест не би био „откривен“.

Графички приказ IRT модела добија се исцртавањем функције која се назива *карактеристична крива ајтема (Item Characteristic Curve)* или скраћено ICC. Карактеристична крива ајтема има форму графика (Wainer, 1989). Два својства ICC-а која дефинишу све конструкте обухваћене IRT моделом су тежина ајтема и дискриминативност (слика 1). У IRT тежина ајтема и мерено својство припадају истој скали (Baker, 2001; Natarajan, 2009; Di Nisio, 2010), односно читавају се са апсцисе графика. Ово је јединствена карактеристика која омогућава директну везу између ајтема у тесту и стварних скорова испитаника. Дискриминативност ајтема је својство на основу кога се прави разлика између испитаника у степену у коме поседују одговарајуће својство (Baker, 2001; Natarajan, 2009).

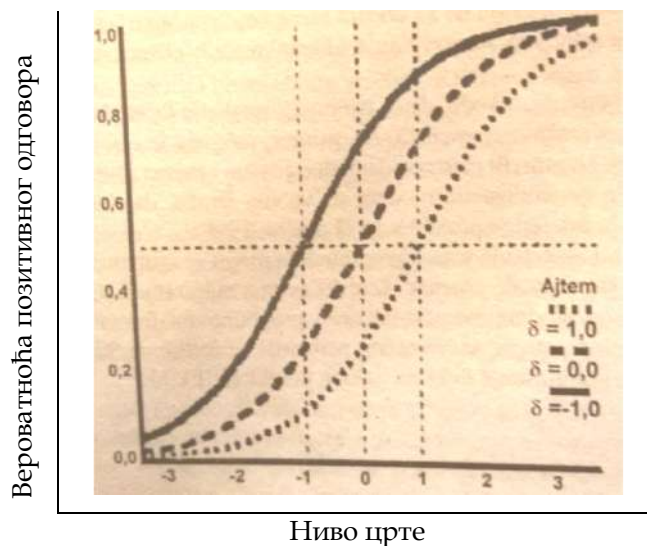
На слици 1. приказана је карактеристична крива ајтема за најједноставнији случај – ајтем се дихотомно скорује, при чему се тачан одговор означава бројем 1, а нетачан са 0. На слици су приказани следећи параметри:

ајтем (j), мерено својство или особина (θ), тежина ајтема (δ_j), нагиб криве (a_j) и доња асимптота криве (C_j). Тежина ајтема на приказаном примеру тумачи се као 50% вероватноће пружања тачног одговора. Приказани ајтем се може означити као умерено лак. Нагиб криве одражава дискриминативност ајтема и, у датом примеру, она износи 0,90 (умерена дискриминативност). Доња асимптота криве означава погађање и у датом примеру вероватноћа пружања тачног одговора погађањем износи 10%. Крећући се од координатног почетка на десно, дуж скале особине (θ), можемо закључити да је у случају испитаника са ниским нивоом мереног својства мала вероватноћа да ће пружити тачан одговор (Фајгељ, 2003).



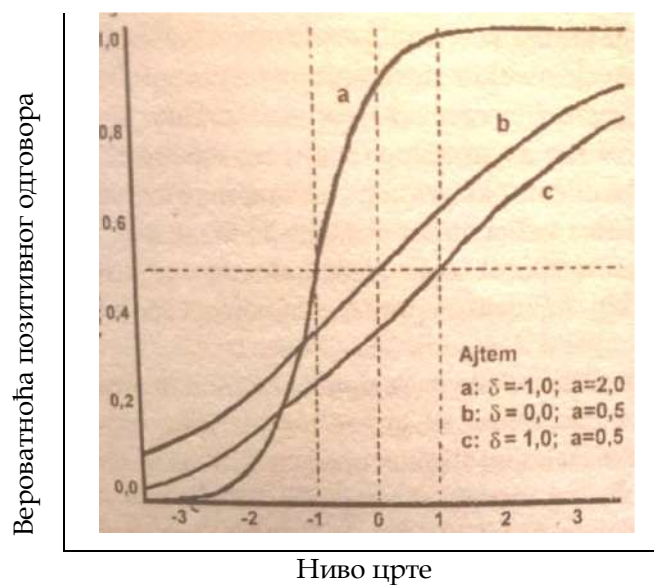
Слика 1. Карактеристична крива ајтема (ICC)
(преузето из: Фајгељ, 2003: 176)

ICC се појављује у три модела. То су Рашов, Бирнбаумов и Лордов модел. Рашов модел за појединачне ајтеме израчунава само један параметар, а то је тежина, док се дискриминативност означава константом (Natarajan, 2009). Слика 2. приказује карактеристичне криве три дихотомне ставке према Рашовом моделу. Ајтеми имају различиту тежину, при чему је најтежи ајтем највише удаљен од јединичне вероватноће пружања тачног одговора.



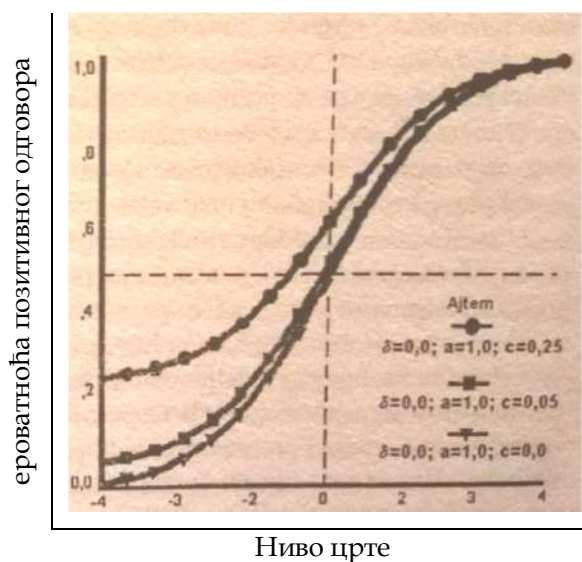
Слика 2. Рашов једнопараметарски модел (1PL)
(преузето из: Фајгељ, 2003: 193)

Бирнбаумов модел дефинише два параметра, тежину ајтема и дискриминативност (Natarajan, 2009). Слика 3. приказује карактеристичне криве три дихотомне ставке према Бирнбаумовом моделу. Три приказана ајтема имају различите тежине, од којих два имају истоветан нагиб, а најтежи ајтем је највише удаљен од јединичне вероватноће пружања тачног одговора.



Слика 3. Бирнбаумов двопараметарски модел (2PL)
позитивног одговора на дихотомни ајтем
(преузето из: Фајгељ, 2003: 197)

Лордов модел дефинише три параметра – тежину, дискриминативност ајтема и „погађање“ тачног одговора („погађање“ се дефинише као константа) (Natarajan, 2009). Слика 4. приказује карактеристичне криве три дихотомне ставке према Лордовом моделу. Сва три ајтема имају исте тежине и нагиб, али им је различито процењена вероватноћа погађања тачног одговора.



Слика 4. Лордов тропараметарски модел (3PL) позитивног одговора на дихотомни ајтем (преузето из: Фајгел, 2003: 197)

Избор модела остварује се применом сложених математичких прорачуна уз помоћ рачунарске технологије (Hambelton & Cook, 1977; Baker, 2001). На основу прорачуна, истраживач се опредељује за онај модел који највише одговара („фитује“) моделу одговора испитаника (Kang, 2006). Лордов модел је посебно погодан за састављање тестова са задацима вишеструког избора, јер решава проблем погађања одговора (Фајгел, 2003).

Поређењем двеју теорија тестирања бавили су се многи истраживачи. Резултати испитивања нису једнозначни. Тако, истраживање Шаркнес и Дианђело (Sharkness & DeAngelo, 2011) даје предност IRT моделу у односу на СТТ модел када је у питању прецизност мерења. Слично, Фајгел и Косановић (2001) утврдили су да IRT и СТТ модели различито оцењују квалитет ајтема у

тесту, дајући предност IRT моделима. Са друге стране, испитивање које спровео је Фан (Fan, 1998), показало је да IRT и СТТ модел израчунавају сличне вредности параметара за ајтеме и испитанике.

Халадина и Роид (Haladyna & Roid, 1981) су поредили класичну и теорију ставског одговора са становишта конструисања критеријумског теста знања. Класична теорија тестирања у конструкцији теста ослања се на насумични избор ајтема из претходно формиране базе задатака, док теорија ставског одговора наглашава да је неопходно тежину ајтема усагласити са нивоом постигнућа конкретног ученика. Резултати су показали да је на тестовима који су састављени уз примену Рашовог модела грешка мерења била значајно мања него у случају тестова који су састављени насумичним узорковањем ајтема.

Иако би се очекивало да су мерења спроведена применом IRT модела поузданија, с обзиром на већу математичку тачност модела, ипак се не може оспорити да модели класичне теорије тестирања и даље имају широку примену у пракси. Део образложења се сигурно може потражити у већој доступности класичних модела. Исто тако, примена IRT модела претпоставља сложене кораке као што је израда банке ајтема и рачунарски адаптивно тестирање који стручној јавности још увек нису довољно познати (Фајгел, 2001).

Последњих година, појавио се и трећи модел мерења латентних својстава, који је базиран на когнитивним процесима (*cognitive-processing-based model*; Hattie, Jaeger & Bond, 1999). Потреба за конструисањем новог модела потиче отуда што ни IRT и СТТ не моделују процесе који стоје у основи испитаниковог понашања, али не кажу много ни о природи самог мереног својства. Математички модели на које се IRT и СТТ ослањају, заправо занемарују процесну основу способности и вештина, и оцењује се да ће се у будућности теорије тестирања оријентисати ка моделима који полазе од такве основе (Hattie, Jaeger & Bond, 1999).

КОНСТРУКЦИЈА ТЕСТА ЗНАЊА

Фазе у конструкцији теста знања

У литератури се срећу два приступа конструкцији тестова: *конвенционални* приступ и приступ утемељен на *доказаној повезаности ајтема и мереног својства* (*Evidence-Centered Assessment Design* или скраћено: *ECD*). Конвенционални приступ је развијен на основама класичне теорије тестирања, док је *ECD* приступ развијен на основама теорије ставског одговора. Реч је о приступима који нису супротстављени један другом. Наиме, *ECD* приступ као историјски каснији, развијен је на основама конвенционалног приступа у циљу његовог унапређивања. Основна тежња била је да се јаснијом операционализацијом задатака у процесу конструкције теста побољша ваљаност тестова као мерних инструмената.

Конструкција теста према конвенционалном приступу обухвата низ фаза које су хронолошки уређене. Класификација Буквића (1996) представљала би пример класификације општијег типа. Према становишту аутора, израда сваког теста обухвата три основне фазе. Прва фаза обухвата израду пробне верзије теста, састављање упутства за задавање теста, упутства за оцењивање и бодовање одговора и пробно испитивање на једном мањем узорку испитаника. У овој фази, истраживач се не ослања превише на статистичку анализу, јер се овог пута испитује само генерална примењивост теста. Утврђује се да ли је упутство преопширно, или можда недовољно јасно, да ли установљени систем бодовања обухвата све могуће одговоре испитаника, да ли се појављују ајтеми који су двосмислени и нејасни и томе слично. Израда друге верзије теста детерминисана је врстом и обимом корекција које је потребно извршити на прелиминарној верзији теста. Примена друге верзије теста на великом узорку испитаника омогућава да се провере метријске карактеристике задатака и да се

изврши избор задатака који ће бити уврштени у коначну верзију теста. Тиме се окончава друга фаза израде теста знања (Буквић, 1996).

Трећа фаза започиње израдом финалне верзије теста. Уобичајено је да се финална верзија теста испитује неколико пута на већим узорцима како би се провериле све метријске карактеристике инструмента. Ако тест има специфичну намену, у овој фази неопходно је израдити норме постигнућа за популацију којој је тест намењен. Рад на тесту се завршава писањем елабората о његовој конструкцији и израдом приручника за примену теста, мада се ниједна верзија теста не сматра коначном докле год има простора за његово унапређивање. С обзиром на вишеетапно проверавање инструмента које доводи до тога да се одређени ајтеми из теста морају искључити, препоручује се да се за прелиминарну верзију теста састави 50 до 100% више ајтема од броја који је предвиђен за коначну верзију (Буквић, 1996).

Банђур и Поткоњак (1999) наведене фазе рапосређују на нешто другачији начин, конституишући низ од укупно четири фазе. Према становишту аутора, прву фазу чини састављање задатака, различите форме и садржаја према унапред одређеној структури. У другој фази се елиминишу задаци који се, на основу првог прегледа, истраживачу учине нејасним, двосмисленим или одступају од садржаја теста. Трећа фаза се односи на пробно испитивање теста. У завршној фази се утврђују метријске карактеристике теста, одређује оптимално време за решавање теста, коригују упутство и кључ за оцењивање, утврђује редослед задатака.

Полазне три фазе у конструкцији теста које описују Банђур и Поткоњак, истоветне су фазама које приказује Мужих (1986; 2004). Приказани опис четврте фазе, код Мужиха је употпуњен анализом тежине и дискриминативности сваког од задатака у тесту. Такође, аутор предвиђа могућност да се након пробног испитивања, тест провери још неколико пута како би коначна форма теста имала што повољније метријске карактеристике. Мужих описује и пету фазу конструкције теста, која подразумева да се коначни облик теста примени на репрезентативном узорку испитаника. На основу

добијених података врши се баждарење инструмента и одређују коначне метријске карактеристике теста.

Конструкција теста знања према становишту Баковљева (1997) обухвата исте кораке као и Мужичева класификација. Разлика постоји само у подели четврте и пете фазе на две, чиме се добија низ од укупно 7 фаза конструкције теста. Конкретно, провера метријских карактеристика теста након пробног испитивања и елиминисање задатака који не покажу задовољавајуће карактеристике, код Баковљева чине две одвојене фазе. Аналогно томе, примена теста на репрезентативном узорку и баждарење теста са утврђивањем метријских карактеристика финалне верзије, такође представљају одвојене фазе конструкције теста код Баковљева.

Богавац и сарадници (1972) утврдили су посебно детаљну листу фаза у изради теста знања. У методолошком смислу, фазе су садржински добро дефинисане и пажљиво анализирају сваки од корака које истраживач или наставник предузима у процесу израде теста знања. Фазе су следеће:

- дефинисање циља теста
- одређивање садржаја теста
- одређивање дужине теста
- одређивање тежине задатака у тесту
- избор типа задатака према форми
- израда задатака
- утврђивање редоследа задатака
- предвиђање начина обраде резултата теста
- вредновање задатака
- предвиђање коректуре за погађање одговора
- пробно испитивање и баждарење теста
- састављање дефинитивне форме теста.

Значајан број фаза (укупно 6) у наведеној класификацији односи се на разматрање питања које се тичу садржаја, типова задатака и њиховог редоследа у тесту знања. Једна од фаза тиче се непосредне израде задатака, а укупно три фазе се односе на оцењивање теста. У вези са провером квалитета теста,

прецизирана је само једна фаза. Наведене карактеристике дату класификацију чине другачијом од претходно анализираних, али је више приближавају ономе што су захтеви конструкције теста знања у образовној пракси.

Крокер и Алцајна (Crocker & Algina, 1986) такође одређење циља означавају полазном фазом у изради теста. У следећем кораку, према становишту аутора, неопходно је дефинисати конструкт који се мери или јасно прецизирати домен садржаја који ће бити обухваћени задацима у тесту (фаза 2). Следећа фаза (фаза 3) је означена као израда спецификације теста. Израда задатака представља четврту фазу код Крокера и Алцајне. Иза тога, следе фазе ревизије задатака (фаза 5), потом фаза пробног тестирања (фаза 6) и тестирања на репрезентативном узорку испитаника (фаза 7). Наредне фазе у класификацији Крокера и Алцајне односе се на ајтем анализу (фаза 8), утврђивање коначних вредности метријских карактеристика теста (фаза 9) и израду упутстава за реализаторе тестирања и интерпретацију резултата остварених на тесту (фаза 10).

Крокер и Алцајна наглашавају да, када је реч о тестовима знања, одређивање конструкта који стоји у основи теста може да се сведе на препознавање образовних циљева којима се тежи у процесу наставе (у зависности од тога да ли се припрема критеријумски или нормативни тест). Такође, за тестове знања посебно је важна израда спецификације теста којом се прецизирају садржаји задатака и когнитивни процеси или операције које испитаник мора применити како би решио задатак.

Класификација фаза израде теста према становишту Изарда (Izard, 2005) обухвата 11 фаза, при чему се прве две фазе односе на доношење одлуке да се приступи тестирању (фаза 1) и доношење одлуке да се ангажују сви ресурси неопходни за реализацију тестирања (фаза 2). Овог пута, одређивање садржаја теста посматра се одвојено од израде спецификације теста, мада сачињавају јединствену фазу конструкције теста знања (фаза 3). Израда задатака представља четврту фазу, а потом следи фаза ревизије задатака (фаза 5). Шеста фаза је вредновање задатака, а потом следи припрема теста за пробно испитивање (фаза 7) и реализација пилот истраживања (фаза 8). Девета фаза је

фаза ревизије теста, а десета се односи на доношење одлука у вези са заменом или искључивањем појединих задатака из коначне форме теста. Једанаеста фаза подразумева израду финалне верзије теста.

Изард (2005) спецификацију теста одређује доста широко. Према становишту аутора спецификација теста укључује: назив теста, циљ примене теста, аспекте курикулума на које се тест односи, прецизирање којим ученицима је тест намењен, типове задатака које тест обухвата, заступљеност задатака одређеног типа у тесту, начине коришћења резултата добијених тестом, услове под којима ће тестирање бити реализовано (просторно-временски услови, особе задужене за задавање теста, особе задужене за оцењивање теста, могућност коришћења дигитрона или уџбеника у току решавања теста, прецизирање мера контроле недозвољених понашања у току тестирања).

Новина коју доноси Изардова класификација је укључивање процена о расположивим ресурсима за реализацију тестирања. Аутор наглашава да је у процесу тестирања питање „можемо ли да приуштимо средства за тестирање“ подједнако важно као и питање у вези са одређењем циља истраживања. Под ресурсима неопходним за реализацију тестирања аутор подразумева ангажовање стручних лица и истраживача, набавку рачунара и софтвера за обраду података, опреме за умножавање и штампање материјала, папира и средстава за писање (Izard, 2005). Укључивање процена о расположивим ресурсима у процес конструкције теста може се довести у везу са реализацијом тестирања стандардизованим тестовима. Ипак, већина наведених питања неће бити релевантна за ситуацију када наставници самостално израђују тестове знања у циљу редовне провере знања ученика.

Наведене класификације су различитог степена општости. Укупан број фаза креће се у распону од 3 до 12. У основи, разлике у броју фаза не указују на препознавање квалитативно другачијих фаза у конструкцији теста – више је реч о другачијем организовању практично истоветних фаза. На пример, фаза израде спецификације теста код Крокера и Алцајне обухвата укупно 4 различите фазе у класификацији Богавца и сарадника (израду садржаја

задатака, одређивање дужине теста, тежине задатака и избор форме задатака). Са друге стране, провера метријских карактеристика теста и елиминисање задатака који не покажу задовољавајуће карактеристике, код Баковљева чине две одвојене фазе, док су код Мужиха то два аспекта јединствене фазе конструкције теста.

Независно од општости класификације, одређене фазе су незаобилазни елемент сваке од њих. Свака конструкција теста садржи фазу састављања задатака, и један или више начина провере квалитета теста. Такође, поједине фазе препознајемо само код одређених аутора. На пример, на израду приручника за примену теста и интерпретацију резултата тестирања указују само Крокер и Алцајна и Буквић, док на фазу доношења одлуке о тестирању и ангажовању неопходних ресурса за тестирање указује само Изард.

Активности кроз које пролази конвенционални процес заступљене су и у ECD приступу, али им се придружују додатне активности усмерене ка развијању модела задатака. Модели задатака израђују се на основу анализе способности и вештина које ученици демонстрирају у процесу решавања задатака (Hendrickson, Huff & Luecht, 2010).

ECD приступ одговара на три питања (Razzouk, 2011):

- које својство/која својства се процењују – модел компетенција (*competency model*)
- која понашања испитаника представљају индикаторе мереног својства на одговарајућем нивоу – модел доказа (*evidence model*)
- који задаци/ситуације изазивају понашања која имају функцију индикатора – модел задатака (*task/activities model*).

Може се приметити да су то питања на која одговара и конвенционални приступ. Разлика је у томе што ECD приступ обезбеђује да се образложи како испољено понашање испитаника (одговори на питања или решења задатака у тесту) служи као доказ за тврдње које изводимо на основу резултата тестирања (Mislevy & Haertel, 2006)³. Да би то било могуће, ECD приступ удружује сва

³ Извођење закључака у вези са мереним својством у суштини се своди на аргументовање. ECD приступ користи термин „аргументовање“ у значењу које је дефинисао Месик (Mislevy &

расположива сазнања из области когнитивне психологије, психометрије, математике, информатике, статистике.

Дизајнирање теста у ECD приступу остварује се на неколико нивоа, односно обухвата неколико „слојева“ (*layers*). Сваки ниво/слој има своју улогу у процесу мерења, одређен је кључним појмовима, ентитетима и алатима који помажу у остваривању задате улоге (детаљнији приказ свих елемената дат је у: Mislevy & Haertel, 2006). Табела 2. приказује 5 нивоа дизајнирања теста у ECD приступу и опис њихових главних улога.

Табела 2. Нивои дизајнирања теста знања⁴

Ниво	Улога
Анализа домена	прикупљање информација о томе како се знање конструише, усваја, примењује и преноси
Моделовање домена	аргументација о постигнућу у наративној форми (засновано на информацијама из анализе домена)
Појмовни оквир	аргументација о постигнућу у структуралној форми; спецификација задатака у тесту; евалуација процедура; модели мерења
Имплементација	спровођење пробног тестирања; калибрација модела мерења
Финална реализација тестирања	координирање интеракције испитаника и задатака; скоровање; извештавање

У односу на конвенционални приступ, ECD приступ омогућава да се боље разумеју кораци које је неопходно предузети у процесу конструкције теста, као и то да читав процес конструкције буде ефикаснији (Mislevy & Haertel, 2006). Вероватно најкрупнија новина коју ECD приступ доноси јесте

Haertel, 2006). О Месиковом учењу нешто више речи ће бити у поглављу о метријским карактеристикама тестова знања.

⁴ Адаптирано према: Mislevy & Haertel, 2006: 8.

гарантовање везе између ајтема и мереног својства. Ипак, мерење одговарајућег својства у ECD приступу прецизно је у оној мери у којој су прецизна и наша сазнања о функционисању људске когниције и у мери у којој се когнитивни процеси могу представити постојећим математичким моделима. С обзиром на то да се научна сазнања континуирано унапређују, ECD платформа је подложна променама и налази се у сталном процесу развијања.

Приступи разради садржаја теста знања

У циљу разраде садржаја теста знања могу се користити различити оквири. Може се поћи од наставног програма, уџбеника, исхода учења, компетенција или стандарда дефинисаних за одговарајући наставни предмет. Избор једног од наведених оквира одражава се на процес конструкције теста знања. Ипак, разграничење наведених оквира није увек једнозначно. На пример, циљеве, задатке и исходе поједини аутори означавају синонимима (детаљније у: Деспотовић, 2010), понекад се истиче разлика на нивоу општости и мерљивости (Илић, 2012), а понекад се наглашава да имају исту основу, али различите функције (Davies, 1976).

Наставни програм као основа за разраду садржаја теста знања, обезбеђује увид у теме које ће бити обрађене у настави, циљ и задатке изучавања наставног предмета. У складу са дефинисаним темама, разрађује се садржај уџбеника. Садржај уџбеника може помоћи наставнику да издвоји делове градива које би тест требало да обухвати. Ипак, то не значи да ће се случајним избором реченица из уџбеника направити добар избор садржаја на основу кога би се конструисао тест. Формирање „узорка“ садржаја обично подразумева процену шта је од обухваћеног садржаја битно за наставак учења (Фајгел, 2005: 338). У извесној мери, процена је увек ослоњена на наставникову лични став који садржаји су важнији од других.

С обзиром на то да се циљ програма односи на изучавање целокупног градива једне школске године, формулација циља је углавном уопштена. Висока општост циљева одговара ономе што је њихова функција и не може се сматрати њиховим недостатком (Davies, 1976). Међутим, пошто нису довољно прецизни, циљеви у крајњој истанци нису ни мерљиви (Деспотовић, 2010). Са становишта конструкције теста то је битно ограничење. На пример, у формулацији циљева се повремено користе језичке конструкције које нису довољно директивне да би се на основу њих формулисали задаци у тесту знања („развијати критичко мишљење“ или „изградити добре грађане друштва“) (Ebel, 1979: 37).

Циљ наставног програма се може конкретизовати како бисмо знали који кораци су неопходни да бисмо се том циљу приближили. Конкретизација циља се остварује извођењем задатака образовног програма. Начелно, циљ се може конкретизовати извођењем општијих и оперативних задатака. Формулација општијих задатака своди се на извесну конкретизацију циља. На пример, циљ изучавања биологије може да буде да ученици разумеју утицај и ефекте утицаја организама на животну средину. Општији задатак, изведен на основу датог циља, може да гласи: „Ученици би требало да се упознају са различитим утицајима организама на животну средину“ (*Настава усмерена на исходе, компетенције и стандарде*, 2015: 91).

Формулација оперативних задатака је конкретнија. Њима се указује на знања и вештине које ће испитаник бити у стању да покаже по завршетку процеса учења (Davies, 1976). Оперативни задаци се у новијој литератури означавају као исходи учења (Илић, 2012; Кундачина, 2012). Образовни исходи пружају повратну информацију о постигнутим резултатима и њиховој подударности са оним што је намеравано (*Опште основе школског програма*, 2003). Исходи учења имају форму изјаве о томе шта се очекује да ученик зна или уме да уради након завршетка процеса учења (Kennedy, 2009).

У односу на наставни програм и уџбеник као оквире, у случају исхода учења, наставник има увид не само у релевантан садржај, већ и у видљива понашања која показују да је садржај усвојен. На пример, исходи за наставну

тему „Грађа људског тела - правила исхрана и последице неправилне исхране“, могу да гласе (*Настава усмерена на исходе, компетенције и стандарде*, 2015: 71):

Исход 1: „Ученик ће бити оспособљен да састави јеловник за правилну исхрану“.

Исход 2: „Ученик ће бити оспособљен да избегава ризична понашања која доводе до болести зависности и нарушавања здравља“.

Иако исходи јесу конкретнији од претходно наведених оквира за разраду садржаја теста, ослањање на исходе такође има одређена ограничења. Наиме, отворена су питања мерљивости исхода, фокусираности на очекивана понашања и питање контекстуализације исхода.

Као оквир за разраду садржаја теста исходи могу највише да помогну наставнику уколико су формулисани тако да упућују на мерљива својства. На пример, исход може да гласи: „Ученик дефинише појам фотосинтезе“ (*Настава усмерена на исходе, компетенције и стандарде*, 2015: 94) или „Ученик користи микроскоп у циљу разликовања грађе ткива“ (*Настава усмерена на исходе, компетенције и стандарде*, 2015: 96). Из наведених формулација могло би се закључити да је исход остварен уколико ученик уме да искаже дефиницију фотосинтезе, односно уколико може да употреби микроскоп како би утврдио грађу узорка ткива. Међутим, формулисање исхода учења није увек тако прецизно. Повремено, формулације исхода упућују на својства која је тешко мерити. На пример, исход може да гласи: „Ученик је свестан потребе очувања животне средине“ или „Ученик разуме појам еколошке нише“ (Kennedy, 2007; према: *Настава усмерена на исходе, компетенције и стандарде*, 2015: 94). Отворено је питање како измерити свесност ученика или разумевање одговарајућег појма.

Формулација исхода се фокусира на очекивана понашања. Ипак, то не искључује могућност да ученик током наставног процеса усвоји одређена знања или развије одређене вештине које су важне за наставак учења, а које нису обухваћене исходима. Имајући то у виду, основа за садржај доброг теста

знања некада се не исцрпљује садржајем исхода учења. Истовремено, исходи не указују на начине како да се одговарајућа знања и вештине стекну (*Реформа образовања у Републици Србији: школски програм, 2003*), и најчешће нису контекстуализовани. Исходи би се могли једнозначније разумети када би се указало на услове у којима се може говорити о усвојености одговарајућих знања и вештина.

Дакле, ослањање на исходе учења у циљу конструкције теста знања није нужно ослобођено недостатака који прате ослањање на општије оквири као што су циљ или општи задаци изучавања одређеног садржаја. У циљу конструкције доброг теста знања, пожељно је да се ослонимо на исходе учења који су више операционализовани и конкретније формулисани.

У пракси се неминовно дешава да не остваре сви ученици жељени исход, односно дешава се да исти исход остваре у различитој мери. До ког су нивоа ученици усвојили одређена знања и вештине, може се утврдити уз помоћ образовних стандарда (Стојановић, 2012). У нашој земљи, образовни стандарди су формулисани на три нивоа постигнућа: основном, средњем и напредном. Нивои описују захтеве различите тежине, когнитивне комплексности и обима знања. Сваки наредни ниво подразумева да је ученик савладао знања и вештине са претходног нивоа. На првом нивоу су описани захтеви који представљају базични или основни ниво знања, вештина и умења. На базичном нивоу налазе се темељна предметна знања и умења, односно функционална и трансферна знања и умења која су неопходна за свакодневни живот и за наставак учења. Очекује се да ће око 80% ученика постићи или превазићи овај ниво. Средњи ниво описује захтеве у погледу знања, вештина и умења које просечан ученик може да достигне. Очекује се да ће око 50% ученика постићи или превазићи овај ниво. На трећем нивоу описани су захтеви који представљају напредни ниво знања, вештина и умења. Очекује се да ће око 25% ученика постићи тај ниво. Знања и умења са овог нивоа су трансферна, пре свега за наставак школовања. Компетенције са напредног нивоа су по правилу и когнитивно сложеније од компетенција са базичног и средњег нивоа. То значи да се од ученика очекује да анализира, упоређује,

разликује, критички суди, износи лични став, повезује различита знања, примењује их и сналази се у новим ситуацијама (*Образовни стандарди за крај обавезног образовања*, 2010).

На пример, исходи и стандарди за наставну тему „Грађа људског тела - правила исхрана и последице неправилне исхране“ могу да гласе (*Настава усмерена на исходе, компетенције и стандарде*, 2015: 71-72):

Исход 1: „Ученик је оспособљен да избегава ризична понашања која доводе до болести зависности и нарушавања здравља.“

Исход 2: „Ученик је оспособљен да затражи стручну помоћ када запази промене у психичком здрављу, поремећају рада органа, организма у целини и теже повреде тела.“

Стандард за основни ниво: „Ученик уме да идентификује елементе здравог начина живота и у односу на њих уме да процени сопствене животне навике и избегава ризична понашања“.

Стандард за средњи ниво: „Ученик је оспособљен да процени када може сам себи да помогне и када је потребно потражити лекарску помоћ“.

Стандард за напредни ниво: „Ученик повезује настанак болести, а посебно болести зависности са ризичним облицима понашања и стресом (односно поремећајима психичког стања и здравља личности).“

Једна врста, условно речено, исхода⁵ о којој има значаја говорити у контексту образовања јесу и компетенције. У ужем смислу, компетенције су способност обављања одређене активности у оквиру датог посла или занимања према дефинисаним критеријумима. У ширем смислу, реч је о композиту специфичних знања, вештина и ставова који омогућује обављање одређене активности (Деспотовић, 2010). Иако компетенције имају своје значајно место, пре свега у стручном образовању и домену припреме младих за живот у грађанском друштву, у основној школи конципирање садржаја теста знања на садржају компетенција је мање уобичајено. То је и разумљиво, јер су компетенције изворно усмерене ка свету рада и њихова процена се дешава у

⁵ Компетенције са ужим фокусом се формулишу као исходи (*Настава усмерена на исходе, компетенције и стандарде*, 2015).

реалним ситуацијама. Критеријуми извршења одговарајућих радњи дефинишу не само ниво прихватљивости извршења већ и услове у којима се извршење мора десити (Деспотовић, 2010).

Заједничко за све наведене оквире за разраду садржаја теста је да су прописани и да могу имати статус законских докумената. Опредељење наставника да се ослони на дефинисане оквире јесте определење да се „верује“ онима који су оквире дефинисали. Истовремено, то је определење да се процесу разраде садржаја теста не приступи самостално или барем не у потпуности самостално. Појединим наставницима постојање дефинисаног оквира значајно помаже у раду, јер немају искуства или се не могу ослонити на знање стечено током студија (Максимовић и Марковић, 2012). Међутим, још су бројнији наставници који ослањање на дефинисане оквире посматрају само као „поштовање форме“, док се у свом раду руководе властитим осећајем шта је оно што је за ученике битно (Максимовић и Марковић, 2012). С једне стране, наставници имају потешкоће у интерпетацији исхода и њиховој имплементацији у пракси (Hargreaves *et al.*, 2001). С друге стране, отпор појединих наставника, али и научника, према исходима утемељен је на њиховом тумачењу као редукционистичких, усмерених ка фаворизовању одређених знања и вештина и занемаривању савремених приступа интелигеницији и стиловима учења (Максимовић, 2013).

Када је полазни оквир за разраду тесту већ осмишљен, као у случају исхода, компетенција и стандарда дефинисаних на нивоу школског система, преостаје нам још да саставимо конкретне задатке или питања. Приликом формулације задатака води се рачуна о: а) градиву које се испитује тестом и б) мисаоним процесима који се примењују над одговарајућим садржајем (Fives & DiDonato-Barnes, 2013). За већину наставника, утврђивање домена садржаја не представља велики проблем. Оно што је значајно теже јесте утврдити који процеси мишљења се налазе у основи решавања задатка (Ebel, 1979).

У препознавању процеса мишљења који су у основи решавања задатака, наставницима могу помоћи различите таксономије образовних циљева.

Најпознатија међу њима је Блумова таксономија образовних циљева у когнитивном домену (Павловић-Бабић и сар, 2003).

Блумова таксономија је први пут описана 1956. године. Изворно, таксономија је осмишљена како би олакшала процес конструкције тестова знања. У тренутку када се појавила, таксономија је била намењена наставницима у високошколској настави и требало је да им помогне у састављању тестова на крају школске године (Krathwohl, 2002).

Од тренутка објављивања, па до данашњих дана, Блумова таксономија привлачи велику пажњу научне и стручне јавности. Таксономија се показала као корисно средство у процесу израде теста. Идеја од које су Блум и његови сарадници пошли је да се образовни циљеви могу представити конкретним облицима понашања појединца. Такво понашање се може посматрати и описивати (Bloom, 1981). Може се рећи да је тежња аутора да у дефинисању образовних циљева пођу од нечега што је више очигледно него апстрактно, управо оно што је таксономију учинило тако атрактивном. Таксономија је развијена за три подручја понашања: когнитивно, афективно и психомоторно. У овом раду анализираћемо само питања које се односе на когнитивно подручје које је релевантно за процес израде тестова знања.

Таксономија образовних циљева у когнитивном подручју коју су Блум и сарадници саставили садржала је шест категорија. То су: знање, схватање (разумевање), примена, анализа, синтеза и евалуација. Главне разлике између таксономских категорија одражавају разлике које наставници уочавају између различитих облика понашања ученика (Bloom, 1981). *Знањем* се означава могућност ученика да се сети и искаже појединачне чињенице, опште појмове, да опише методе и процесе, моделе и структуре. *Схватање* је, у изворном значењу, означено као простија форма разумевања која подразумева да ученик зна о чему се говори у задатку, и да може да употреби полазну идеју или податке, а да при том не мора видети у каквом односу се они налазе са другим деловима градива. *Примена* се односи на употребу апстракција у одређеним конкретним ситуацијама. Апстракције могу имати форму општих идеја, правила и поступака, стручних начела, сложенијих идеја или теорије које је

ученик запамтио и у стању је да их успешно примени. *Анализа* подразумева растављање саопштења на саставне елементе, утврђивање хијерархије и односа међу њима. Циљ анализе је да се саопштење учини јаснијим, да се експлицира његов састав и начин на који преноси информативни садржај. *Синтеза* је, пак, састављање делова у целину. Укључује процес сређивања елемената и њихово комбиновање у циљу стварања модела или структуре које пре тога није било. *Евалуација* подразумева просуђивање вредности садржаја и метода на неком подручју у односу на задате или самостално конструисане критеријуме. Процена може бити квантитативна и квалитативна (Bloom, 1981). Образовни циљеви једне категорије надограђују се и укључују облик понашања претходних категорија у наведеној листи.

Табела 3. Блумова таксономија циљева у когнитивном домену⁶

Категорија циља	Тип очекиваног мишљења	Примери формулације задатака
Знање	присећање или препознавање информације која је научена	<i>Дефиниши (одреди)... Који је главни град...?</i>
Разумевање	демонстрирање разумевања материјала; трансформисање, реорганизација или интерпретирање	<i>Објасни својим речима... Упореди... Која је главна идеја у ...? Опиши шта си видео...</i>
Примена	употреба информација у решавању проблема који имају један тачан одговор	<i>Који принцип је демонстриран у...? Примени правило...да би решио...</i>
Анализа	критичко мишљење: идентификовање разлога и мотива; извођење закључака који се заснивају на одређеним подацима; анализирање закључака да би се утврдило да ли су засновани на доказима	<i>Зашто је важан одређени историјски догађај? На основу датог експеримента, која хемијска реакција доводи до...?</i>
Синтеза	дивергентно, оригинално мишљење; оригиналан план, предлог, нацрт или прича	<i>Како бисмо могли да прикупимо новац за...? Како би изгледала Европа да је Немачка победила у Другом светском рату?</i>
Евалуација	процењивање вредности идеја, изношење мишљења, процењивање стандарда	<i>Према твом мишљењу, који је кошаркаш тренутно најбољи? Због чега? Шта мислиш, који је сликар бољи...? Зашто тако мислиш?</i>

⁶ Преузето из: Павловић-Бабић и сар., 2003.

Табела 3. приказује категорије циљева који чине Блумову таксономију, тип мишљења који одговара конкретом циљу и примере како се може формулисати задатак који испитује одговарајући ниво знања. За сваки од когнитивних домена у Блумовој таксономији, могу се утврдити листе глагола који упућују на испитивање одговарајуће мисаоне активности ученика (Прилог 2).

Бројни аутори су емпиријски проверавали Блумову таксономију, а провере су водиле ка ревизијама полазне таксономије. Једна од најпознатијих ревизија је ревизија Андерсона и Кратвола (Anderson & Krathwohl) из 2001. године. У односу на изворну таксономију која је једнодимензионална, ревидирана верзија садржи две димензије. Једну димензију чине знања, а другу когнитивни процеси.

Аутори ревидиране верзије таксономије уочили су да се у пракси у формулацији сваког образовног циља употребљава именица или групе речи, са једне стране, и глагол са друге стране. Именица или група речи се обично односи на садржај учења, а глагол на когитивни процес. На пример, „Кандидат би требало да се сети закона понуде и потражње“. У датом примеру, закон понуде и потражње представља садржај учења, док глагол „сетити се“ указује на когнитивну радњу коју би кандидат требало да покаже у испитној ситуацији (Krathwohl, 2002: 213). Андерсон и Кратвол су приметили да у Блумовој таксономији само димензија „знања“ у опису садржи и именице и глаголе, док су све остале категорије описане употребом глагола. То их је подстакло да измене категорију „знање“, односно да таксономију обликују у две димензије, одвојивши знања од когнитивних процеса.

Димензија знања обухвата следеће категорије: познавање чињеница (терминологија, специфични детаљи и елементи), концептуална знања (класификације, категорије, принципи, теорије, методе и структуре), процедурална знања („знати како“, познавање процедура, услова под којима се примењују одговарајуће процедуре, овладаност специфичним вештинама у одговарајућој области знања) и метакогнитивна знања (познавање стратегија, свест о когнитивним процесима, самоспознаја) (Krathwohl, 2002).

Када је реч о когнитивним процесима, ревидирана верзија таксономије задржала је 6 категорија, али су појединим категоријама промењени називи. Тако, категорија „знање“ добија назив „сетити се“, „схватање“ постаје „разумети“, „примена“ добија назив „применити“, „анализа“ постаје „анализирати“, „синтеза“ постаје категорија „евалуирати (вредновати)“, а „евалуација“ постаје категорија „креирати“. Пета и шеста категорија које дефинише Блум су у ревидираној верзији замениле места. Као и Блумова, и ревидирана таксономија има хијерархијски уређену структуру. Очекује се да је свака следећа категорија сложенија од претходне (Krathwohl, 2002).

С обзиром на чињеницу да се сваки образовни циљ може представити у две димензије, Кратвол (Krathwohl, 2002: 214) предлаже да се коришћење ревидиране верзије таксономије ослони на примену *таксономске табеле* (*Taxonomy Table*) (табела 4).

Табела 4. Пример примењене таксономске табеле ⁷

Когнитивни процеси						
Знање	сетити се	разумети	применити	анализирати	евалуирати	креирати
познавање чињеница	x					x
концептуална знања		x			x	x
процедурална знања						
метакогнитивна знања						

У датој табели, редови представљају категорије знања, а колоне когнитивне процесе. У пресеку реда и колоне налази се ћелија која показује шта је то што би ученик у одговарајућој области знања требало да „зна“ и шта би требало да „уради“ са тим садржајем (односно које когнитивне процесе би требало да покаже). Кратвол (2002) оцењује да табела обезбеђује јасну, визуелну репрезентацију одговарајућег наставног садржаја. Прегледом табеле,

⁷ Адаптирано према: Krathwohl, 2002: 217.

наставник може донети одлуку да ли је потребно нешто да промени у свом раду, како у домену планирања, тако и у домену непосредне реализације часа.

Поред Блумове таксономије, користи се и таксономија структуре регистрованих исхода учења (SOLO таксономија) која даје квалитативан опис повећања компетенција код ученика на пет нивоа. SOLO је скраћеница од „*structure of the observed learning outcome*“ (<http://www.johnbiggs.com.au/academic/solo-taxonomy/>). Реч је о таксономији аутора Бигса и Колинса (Biggs & Collis) први пут објављеној 1982. године. SOLO таксономија омогућава да се, у ширем смислу, идентификује стадијум на коме учениково мишљење тренутно оперише. Стадијуми су следећи (Hattie, Jaeger & Bond, 1999: 406):

- *преструктурални стадијум* – ученик задатку не прилази на прави начин
- *једноструктурални стадијум* – ученик добро разуме један аспект задатка, али још увек чињенице не доводи у међусобну везу
- *мултиструктурални стадијум* – ученик добро разуме два ли више аспеката задатка, али их не доводи у међусобну везу
- *релациони стадијум* – ученик успешно интегрише неколико аспеката задатка у јединствену целину која има дефинисану структуру и јасно значење
- *проширена апстракција* – кохерентна целина је уопштена и подигнута на виши ниво апстракције.

SOLO таксономија још увек нема тако широку примену као Блумова таксономија. У литератури се оцењује да је њена примена најбоље прилагођена оцењивању усмених одговора ученика (Бјекић и Папић, 2005).

Као и критика образовних исхода, критика таксономија образовних циљева усмерена је ка њиховој вредносној обојености. Истиче се да класификације не могу бити вредносно неутралне уколико образовне циљеве постављају у хијерархијски низ (Furst, 1981; према: Максимовић, 2013), што је одлика и Блумове и SOLO таксономије. Другим речима, ослањање на дефинисане таксономије образовних циљева значи и прихватање вредносног система који је у њиховој основи.

Таксономије су разрађене као помоћ у конструкцији теста. С обзиром на више ревизија које су претрпеле и с обзиром на напредовање сазнања у вези са процесом мишљења, разумљиво је да се таксономије не могу сматрати коначним. Такође, чињеница да су разрађене листе глагола као додатна помоћ, упућује на закључак да интерпретација одговарајућих категорија циљева није увек једнозначна или да у пракси постоје потешкоће са њиховим разумевањем. Такође, конструктор теста не мора тежити томе да тест конципира имајући у виду сваку од наведених категорија циља или сваки стадијум. С обзиром на различите карактеристике ученика са којима наставници раде, адаптирање таксономских категорија у домену примене је неминовно.

Задатак – основна јединица теста знања

Терминолошке дилеме у вези са одређењем задатка

Основне јединице теста су задаци. Задаци се састављају на основу индикатора или показатеља стеченог знања у одговарајућој области (Буквић, 2007). Индикатори су увек описи одређених понашања, а задаци подстицаји који испитанике усмеравају у правцу одговарајуће активности. Задатак може да гласи:

„Истраживањем се испитује став средњошколаца према настави ликовне културе. Формулишите једну позитивну и једну негативну тврдњу у односу на ставски објекат“.

Како су индикатори шири од задатака, из једног индикатора могуће је извести више задатака. У односу на претходно наведени пример задатка, индикатор може да гласи:

„Студент успешно конструише позитивне и негативне тврдње у скали ставова Ликертовог типа“.

Овако формулисан индикатор омогућава да се конструишу различити задаци у којима се очекује да испитаник на задатим примерима покаже своје знање у домену конструкције тврдњи у скали ставова Ликертовог типа. Док индикатори могу бити формулисани стручним језиком, задаци морају добити језички облик који је јасан испитаницима (Буквић, 2007).

Неретко, уместо о задацима говоримо о питањима у тесту. Са методолошког становишта, има основа да се нагласи разлика питања у односу на задатак у тесту знања. За разлику од задатка, питање не садржи специфичне информације које су неопходне за пружање тачног одговора (Фридман и Джумаев, 1973; према: Дубљанин, 2010). Како би било могуће понудити одговор на питање, испитаник мора да открије податке који недостају. Задатак, напротив, садржи све податке неопходне за продуковање тачног одговора (Ward, Carlson & Woisetschlaeger, 1983; Дубљанин, 2010). Ипак, то не значи да је довољно репродуковати дате податке, већ је за решавање задатка најчешће потребно применити их на одређени начин, или на темељу задатих, формирати нова знања и појмове (Дубљанин, 2010). Питање се посматра и као структурални елемент задатка, јер се стабло одговарајућих форми задатака може састојати од питања.

На основу наведеног, могло би се закључити да су питања једноставнија од задатака. То, ипак, не мора нужно бити тако. Постоје околности у којима је тешко разграничити питање од задатка. Такав пример су сложена проблемска питања у науци која се могу упоредити са најсложнијим задацима (Дубљанин, 2010).

У основи задатка налази се вербална формулација проблема. Управо вербална формулација одређује тежину задатка и тип анализе који ће бити примењен. Променом формулације, може се значајно променити и тежина задатка. Према становишту Рубинштајна (1981), то показује да формулација

задатка није само вербална, него и мисаона чињеница, а вербално формулисање захтева истовремено и акт анализе.

Дакле, иако се често не наглашава разлика између питања и задатка у тесту знања, исправније је јединице садржаја теста означавати задацима. Осмишљавање задатака је најчешће сложеније од састављања питања, јер је неопходно да задатак буде довољно директиван, и када је реч о садржају који ће сачињавати одговор, и у погледу операција над садржајем које воде решењу задатка.

У методолошкој литератури разматра се и питање односа задатка и ајтема. Ајтем, „ставка“ или „честица“ одређује се као јединица скоровања, док је задатак скуп ајтема који су структурално повезани (Ebel, 1981). На пример:

Са леве стране наведене су врсте истраживачких пројеката, а са десне поједини њихови елементи. Означите који елемент припада којој врсти пројекта тако што ћете бројеве 1, 2 и 3 уписати испред одговарајућег елемента са десне стране.

1) идејни пројекат	_____узорак
2) студијски пројекат	__1__дефинисање кључних појмова
3) технички пројекат	_____инструменти
	_____хипотезе
	_____методе и технике
	_____преглед планираних трошкова

У датом примеру, повезивање идејног пројекта истраживања са дефинисањем кључних појмова у истраживању, као његовим саставним елементом, представља један ајтем у задатку. Приказани задатак обухвата укупно 6 ајтема.

У страниј литератури, значење задатка се повремено одређује веома широко. Задатком се у том случају означавају различити начини испитивања ученика који се не изводе писаним путем. На пример, испитивање ученика путем извођења одређених радњи, демохстрирањем, па чак и посредством разговора у неформалним условима (Marzano, 2006). У домаћој литератури,

ајтем се код тестова способности поистовећује са задатком (Фајгељ, 2004). Иако су појмови „ајтем“ или „ставка“ прихваћени термини у психолошким и педагошким речницима, у свакодневној комуникацији наставника они немају прецизно значење (Бјекић и Папић, 2006).

Структуру задатка чине *стабло* и *одговор* (Postlethwaite, 2005). Стабло задатка представља питање, проблемску ситуацију коју је потребно решити или упут да се обави одређена рачунска операција. На пример, стабло може да гласи:

Шта се подразумева под природним експериментом?

У зависности од форме задатка, одговор подразумева више или мање слободну форму саопштавања мисли испитаника, односно систем бирања понуђених одговора, спаривања понуђених одговора или уређивања понуђених одговора у одговарајући низ.

На пример, одговор може подразумевати слободну форму образложења коју испитаник уписује у простор предвиђен за одговор:

Шта се подразумева под природним експериментом?

(простор за уписивање одговора)

Такође, одговор може подразумевати да испитаник у низу понуђених одабере одговор који сматра тачним у односу на стабло задатка. На пример:

Шта се подразумева под природним експериментом?

1. експеримент који се спроводи изван лабораторије
2. *ex-post-facto* поступак
3. експеримент који се спроводи са једном групом испитаника
4. експеримент који се спроводи са две групе испитаника
5. факторски експеримент

Добар задатак у тесту знања испуњава два основна критеријума: испитује одређено знање или одређену вештину (представља адекватан индикатор понашања које се мери) и прецизно је формулисан. Други захтев подразумева да је задатак довољно информативан за очекивани одговор и јасан у вези са дужином и формом очекиваног одговора, логично структуриран, граматички и правописно исправан (Пејић и Тодоровић, 2007).

Врсте задатака и њихове карактеристике

Најопштија подела задатака у тесту је подела на задатке са бирањем одговора и на задатке са уписивањем одговора (Taylor, Greer & Mussio, 1978; Popham, 2003). Задаци бирања одговора подразумевају да испитаник, у односу на захтев задатка, међу понуђеним одговорима одабере један или више одговора које сматра тачним. Задаци са уписивањем одговора подразумевају да испитаник уписује свој одговор на постављени захтев.

Подврсте задатака са бирањем одговора су задаци *двочланог избора, вишеструког избора, спаривања, сређивања* (Баковљев, 1997) и задаци типа „К“ (Haladyna & Downing, 1989) као подврста задатака вишеструког избора.

Задаци двочланог, двоструког или алтернативног избора. Структуру задатка двочланог избора најчешће чине: полазна тврдња и два понуђена одговора. Испитаник одговара на задатак избором једног од два наведена одговора. Задаци двочланог избора обично се састоје од тврдње у односу на коју је потребно одговорити да ли је „тачна“ или „нетачна“. На пример:

Цитирање је дословно навођење мисли другог аутора.

тачно

нетачно

Нешто мање заступљена форма задатака двочланог избора је форма вишеструких двочланих избора. У датој форми, ученици имају задатак да прочитају одређени текст. Потом је потребно да одреде које у низу понуђених тврдњи у вези са прочитаним текстом су тачне, односно нетачне. На пример (Slideplayer.com/slide/271053):

Решавајући 10 задатака вишеструког избора на тесту, 12 ученика је остварило следеће скорове: 5,6,7,7,7,7,8,8,8,9,10. На основу наведених података, утврдите тачност следећих тврдњи:

- | | | |
|---|-------|---------|
| 1) медијана за наведене скорове износи 7.5 | | |
| | тачно | нетачно |
| 2) мод за наведени низ скорова износи 8.0 | | |
| | тачно | нетачно |
| 3) распон наведених скорова је 5.0 | | |
| | тачно | нетачно |
| 4) вредност аритметичке средине скорова је иста као вредност медијане | | |
| | тачно | нетачно |

Иако се најчешће користе за испитивање нивоа познавања чињеница, задаци двочланог избора нису нужно лаки за испитанике. На пример, искуство у примени задатака са вишеструким двочланим изборима показало је да их ученици доживљавају као тешке (Porham, 2003).

Предности задатака двочланог избора огледају се у релативно једноставној конструкцији, брзом решавању и оцењивању (*Guide to Assessment*, 2015). Кључним недостатком задатака алтернативног избора сматра се висока вероватноћа погађања тачног одговора. На нивоу појединачног задатка двочланог избора вероватноћа погађања тачног одговора је 50% (Porham, 2003).

У наставку, указаћемо на неке од препорука за исправно конструисање задатака двочланог избора (Porham, 2003):

- потребно је избегавати негативне тврдње, као и двоструку негацију у формулацији тврдње
- једном тврдњом потребно је обухватити један однос између појмова

- пожељно је формулисати паран број задатака двочланог избора на нивоу теста, тако да је половина тврдњи тачна, а друга половина нетачна
- тврдње се могу формулисати тако да површно читање води ка нетачном одговору
- потребно је водити рачуна да тврдње које су тачне, као и оне које су нетачне буду приближно исте дужине.

Све наведене препоруке, изузев препоруке која се односи на формулисање тврдњи тако да површно читање води ка нетачном одговору, саставни су део различитих методолошких приручника које анализирају Фреј и сарадници (Frey *et al.*, 2005). Суштина наведене препоруке је у томе да она подсећа истраживача да осмишљавање доброг задатка алтернативног типа није нужно једноставно и да је потребно промислити на који начин је задатак формулисан.

За разлику од препоруке да се формулише половина тачних и половина нетачних тврдњи, у литератури се може срести и мишљење да је пожељније да укупан број нетачних тврдњи на нивоу теста буде нешто већи од укупног броја тачних тврдњи. Показало се да су нетачни ајтеми дискриминативнији у односу на ајтеме који су тачни, управо због склоности испитаника да приликом „погађања“ одговора чешће бирају афирмативни облик (Jacobs, 2002). Такође, уколико је испитанику познато да је број тачних и нетачних тврдњи уједначен, сигурност у исправно решење половине задатака, омогућава испитанику да исправно одговори на преостале задатке алтернативног избора.

Према резултатима истраживања која разматрају Халадајна и Даунинг у свом раду (Haladyna & Downing, 1989) негативне формулације чине задатак тежим. С обзиром на тенденцију да се задаци двочланог избора користе за испитивање нивоа познавања чињеница, очекује се да су то уједно и релативно лаки задаци за ученике. У оној мери у којој се задатком жели испитати неки виши ниво знања, прихватљиво је коришћење негативне формулације у поставци задатка.

Задаци вишеструког избора. Задаци вишеструког избора имају два структурална елемента: стабло и понуђене одговоре (алтернативе или дистракторе) (Баковљевић, 1997; Popham, 2003). У тестовима знања се могу срести четири варијанте задатака вишеструког избора које комбинују следећа решења: а) стабло се састоји од: питања или започете мисли; б) од испитаника се очекује да одабере: тачан одговор или најбољи одговор међу понуђенима. У начелу, сматра се да је боље решење да се стабло састоји од питања, јер започета мисао може да збуни испитаника (Haladyna & Downing, 1989; Popham, 2003). На пример:

Којој техници припада Фландерсов протокол интеракцијске анализе?

- 1) анкетирању
- 2) интервјуисању
- 3) посматрању
- 4) анализи садржаја
- 5) скалирању

Одабир најбољег од понуђених одговора представља већи изазов за испитаника (Popham, 2003), а његова примена је могућа уколико је једна или више алтернатива делимично тачна (Cohen & Wollack, 2012).

Предности задатака вишеструког избора огледају се у чињеници да се њима могу испитати различити нивои знања, односно могу представљати сложен когнитивни задатак за ученике (Buser, 1996; Popham, 2003). Истовремено, реч је о задацима који су једноставни за оцењивање, а типичне грешке у одговорима ученика пружају корисну повратну информацију наставнику у вези са тим где постоје најчешћа погрешна тумачења (Буквић, 1996; Buser, 1996; Popham, 2003).

Основна слабост задатака вишеструког избора јесте могућност препознавања тачног одговора у низу понуђених, који ученик вероватно не би могао самостално да формулише да није понуђен. Успешност ученика у препознавању одговора може да доведе до погрешног закључка у вези са

његовом укупном овладаношћу градивом (Popham, 2003). Састављање задатака вишеструког избора није једноставно и захтева доста времена, а поједини аутори сматрају да се овај тип задатака ипак најуспешније користи за испитивање нивоа познавања чињеница (Гојков, 2012).

Неке од препорука за исправно конструисање задатака вишеструког избора су (Popham, 2003):

- потребно је да питање или започета мисао у стаблу задатка буду самодовољни, односно да није потребно да захтев буде додатно појашњен
- избегавати негативне формулације у стаблу задатка
- свака понуђена алтернатива мора бити граматички уклопљена са стаблом задатка
- пожељно је да свака алтернатива буде формулисана тако да може деловати као прихватљив одговор на питање, али да и поред тога остаје само један тачан одговор, односно најбољи од понуђених
- позиција тачног одговора би требало да буде насумично одређена у низу задатака вишеструког избора
- није пожељно као алтернативу користити „све горе наведено“ – препознавање једног погрешног одговора аутоматски елиминише дату опцију, а препознавање два тачна одговора идентификује дати одговор као тачан, чак иако су испитанику сви остали одговори непознати (Rudner & Schafer, 2002).

Занимљиво је да Пофам препоручује да се као алтернатива користи „ништа од наведеног“. Анализа 20 методолошких приручника коју су спровели Фреј и сарадници (2005) показује да већина методолога сматра да то није пожељно, осим уколико је у питању рачунски задатак (Rudner & Schafer, 2002). Све остале наведене препоруке препознају се и у методолошким приручницима које Фреј и сарадници анализирају.

Иако су поменути аутори сагласни око тога да задатак може имати један тачан одговор, могуће је предвидети и више тачних одговора на нивоу задатка вишеструког избора. На пример, задатак са 5 понуђених алтернатива може

имати два тачна одговора. Наведена могућност повлачи са собом специфичан приступ бодовању одговора. Истраживач мора одлучити, најпре, да ли сваки од тачних одговора носи по један поен, или сваки носи по 0,5 поена. Такође, потребно је одлучити да ли се уводе негативни поени за погрешне одговоре и колики је износ негативних поена по изабраном дистрактору? На пример, како поступити у ситуацији када испитаник изабере један тачан и један погрешан одговор – да ли испитаник у описаном случају добија 0 поена, 0,5 или – 0,5 поена? Опредељење за увођење већег броја тачних одговора на нивоу задатка, захтева доношење одлука о начину бодовања одговора, како теста у целини, тако и на нивоу појединачних задатака. Такође, истраживач пре тестирања доноси одлуку и о томе да ли ће испитаницима скренути пажњу да постоји више тачних одговора на нивоу поједних задатака, у случају да таква информација није понуђена у поставци задатака (у загради или формулацијом задатка која упућује на множину тачних одговора). Уколико се испитаницима не даје таква информација, задатак постаје тежи.

Када је реч о састављању дистрактора, у литератури се описују три начина (Haladyna & Downing, 1989). Један начин је да истраживач самостално осмисли алтернативе које могу деловати прихватљиво испитанику који није сигуран у тачан одговор. Други начин је да се као алтернативе понуде најчешћи погрешни одговори испитаника на задато питање, посебно ако су истраживачу доступни одговори на задатак истоветног садржаја, али са формом задатка са уписивањем одговора. Трећи начин је израчунавање индекса дискриминативности дистрактора и избор дистрактора чији индекс има одговарајућу вредност. Испитивања упућују на закључак да је за састављање дистрактора најбоље користити први или други начин (Haladyna & Downing, 1989).

У погледу препорученог броја понуђених алтернатива, постоје различита уверења. На пример, Баковљев (1997) предлаже да понуђених алтернатива у задатку вишеструког избора буде четири или пет. Бусер (Buser, 1996) оцењује да је пожељно понудити између три и пет алтернатива, а подршку таквом становишту даје и анализа различитих методолошких

приручника Фреја и сарадника (2005). Руднер и Шафер (2002) оцењују да је довољно понудити три или четири квалитетна дистрактора. Лорд се позива на нека ранија испитивања (Tversky, 1964; Grier, 1975; према: Lord, 1977) која показују да је довољно да задатак вишеструког избора има три понуђена одговора.

Мањи број алтернатива, повећава вероватноћу погађања тачног одговора. На пример, уколико је понуђено 5 алтернатива у задатку, вероватноћа погађања тачног одговора износи 20%, док у случају три понуђене алтернативе вероватноћа расте на 33%. Аутори који не наглашавају важност великог броја дистрактора, заправо наглашавају њихов квалитет. Између задатка са три квалитетна дистрактора и задатка са 6 мање квалитетних дистрактора, већина методолога би се определила за први задатак. Израда квалитетних дистрактора није лак задатак. Посвећеност том циљу као резултат има задатак са мањим бројем дистрактора међу којима ће ученик без довољно знања тешко „препознати“ тачан одговор.

Позиција тачног одговора у низу алтернатива требало би да буде насумично уређена како би се обезбедила већа поузданост тестирања. Наиме, увођење система у позиционирање тачног одговора оставља могућност испитаницима да утврде систем. На пример, испитаници могу закључити да се тачан одговор увек налази у средини низа понуђених алтернатива; на почетку или на крају низа; да се са ређањем низа питања систематски помера и позиција тачног одговора и томе слично. Утврђивање система позиционирања тачног одговора омогућава испитаницима да исправно реше задатак и онда када, заправо, не знају тачан одговор.

У литератури се разматра и модел задатка који се означава као тип „К“ (Haladyna & Downing, 1989). Назив је скраћеница за „комплексан модел задатка вишеструког избора“ (у оригиналу: „*complex multiple choice*“). С обзиром на структуралну сличност са задацима вишеструког избора, могу се означити њиховом подврстом. Задатак типа „К“ се састоји од стабла, и два низа понуђених одговора. Модел се може представити следећим примером:

Шта одређује поузданост теста?

1. дужина теста
2. хомогеност задатака
3. варијабилност скорова на тесту
4. дужина ајтема

- а) 1,2 и 3
- б) 1 и 3
- в) 2 и 4
- г) 1 и 4
- д) само 4

У датом примеру, испитаник би требало да заокружи слово испред тачног одговора.

Предности и недостаци задатака наведеног типа истоветне су предностима и задацима задатака вишеструког избора, с додатком још једног недостатка. Наиме, тип „К“ сматра се формом која је сложена за испитанике, и прати је детаљно упутство за решавање. Из тог разлога, већина методолога наведену форму задатака не разматра у својим приручницима, док је значајан број аутора негативно оцењује са становишта примене у пракси (Haladyna & Downing, 1989).

Препоруке у вези са конструкцијом задатака типа „К“ практично су истоветне препорукама које се односе на задатке са вишеструким избором.

Задаци спаривања. Задаци спаривања састоје се од стабла задатка и понуђених одговора, организованих у више низова података. Најчешће, задатак садржи два низа података. Задатак испитаника је да повеже елементе једног низа – премисе (обично наведене са леве стране) са елементима другог низа – одговорима (обично наведеним са десне стране) према задатом критеријуму (Бјекић и Папић, 2006). На пример:

Са леве стране наведене су истраживачке технике, а са десне инструменти. Уписивањем одговарајућег броја испред назива инструмента, означите који инструмент припада којој техници.

- | | | |
|---------------------|-------|------------------|
| 1) интервјуисање | _____ | протокол |
| 2) анкетање | _____ | чек листа |
| 3) посматрање | _____ | тест способности |
| 4) анализа садржаја | | |
| 5) скалирање | | |
| 6) тестирање | | |

Предност задатака спаривања је та што могу да обухвате велики обим градива, и да испитају различите нивое знања (Џордан, 1966; Мужич, 1968; Porham, 2003). Заједно са задацима сређивања, ова врста задатака се сматра најпоузданијом, јер је могућност погађања тачног одговора врло мала (Бјекић и Папић, 2006). Могућност погађања одговора је посебно мала уколико се исти податак може користити више пута приликом формирања парова података.

Предношћу задатака спаривања обично се сматра и њихово оцењивање, које је брзо и лако (Гојков, 2012). О наведеној предности могуће је говорити уколико су претходно донете одлуке о поступању у ситуацији када испитаник делимично исправно одговори на задатак. Дилема коју оцењивач теста може имати је да ли има основа да се задатак призна као делимично тачан и на који начин то остварити. На пример, на нивоу задатка је потребно формирати 6 парова података. Уколико испитаник успешно формира четири пара, а код преосталих парова погрешно, да ли је исправно задатак признати као делимично тачан? Уколико се донесе одлука о делимичном признавању, како би се бодовао добијени одговор? Један од начина је да се бодују само парови који су исправно формиран, а да се преостали парови занемаре. Друга могућност је да се за погрешно формиране парове предвиде негативни поени. Следеће питање је колико поена би требало да носи сваки исправно формиран пар, а колико поена би се одузимало у случају формирања неодговарајућег пара. Наведена питања истраживач би требало да размотри пре реализације

тестирања. Поступак бодовања за који се одлучи, истраживач би требало доследно да примењују у свим задацима спаривања у тесту.

Основни недостатак задатака спаривања је сложено конструисање. Састављање обично дуго траје, пре свега због потребе да упутство за решавање задатка буде довољно једноставно и истовремено довољно прецизно (Гојков, 2012). Упркос потенцијалу који носе, у пракси се за ову форму задатака најчешће опредељујемо у циљу утврђивања нивоа познавања чињеницама (Popham, 2003).

Неке од препорука за исправно конструисање задатака спаривања су (Popham, 2003):

- пожељно је да сви елементи у листама буду кратко формулисани, са тенденцијом да одговори буду краћи од премиса
- сви елементи једне листе требало би да припадају истој класи чињеница, како би листа била хомогена
- пожељно је да одговора буде више него премиса
- препоручује се да листа одговора буде уређена по одређеном принципу, на пример по абecedном реду
- потребно је дефинисати критеријум повезивања премиса и одговора, али и прецизирати колико пута се исти одговор може употребити приликом спаривања (једном, два или више пута)
- пожељно је да листа са премисама и листа са одговорима технички припадају истој страници на тесту.

Предлог да листа одговора буда дужа од листе премиса, доприноси већој поузданости тестирања. Уколико би број био изједначен, а испитаницима познато да се сваки податак може употребити једанпут, испитаници би најмање један пар података могли формирати без уочавања њихове повезаности. На пример, уколико је број премиса и одговора 5, након формирања четири пара, испитанику преостаје само да повезе последњи пар података.

Хомогеност података у задацима спаривања, доводи се у везу са тежином задатка. Наиме, што су елементи једног низа хомогенији међу собом, задатак

спаривања је тежи за решавање (Џордан, 1966). Претпоследња у низу препорука, такође се може довести у везу са тежином задатка. Уколико се наведена препорука не примени у изради задатка спаривања, задатак постаје тежи за решавање. На пример, уколико испитаник не уочи да се одређени податак користи за формирање два пара података, а инструкција за решавање то не експлицира, јасно је да испитаник не може остварити максималан број поена на задатку. Такође, може доћи у ситуацију да на задатку не оствари ниједан поен, уколико делимично признавање одговора није предвиђено.

Све наведене препоруке, изузев препоруке да листа одговора буде уређена по одређеном принципу, саставни су део различитих методолошких приручника које анализирају Фреј и сарадници (Frey *et al.*, 2005). Додатно, њихова анализа показује да број одговора у задацима спаривања може да варира са узрастом ученика. Препоручује се да број одговора у тестовима намењеним ученицима основне школе не прелази 7, док је за ученике средње школе лимит 17 одговора.

Задаци сређивања. Задаци сређивања састоје се од стабла задатка и неуређеног низа понуђених одговора. Задатак испитаника је да наведене елементе уреди према задатом критеријуму, уписивањем редних бројева поред сваког од понуђених одговора. На пример:

Наведене елементе распоредите тако да одговарају фазама израде скале Ликертовог типа. Распоређивање извршите уписивањем бројева 1, 2, 3, 4, 5 и 6 на празне црте. Број 1 упишите поред прве фазе.

- _____ бодовање одговора
- _____ припрема прелиминарне листе тврдњи
- _____ утврђивање дискриминативности тврдњи
- _____ анализа и редиговање тврдњи од стране истраживача
- _____ избор коначне листе тврдњи
- _____ задавање прелиминарне листе испитаницима

Задаци сређивања могу испитати усвојеност градива већег обима и различите нивое знања, што су њихове основне предности (Мужић, 1968). Остале предности задатака сређивања су да се лако састављају, брзо решавају, а могућност погађања тачног одговора је мала (Бјекић и Папић, 2006).

Недостаци задатака сређивања се огледају у тешкоћама бодовања могућих одступања у редоследу одговора, у чињеници да заузимају доста простора у тесту, захтевају детаљну инструкцију за решавање, док одговори ученика могу бити одраз механичког памћења редоследа радњи, односно учења без разумевања (Бјекић и Папић, 2006).

Неке од препорука за исправно конструисање задатака сређивања су (Бјекић и Папић, 2006):

- пожељно је изабрати редоследе радњи или процеса који су непроменљиви
- потребно је саставити јасну инструкцију према ком критеријуму би ученици требало да уреде елементе низа
- елементе које је потребно уредити пожељно је распоредити по одређеном принципу, на пример по абecedном реду, или се одлучити за случајни редослед
- препоручује се да се предвиди начин бодовања тачних одговора, као и начин бодовања у случају да испитаник промени редослед појединих етапа.

Прва у низу наведених препорука подсећа на важност избора низа радњи или процеса који су коначни и непроменљиви. Уколико би редослед радњи/процеса имао више прихватљивих варијанти, оцењивач теста би био у обавези да сваку од варијанти предвиди као тачан одговор и да поступак оцењивања прилагоди тој околности. Поступак оцењивања задатака сређивања битно је отежан чињеницом да испитаник може одређене елементе редоследа да замени, односно да делимично тачно одговори на постављени захтев. Слично као у случају задатака спаривања, неопходно је да пре реализације тестирања истраживач предвиди да ли се делимично тачан одговор може прихватити, да ли се уводе негативни поени у случају погрешног

распореда појединих елемената у низу, и како се утврђује коначан број поена на нивоу задатка.

Подврсте задатака са уписивањем одговора су *задачи испуњавања налога, есејска питања, задачи са кратким одговорима, задачи допуњавања* (Баковљевић, 1997) и *структурирани задачи* (Guide to Assessment, 2015).

Задачи испуњавања налога. Структуру задатка испуњавања налога чини стабло задатка и простор за уписивање одговора. Стабло задатка овог пута има форму питања, описа проблемске ситуације или поставке рачунског задатка. У зависности од структурирања стабла задатка, испитаник у простор за одговор уписује: одговор на постављено питање, предлаже решење проблемске ситуације или решава рачунски задатак. Одговор на питање у задатку (пример а)) или предлог решења проблемске ситуације (пример б)) обично се састоји од неколико реченица које испитаник уписује у простор за одговор. На пример:

а) Шта се подразумева под корелационим везама и односима у педагогији?

(простор за уписивање одговора)

б) Предмет истраживања је став ученика ромске популације према образовању. Опишите на који начин се може формирати репрезентативан узорак испитаника у датом истраживању.

(простор за уписивање одговора)

Рачунски задатак подразумева да се примене одговарајуће математичке операције над подацима који су саопштени у стаблу задатка. На пример:

Једна школа има 612 ученица и 588 ученика. Израчунај и наведи колики је проценат ученица.

(простор за уписивање одговора)

Предности задатака испуњавања налога су могућност примене за испитивање свих нивоа знања и релативно једноставна конструкција (*Guide to Assessment*, 2015). Недостаци се огледају у мањем обухвату наставног градива (*Reiner et al.*, 2002; *Guide to Assessment*, 2015), дуготрајном решавању и сложенем оцењивању ученичких одговора (*Buser*, 1996; *Porham*, 2003; *Guide to Assessment*, 2015).

Препоруке за исправно конструисање задатака испуњавања налога су (*Reiner et al.*, 2002; *Porham*, 2003):

- формулација захтева би требало да буде довољно прецизна како би омеђила ширину ученичких одговора
- потребно је унапред одредити шта се све може признати као тачан одговор, али се препоручује и да се испитаницима одреди дужина одговора (најчешће број речи или редова) и да се препоручи време за писање одговора
- уколико је потребно направити избор између већег броја задатака који захтевају краћи одговор и неколицине задатака који захтевају дужи одговор, препоручује се да се одабере више задатака са краћим одговорима

- не препоручује се остављање могућности испитаницима да међу неколико понуђених питања изабере питање на које желе да дају одговор.

Напоменимо да друга препорука у извесној мери одступа од онога што је уобичајена пракса у нашем школском систему. Наиме, прецизно дефинисање броја речи или редова које би одговор требало да садржи, односно препоручивање времена за писање одговора, готово су непознате праксе у нашим школама. У писању препорука, аутори су вероватно имали у виду амерички школски систем у коме је наглашена висока стандардизација поступка тестирања. Строга стандардизација услова у којима се тестирање одвија доводи до тога да се „уситњавају“ кораци које је потребно начинити како би се стигло до жељеног циља. У том смислу, може се разумети због чега се ученицима дају детаљне инструкције о поступању у ситуацији тестирања. Стиче се утисак да би дефинисање дужине одговора или трајања одговарања учинило ситуацију тестирања још више стресном. Могуће је, међутим, такве препоруке посматрати и као начин да се помогне ученику да направи разлику између више и мање сложених захтева. Такође, у Америци је уобичајено да се припреми за тестирање посвећује доста пажње. Могуће је да се припреме за тест тако организују да ученици на часовима припреме увежбавају вештину усаглашавања дужине свог одговора са препорученом дужином, тако да у ситуацији решавања теста нису суочени са захтевом који им је непознат.

Последња препорука у приказаном низу упозорава да решавање задатака који су у основи другачији, резултате испитаника не чини упоредивим (Buser, 1996; Cohen & Wollack, 2012).

Есејска питања. Структуру есејских питања чини стабло задатка, обично у форми питања, и простор за одговор. Задатак испитаника је да упише свој одговор у за то предвиђен простор.

Есејска питања имају најдужу историју од како постоје подаци о испитивању ученичких знања писаним путем (Porham, 2003). Своју дугу традицију дугују флексибилности и чињеници да се њиховом применом могу

испитивати различити нивои знања. Када је потребно извршити поређења, супротставити контрасте, изложити претпоставке, извести дедукције или донети закључке, најбоље је одредити се за задатке есејског типа (Џордан, 1966). Недостаци који прате примену есејских питања су дуго време решавања, нејасан или неуједначен критеријум оцењивања ученичких одговора од стране наставника (Buser, 1996; Popham, 2003) и мањи обухват градива (Reiner *et al.*, 2002). С обзиром на мањи обухват градива, увек остаје могућност да процена ученичког знања не буде реална. Наиме, може се догодити да је ученик успешно савладао само део градива који се обухвата есејским питањима (Рајстон, Џастман и Робинс, 1966). Високо постигнуће које ученик оствари у датим околностима не значи да је успешно савладао целокупно градиво.

Препоруке за конструисање есејских питања, у основи су истоветне препорукама за састављање задатака испуњавања налога, са нагласком на избору глагола који ће покренути одговарајуће мисаоне процесе код ученика (Reiner *et al.*, 2002). Како би се есејским питањем испитало разумевање градива или примена знања, препоручује се да се формулација питања започне речима и синтагмама као што су: „због чега“, „објасните“, „упоредите“, „анализирајте“, „изнесите своју критику“ (Taylor, Greer & Mussio, 1978). Прецизно формулисање захтева подразумева употребу директивног глагола и објекта глагола. На пример (Reiner *et al.*, 2002: 25):

Која истраживачка техника је погодна за испитивање облика рада на часу?

Образложите (директивни глагол) свој одговор (објекат глагола).

(простор за уписивање одговора)

У литератури се саветује да се у формулацији есејског питања избегава глагол „дискутујте“ (или „продискутујте“), управо зато што није довољно

директиван. Захтев се може формулисати у форми питања или императивне изјаве, као у претходно приказаном примеру (Reiner *et al.*, 2002).

Величина простора за одговор углавном одређује дужину одговора који ће испитаник понудити. Такво ограничење је у извесној мери умањено уколико испитаник уписује одговоре на посебно припремљен лист за одговоре. За испитанике који су склони давању опширнијих одговора просторна ограничења превазилазе се писањем на маргини или полеђини хартије, потраживањем додатних листова за одговоре и слично. Потребно је имати у виду да ће одређени испитаници строго поштовати просторно ограничење и у том смислу можда угрозити квалитет свог одговора.

Потребно је указати и на нека уврежена, али погрешна мишљења у вези са есејским питањима. Наиме, постоји уверење да есејска питања увек испитују сложене процесе мишљења, без обзира на то како су формулисана. Међутим, у пракси, есејска питања често испитују само ниво познавања чињеница управо због недовољне посвећености истраживача правилној формулацији питања. Пример таквог питања је следећи:

Које су основне предности есејских питања?

(простор за уписивање одговора)

Такође, погрешно је уверење да ученици не могу да „варају“ на есејским питањима. Уколико се потпуно искључи могућност да ученик препише тачан одговор (коришћењем књиге или „пушкице“), остаје могућност импровизације у одговору, где се вештим избором речи прикрива непознавање градива (Reiner *et al.*, 2002).

Задаци са кратким одговорима. Структуру задатака са кратким одговорима чини стабло задатка, обично у форми питања, и простор за одговор. Простор за одговор се понекад представља кратком цртом. Задатак испитаника је да на

питање одговори уписивањем једне речи, броја, симбола или кратке реченице.
На пример:

Којој истраживачкој техници припада семантички диференцијал?

Адекватно формулисани задаци са кратким одговорима могу испитивати различите нивое знања, релативно лако се конструишу, брзо се решавају и умањују могућност погађања тачног одговора. Ипак, у пракси, успешно решавање задатака са кратким одговорима обично се своди на репродукцију знања (Popham, 2003; *Guide to Assessment*, 2015). Такође, недостаци који прате њихову примену везани су за мањи обухват градива и сложено оцењивање, упркос томе што је тачан одговор једна реч, број, симбол или група речи (Popham, 2003).

Правилно конструисање задатака са кратким одговорима ослања се на следеће препоруке (*Guide to Assessment*, 2015):

- пожељно је да задаци имају форму питања
- потребно је да формулација задатка буде довољно директивна у односу на одговор који се очекује
- пожељно је да задатак упућује испитаника да пружи сажет одговор
- потребно је унапред одредити шта се све може признати као тачан одговор.

Једноставност конструисања задатака са кратким одговорима може се довести у питање ако се има у виду чињеница да се њихова примена обично своди на испитивање нивоа познавања чињеница, иако се тиме не исцрпљује потенцијал који ова форма задатака носи. Истраживач мора да тежи формулацији која искључује могућност да испитаник погрешно протумачи захтев, па самим тим да погрешан одговор, односно да приступи детаљном образлагању свог становишта уместо да понуди прецизан одговор.

Задаци допуњавања. Структуру задатака допуњавања чини започета/непотпуна мисао и простор за уписивање одговора, који је интегрални део започете/непотпуне мисли. Простор за одговор се обично представља цртом, која може имати различиту позицију у реченици. Задатак испитаника је да допуни непотпуну мисао уписивањем речи или синтагме у простор предвиђен за одговор. На пример:

Скуп свих јединица на које се резултати истраживања односе назива се _____.

У литератури се разматра и подврста задатака допуњавања која се означава као *задаци допуњавања повезаног или дискурсивног типа* (Рајстон, Цастман и Робинс, 1966). Следећи пример илуструје дату форму задатака:

Оно што су код упитника питања, код теста знања су _____, а код скале ставова Ликертовог типа _____.

Предност задатака допуњавања је економичност, брзо и лако решавање (*Guide to Assessment*, 2015). Недостаци који прате њихову примену везани су за ограничену примењивост у домену испитивања виших нивоа знања и сложено оцењивање, упркос томе што је тачан одговор најчешће једна реч или синтагма (Popham, 2003). Специфичну тешкоћу код задатака допуњавања представља постизање равнотеже између количине информација које решење чине могућим и количине информација које омогућавају да се одговор „погоди“ (Џордан, 1966).

Правилно конструисање задатака допуњавања ослања се на следеће препоруке (Баковљев, 1997; Popham, 2003):

- пожељно је да изостављена реч/синтагма буде на крају реченице, и да је у питању реч/синтагма која „носи“ смисао реченице
- пожељно је да задатак упућује испитаника да пружи сажет одговор

- потребно је унапред одредити шта се све може признати као тачан одговор.

Састављање задатака допуњавања прате сличне потешкоће као и конструисање задатака са кратким одговорима. Наиме, и овог пута тежња ка састављању добрих задатака подразумева тежњу да непотпуна мисао буде довољно директивна како би испитаник/ученик адекватно разумео захтев, и понудио адекватан одговор. У извесном смислу, захтев директивности је код задатака допуњавања још више наглашен. Истраживач/наставник овог пута мора бити сигуран да је изоставио праву реч/речи, односно реч/речи које носе смисао реченице. Додатно, код задатака допуњавања испитаник мора показати концизност и адекватно језички обликовати одговор.

Структурирана питања. Као посебна врста задатака, у литератури се разматрају и структурирана питања. Њихову структуру чини стабло и низ питања која се изводе из поставке задатка. Стабло најчешће има форму текста, мада се могу користити и графикони и слике (*Guide to Assessment, 2015*), што зависи од садржаја на који се задатак односи. Задатак испитаника је да упише свој одговор у за то предвиђен простор. На пример:

Предмет истраживања гласи: „Школски успех и социјална прихваћености ученика у одељењу“.

Како гласи циљ наведеног истраживања?

Који су кључни појмови у наведеном истраживању?

Какав статус у истраживању има варијабла „социјална прихваћеност ученика“?

На који начин би се могао испитати школски успех ученика у датом истраживању?

Предности и ограничења структурираних питања обично се посматрају у односу на задатке испуњавања налога, јер су њима најсличнија. Као предност у односу на задатке испуњавања налога истиче се већа поузданост у испитивању знања ученика. Образложење се налази у томе што низ питања која прате стабло задатака не дозвољавају испитанику да се удаљи од теме. Такође, квалитет одговора на структурирано питање мање је одређен стилем писања код испитаника (*Guide to Assessment, 2015*). Као и задаци испуњавања налога, структурирана питања се могу користити за испитивање свих нивоа знања, мада су најпогоднија за испитивање нивоа разумевања градива и примене знања.

Као и у случају задатака испуњавања налога, ограничење структурираних питања је то што се односе на мањи део градива. Специфични недостаци структурираних питања су сложеније конструисање и чињеница да погрешан одговор на полазна питања може да води ка неуспешном решавању питања која следе (*Guide to Assessment, 2015*).

Правилно конструисање структурираних питања ослања се на следеће препоруке (*Guide to Assessment, 2015*):

- питања која су саставни део задатака морају логички следити из стабла задатака
- потребно је да формулација питања буде довољно директивна у односу на одговор који се очекује
- пожељно је избегавати питања која испитују ниво познавања чињеница
- потребно је унапред одредити шта се све може признати као тачан одговор.

У методолошком смислу, специфична тешкоћа у обликовању структурираних питања јесте тежња да се низ питања логички изведе из стабла задатака, али да то буду питања која су независна једно од другог. Такав захтев је једини поуздан начин да се избегне околност да погрешан одговор на неко од полазних питања води ка низу погрешних одговора на питања која следе. Ипак, тај циљ није лако остварити у практичној примени дате форме задатака.

Можемо закључити да се за утврђивање нивоа познавања чињеница препоручују задаци: двочланог и вишеструког избора, спаривања, допуњавања и кратких одговора. За испитивање нивоа разумевања и примене знања, препоручују се задаци вишеструког избора, задаци са кратким одговорима, задаци испуњавања налога и структурирана питања. За испитивање највиших нивоа знања (анализа, синтеза и евакуација), препоручују се задаци вишеструког избора, задаци са кратким одговорима, испуњавања налога, структурирана и есејска питања. У редовној наставној пракси, препоручује се да број форми задатака на нивоу теста буде три или четири (Боговац и сар, 1972).

У разматрању препорука које се баве правилним конструисањем задатака одређене форме, приметна је разлика између препорука које имају статус вредности и препорука која се могу емпиријски проверити (Haladyna & Downing, 1989). Препоруке које имају статус вредности, углавном се не проверавају, али имају ширу прихваћеност међу истраживачима у односу на препоруке које се емпиријски проверавају. На пример, препорука да је у формулацији задатка потребно фокусирати се на један проблем, има статус вредности и нико од истраживача који је подржавају нема потребу да је проверава. Испитивања која су емпиријски спроведена до тренутка писања извештаја Халадајне и Даунинга (Haladyna & Downing, 1989) углавном су се бавила испитивањем ефеката одговарајуће методолошке препоруке на тежину задатака у тесту. На пример, емпиријски је проверавана повезаност тежине задатка спаривања и хомогености понуђених података у задатку. На основу добијених процена, одређене препоруке за састављање задатака су биле усвојене, а неке одбачене. Аутори упозоравају да је много важније да се ефекти испитују са становишта валидности, дискриминативности и поузданости задатака, а посебно са становишта утврђивања валидности као најважније карактеристике тестова знања. У том смислу, листе препорука се не могу сматрати коначним. Њихова релевантност посебно је дискутабилна у условима дизајнирања теста применом IRT модела мерења.

Критеријуми за утврђивање и бодовање тачних одговора

Сваки поступак мерења квалитативних варијабли одликује додавање квантитативне ознаке (броја) квалитету који је предмет мерења. Одговори испитаника представљају квалитативне индикаторе испитиване варијабле, па је неопходно одредити критеријуме за њихово означавање бројевима. Први критеријум је утврђивање позитивних одговора, односно одговора који указују на присуство варијабле. Следећи критеријуми су систем бодовања одговора и систем утврђивања укупног скорa на тесту (Буквић, 1996).

Поступак вредновања одговора испитаника на тесту започиње израдом кључа. Кључем за вредновање одговора дефинишу се критеријуми тачних одговора, за сваки појединачни ајтем и задатак у тесту.

Најчешће примењиван начин бодовања одговора у тесту знања је дихотомно скоровање (Буквић, 1996). Пошто се утврди да ли је одговор тачан или нетачан, у систему дихотомног скоровања, тачном одговору се приписује један поен, а нетачном нула поена. Дихотомно скоровање има своје недостатке, међутим нема довољно агрумената који би оспорили његову примену. На пример, јасно је да нису сви задаци у тесту исте тежине. Међутим, уколико би се тежем задатку приписала два бода, а лакшем задатку један, то и даље не би значило да је тежи задатак заиста двоструко јача мера варијабле од претпостављеног лакшег задатка (Буквић, 1996).

У другачијим системима скоровања, кључ за вредновање одговора може предвидети признавање делимично тачних одговора, али и негативне поене за погрешне одговоре. Признавање делимично тачних одговора праћено је низом методолошких дилема, међу којима издвајамо питања: шта се све може признати као делимично тачан одговор; да ли се може признати само половина тачног одговора; шта је мера „половине“ тачног одговора; да ли се може признати више од половине тачног одговора; да ли се може признати мање од половине тачног одговора; колико поена носи одговор који представља више од половине тачног одговора; колико поена носи одговор који представља мање од половине тачног одговора? Како би се утврдио систем

бодовања одговора који ће се применити на тест у целини, неопходно је да се пре задавања теста дају коначни одговори на наведена питања.

Уколико је тест знања сачињен првенствено од задатка двочланог избора увођење негативних поена за погрешне одговоре обезбеђује већу дискриминативност теста знања (Petz, 1978). У свом раду, Пец доказује да је распон поена на таквом тесту знања готово душло већи од распона поена на тесту где нема негативних поена. Образложење се проналази у чињеници да распон поена у околностима када се користе негативни поени више одговара стварном распону знања међу ученицима. Пец препоручује да се приликом решавања теста у коме доминирају задаци алтернативног избора, испитаници упозоре да ће се негативно бодовати погрешни одговори, те да прескоче задатке у чије решење нису сигурни. То, ипак, неће спречити поједине испитанике да погађају одговоре, али се очекује да ће негативно бодовање погрешних одговора анулирати разлику која на тај начин настаје.

Буквић (1996) сматра да увођење негативних поена није увек пожељно, јер може нарушити изглед нормалне дистрибуције скорова. Такође, погрешно би било закључити да негативни поени говоре о томе да испитаник поседује неко својство које потиरे варијаблу која се тестом мери. Негативни поени, заправо, само сведоче о одсуству одређених показатеља варијабле (Taylor, Greer & Mussio, 1978).

У литератури се могу срести и обрасци за израчунавање скорa на тесту којима се урачунава и могућност случајног погађања тачних одговора (Мужић, 1968; Буквић, 1996). Систем корекције узима у обзир следеће чињенице: на колико питања двочланог/вишечланог избора је испитаник тачно одговорио, а на колико је дао погрешан одговор? Формула која се примењује у наведеном случају гласи: а) $R = I - N$ (уколико се тачан одговор бира између два понуђена), односно б) $R = I - N / b_0 - 1$ (уколико се тачан одговор бира између више од два понуђена одговора). У наведеним формулама, R означава коначан резултат, I означава укупан број тачних одговора на питања двочланог избора, N означава укупан број нетачних одговора на питања двочланог избора, а ознака b_0 подразумева број могућих одговора (Мужић, 1968).

Описани систем претпоставља да ће сви испитаници погађати одговоре, што не мора бити случај. Погађање одговора је карактеристика одређених испитаника у ситуацији када се нађу пред задатком који не могу да реше, али свакако не свих (Буквић, 1996). Такође, примена описаног система корекције је ограничена у случају краћих тестова (Џордан, 1966).

Вероватноћа случајног погађања тачних одговора смањује се повећањем броја понуђених алтернатива, тако да се уместо примене наведеног корективног обрасца истраживач може одредити за састављање задатака са вишеструким понуђеним одговорима (Буквић, 1996). У циљу корекције погађања испитаницима се може скренути пажња да не погађају одговоре, као што саветује Пец, или можемо инструкисати испитанике да на сваки задатак понуде одговор чак и када нису сигурни у његову исправност. Најбоље решење за умањење ефеката погађања Фајгел и Кнебл (2004) виде у коришћењу тропараметарског IRT модела (који узима у обзир вероватноћу „погађања“ одговора), јер поједини испитаници ипак неће поступити у складу са инструкцијама.

Поред наведеног система бодовања, у пракси се могу користити и други модели. На пример, Марзано (Marzano, 2006) је развио скалу за оцењивање ученичких одговора на тесту којом се у начину одговарања ученика трага за одговарајућим обрасцем одговора. Образац се сравњује са унапред дефинисаним моделима одговора и спрема тога ученик добија одговарајући број бодова на скали. На пример, 1 бод добија ученик који начином решавања задатака у тесту показује делимично разумевање једноставних чињеница и процеса, неких комплекснијих идеја и процеса. Ученик може добити 3 бода уколико решавајући задатке у тесту не прави озбиљније грешке у вези са идејама и процесима који су експлицитно разматрани на часовима наставе. Ученик добија максимална 4 бода уколико самостално закључује и примењује знање у проблемским ситуацијама које нису разматране у настави.

Систем извођења укупног скорa на тесту зависи од тога да ли су предвиђени негативни поени за погрешне одговоре или не. Уколико нема негативних поена систем је најчешће такав да се само саберу позитивни поени

које је испитаник остварио. Уколико су предвиђени негативни поени, систем може бити тако осмишљен да се од суме позитивних поена одузима сума негативних поена, половина или четвртина укупне суме негативних поена (Буквић, 1996). Одлуку о томе да ли се одузима укупна сума, половина или четвртина укупне суме негативних поена доноси истраживач. Одлука се доноси у односу на број поена који носе у целини тачни одговори. На пример, у целини исправно решен задатак може да носи два поена. Сходно томе, да ли је адекватно да погрешан одговор на задатак носи - 2 поена? Да ли је коректније да то буде само - 1 или, пак, - 0,5 поена? Не би требало изгубити из вида да се испитаник који погрешно у одговору на наведени задатак, само том једном грешком, од максималног броја поена на тесту удаљава у распону од 2,5 до 4 поена.

Када постоји детаљно разрађен кључ за вредновање одговора, оцењивање задатака са бирањем одговора је прилично једноставно. Вредновање одговора на задатаке са уписивањем одговора је сложеније, посебно ако се има у виду чињеница да њихово конструисање ретко прати састављање разрађеног кључа за вредновање одговора. Испитивања показују да више различитих процењивача потпуно различито оцењује један исти одговор на задатак есејског типа. Разлике настају услед неусаглашености процењивача око тога шта се оцењује на нивоу одговора, колико су поједини аспекти одговора важни, како се бодују одговори различитог квалитета (Рајстон, Џастман и Робинс, 1966). Међутим, када се утврде јасни критеријуми вредновања постиже се висок ниво сагласности међу проценама (Џордан, 1966). Имајући то у виду, у литератури се за оцењивање задатака са уписивањем одговора препоручује израда *рубрика* (Rudner & Schafer, 2002; Popham, 2003). Најважнија компонента рубрике су критеријуми вредновања. Критеријуми могу бити стил, организација текста, аргументација, оригиналност. Пошто се одреде критеријуми вредновања, неопходно је дати њихову квалитативну дефиницију. Уколико се, на пример, вреднује организација текста, квалитативна дефиниција помаже да се утврди који типови организације текста ће се сматрати прихватљивим, а који не. Следећи корак у конструисању

рубрике је одабир стратегије оцењивања. Избор је могуће направити између две стратегије – *холистичке* и *аналитичке*. Холистичка стратегија подразумева давање глобалне оцене, односно пружање одговора на питање у којој мери је испитаник испунио постављене критеријуме вредновања (Porham, 2003). У аналитичком приступу, потребно је направити листу могућих одговора и прецизирати колико поена се може добити за сваки од њих. На пример, оцењујући решавање сложеног задатка из математике, наставник може дати ученику један поен за идентификовање кључних променљивих у задатку, један поен за састављање једначине којом се проблем решава и један поен за исправно решавање једначине (Livingston, 2009). Бодови се, потом, сабирају и формирају финални скор. У пракси, чешће је заступљена холистичка стратегија, јер је такав поступак оцењивања бржи (Porham, 2003).

Метријске карактеристике теста знања

Провера метријских карактеристика теста знања

Процес конструкције теста често се означава као итеративан процес (Фајгел, 2004) или процес у коме се више пута понављају одређени кораци како би се остварио постављени циљ. Уколико се као циљ постави конструкција „доброг“ теста знања, вишеструким проверама квалитета теста тежи се постизању задовољавајућих метријских карактеристика валидности, објективности, дискриминативности и релијабилности. Традиционално, у теорији тестирања две основне мерне карактеристике тестова су валидност и релијабилност. По значају и учесталости испитивања, иза валидности и поузданости следи дискриминативност, док је објективност на зачељу (Фајгел, 2003). Ретко испитивање појединих карактеристика теста може се објаснити променом приступа тестовима у области психометрије и све интензивнијим ослањањем на принципе теорије ставског одговора у процесу конструкције тестова.

У циљу провере метријских карактеристика теста знања, неопходно је организовати пробно испитивање. Пробно испитивање подразумева да се тест примени на узорку испитаника који је према кључним својствима максимално сличан популацији којој је тест намењен. Неопходно је да се пробно испитивање реализује према упутствима која ће важити и за финалну примену теста. То је једини начин да се евентуални процедурални недостаци уоче и на време уклоне (Konrad, 1981).

У значајној мери, провера метријских својстава теста остварује се применом одговарајућих статистичких поступака. Поред тога, пробно испитивање омогућава да се без сложеније статистичке анализе уоче одређени недостаци теста као инструмента (на пример, недовољно простора за уписивање одговора у тесту, употреба термина у формулацији задатака који испитаницима нису познати) или задате процедуре (на пример, неодговарајући временски оквир тестирања). Циљеве провере квалитета теста Конрад (Konrad, 1981) сумира на следећи начин:

- идентификовање задатака који су двосмислени или нејасни и нефункционалних дистрактора
- утврђивање индекса тежине сваког задатка
- утврђивање индекса дискриминативности сваког задатка
- одређивање укупног број задатака у тесту
- одређивање временског трајања тестирања
- утврђивање евентуалних слабости процедуре за спровођење тестирања, као и упутстава намењених испитивачима и испитаницима
- израчунавање корелације у успешности решавања задатака ради њиховог организовања у субтестове.

У литератури се разликује и „претходна провера“ која се реализује на малим узорцима (10 до 50 испитаника) са циљем идентификовања крупнијих недостатака теста и планиране процедуре за спровођење тестирања. Обично, претходна провера омогућава истраживачима да издвоје задатке који су двосмислени, неодговарајуће тежине, да се уоче неадекватности у садржају упутстава и прецизније одреди време реализације теста (Konrad, 1981). Може се

приметити да су циљеви примене претходне провере подударни са циљевима пробног испитивања, с том разликом што се код претходне провере сви закључци изводе без статистичке анализе.

Валидност теста. До средине XX века, валидност теста се утврђивала одговором на питање „да ли тест мери својство за чије мерење је конструисан“. Тест се сматрао валидним за испитивање сваког својства које се налазило у корелацији са резултатима теста (Guilford, 1946; према: Shepard, 1993: 410) и такво становиште одржало се све до дефинисања стандарда за валидан тест. Стандарди валидности дефинисани су први пут 1954. године од стране Америчке асоцијације психолога (Anastasi, 1986). Идентификована су четири типа валидности која су се односила на различите циљеве тестирања. Стандарди су били следећи: садржинска валидност, предиктивна валидност, конкурентна валидност и конструктивна валидност. Садржинска валидност утврђивала се за тестове чији циљ је био да опишу постигнуће испитаника на унапред дефинисаном низу задатака. Предиктивна валидност се утврђивала за тестове који су коришћени да предвиде будуће понашање испитаника. Конкурентна валидност се утврђивала када би се уводио нови тест као замена за већ постојећи инструмент мерења (на пример, употреба теста знања са задацима са понуђеним одговорима уместо есеја у настави историје). Конструктивна валидност се утврђивала када је било потребно извести одређене закључке у вези са својствима која нису манифестна, као што је интелигенција или анксиозност. Ревизијом стандарда валидности, 1974. године, предиктивна и конкурентна валидност су преиначене у јединствени стандард – критеријумску валидност (Anastasi, 1986; Shepard, 1993).

У време утврђивања стандарда валидности, конструктивна валидност је дефинисана од стране Кронбаха и Мила (Cronbach & Meehl). Кронбах и Мил су развили концепт „номолошке мреже“, тежећи дефинисању законитих веза између психолошких конструката. На теоријској равни, претпоставили су да постоје везе између конструкта који се мери, других конструката и манифестног понашања. Претпостављене везе било је потребно емпиријски проверити путем корелационих и експерименталних студија. Конструктивна

валидност, дакле, претпостављала је да се одређена предвиђања у вези са својствима која меримо најпре изведу на теоријској равни, а потом проверавају емпиријским путем (Shepard, 1993).

У складу са Кромбаховим учењем, тестови који су објављивани седамдесетих и осамдесетих година XX века конструисани су уз озбиљно разматрање теоријских основа полазних конструката (Anastasi, 1986; Moss, 2007). Међутим, покушаји утемељења валидности теста трагањем за номолошким везама конструката имали су ограничен домет. Последично, крај XX века довешће до тога да се одустане од проверавања валидности теста у односу на неку формалну теорију (Kane, 2001). У поступак провере валидности теста укључују се емпиријске и теоријске процене одрживости *интерпретације резултата теста и последица његове примене*. Таквој промени највише је допринео рад америчког психолога Самјула Месика (Samuel Messick).

Месик проглашава конструктивну валидност јединственим концептом валидности, истовремено проширујући границе валидности изван значења скорова на тесту како би се обухватили релевантност, корисност, вредносне импликације и друштвене последице одлука изведених на основу резултата тестирања (Shepard, 1993). Уз извесне ревизије и проширивања од стране Месикових следбеника (на пример: Kane, Huble, Zumbo), Месиково учење остало је наутицајније у области тумачења валидности и до данашњих дана.

Месиково учење поступак испитивања валидности претвара у поступак аргументовања који чине следећи кораци (Kane, 2001):

- саопштити интерпретацију резултата теста (полазна интерпретација) на што јаснији начин
- утврдити на којим претпоставка почива полазна интерпретација
- вредновати (емпиријски или логички) најпроблематичнију претпоставку у полазној интерпретацији; као резултат вредновања, полазна интерпретација би могла бити одбачена или унапређена или би поступак мерења могао бити унапређен како би интерпретација остала одржива

- трећи корак је потребно понављати док се сви закључки који следе на основу полазне интерпретације не прихвате или док се полазни аргумент не одбаци.

Полазна интерпретација састоји се од два дела: дескриптивног којим се повезују резултати теста са конкретним испитаником и прескриптивног дела који укључује одлуке/акције које се предлажу у односу на претходна сазнања. На пример, дескриптивни део аргумента говори о томе колико је ученик успешно решио тест знања, док прескриптивни део подразумева доношење одлуке о томе који образовни модул да се предложи ученику, с обзиром на постигнуте резултате.

Месикова интенција је била да се освести вредносна основа научних истраживања. Уколико се експлицира вредносни оквир, истраживању се може приступити и са другачије позиције (Messick, 1975; Shepard, 1993). Одговорност истраживача је да предвиди какве последице примена теста може имати у зависности од вредносне позиције са које се полази (Messick, 1981). На пример, какве последице по наставу и учење могу имати погрешно интерпретирани резултати тестова знања? Као што фармацеут мора размотрити све очекиване и нежељене ефекте лека пре него што га означи „безбедним и ефикасним“, исто тако би и конструктори теста требало да дају гаранције за коришћење тестова знања (Shepard, 1993: 426).

Дакле, процена последица које примена теста може да има, подразумева разматрање и очекиваних и неочекиваних исхода тестирања (Messick, 1975; Messick, 1980; Messick, 1981). Ипак, предвиђање исхода могуће је тек онда када се добро разуме номолошка мрежа којој одређени конструкт припада (Messick, 1981). Другим речима, практичар који жели успешно да примени одређени тест, мора добро да познаје теоријску основу конструкције теста. У домену васпитно-образовне праксе, конструкти на које се ослањамо када правимо тест знања јесу васпитно-образовни циљеви, преведени у облике мишљења или понашања који се могу мерити тестом (Anastasi, 1986). Дакле, за конструкцију валидног теста знања у школској пракси, неопходно је добро познавање васпитно-образовних циљева и задатака и индикатора њихове остварености.

Укажимо на још неке оригиналне приступе валидности теста који налазе утемељење у Месиковом учењу. Најпре, процес испитивања валидности теста – није могуће окончати. Сви тестови које конструишемо, од прве примене, до актуелног тренутка, пролазе „континуирани процес валидације“ (Brennan, 2001: 13). Интерпретација резултата добијених тестом подложна је променама у складу са сазнањима до којих долазимо научним истраживањима (Anastasi, 1986). Такође, немогуће је предвидети све последице примене резултата теста, поготову када је реч о последицама које нису очекиване (Messick, 1980). Дакле, валидност није инхерентно својство тестова, већ зависи од начина на који се тест користи. За различите намене, тест ће у различитој мери бити валидан (Cronbach, 1971; према: Shepard, 1980). Сам по себи тест *не може* бити валидан – валидни могу бити само закључци који се изводе на основу резултата теста (Popham, 2001; Popham, 2003).

Испитивање у коме је анализирана научна продукција посвећена питањима валидности теста у периоду од 1999. до 2008, показало је, међутим, да истраживачи врло ретко прибегавају тумачењу валидности теста утемељеној на последицама његовог коришћења (Cizek, Bowen & Church 2010). У научној јавности је присутно мишљење да последице коришћења теста логички не припадају валидности теста. Једноставно, није их могуће уклопити у низ осталих доказа о валидности и добити једну кохерентну целину. Хабли и Замбо (Huble & Zumbo, 2011), пак, сматрају да је разлоге потребно тражити на другој страни. Наиме, аутори истичу да је ретко тумачење валидности теста са становишта последица његовог коришћења више потиче од неразумевања или лоше интерпретације Месиковог становишта, које је само по себи сложено. Чак је и Месик у каснијим радовима покушавао боље да објасни шта је подразумевао под вредносним импликацијама и друштвеним последицама тестирања. Хабли и Замбо (2011) подржавају Месикове идеје, подсећајући да све што чинимо у процесу израде и примене теста знања – од развоја конструката, преко мерења, задавања теста до интерпретације добијених резултата – одражава наш систем вредности. Новина којом аутори употпуњују Месиково

учење је да се међу последицама примене теста могу разликовати не само друштвене већ и очекиване и неочекиване последице на персоналној равни.

Критика Месиковог учења у доба постмодернизма изнедрила је поделу одговорности за утврђивање различитих аспеката валидности теста на следећи начин (Gipps, 2003):

- особе које састављају тест морају прецизно дефинисати конструкт на који се тест односи, као и услове адекватне употребе теста
- творци образовне политике у друштву морају размотрити последице употребе теста
- представници установа које објављују резултате теста обавезне су да резултате ученика представе објективно и опишу контекст у коме се одвија наставни процес.

Наведена подела одговорности, понекад је само условна. Подела на „састављаче“ и „кориснике“ теста често није могућа, као, на пример, у ситуацији када тестове састављају управо они који ће их касније примењивати (Kane, 2001).

Питања валидности теста и даље привлаче пажњу истраживача. Последњих година актуелни су покушаји да се поједностави приступ утврђивању валидности. На пример, Лисиц и Самјулсен (Lissitz & Samuelsen, 2007) уводе појам *интерне валидности теста* коју чине садржинска валидност теста и релијабилност. Домену *екстерне валидности* припадали би сви други аспекти валидности, међу које се сврстава и тумачење резултата теста и последице њихове примене. Аутори наглашавају да је за процену квалитета теста утврђивање интерне валидности, ипак, од пресудне важности.

Пофам (2003) разматра могућности да се валидност теста у наставној пракси испитује на нивоу критеријумске, садржинске и конструктивне валидности. Када је реч о критеријумској валидности, Пофам истиче да је ова врста валидности теста мање релевантна за редовни наставни рад. Такође, у редовном раду, наставници немају могућности да испитују конструктивну валидност теста због сложености самог поступка (Taylor, Greer & Mussio, 1978; Popham, 2003). За разлику од критеријумске и конструктивне валидности,

садржинска валидност теста је релевантна у свакој ситуацији испитивања, па самим тим и у ситуацији када наставник самостално конструише тест знања.

Садржинска валидност се одређује дедуктивно, тако што се најпре одреди *универзум задатака*, а потом се на основу универзума формира *узорак задатака* за конкретан тест (Фајгељ, 2005). Универзум задатака може се дефинисати на основу теорије, наставног програма, садржаја уџбеника, практичних сазнања или *a priori*. Један од битних чинилаца валидности теста знања је правилна заступљеност наставних целина у тесту. То значи да све целине које су дефинисане нацртом теста морају бити покривене задацима. Такође, потребно је да број задатака у свакој целини одговара обиму и значају наставне целине. У школској пракси, садржинску валидност утврђује наставник, субјективним просуђивањем, односно логичко-психолошком анализом усаглашености узорка и универзума задатака (Буквић, 1996; Фајгељ, 2005).

Специфично за утврђивање садржинске валидности теста је и то да се она одређује пре и током конструкције теста, а не након што је тест конципиран. За тестове постигнућа то је од суштинске важности. Говорећи о квалитету тестова постигнућа, Фајгељ упозорава (2005: 337): „Ништа неће помоћи њиховом квалитету ако се ваљаност проверава након, рецимо, пријемног испита, када су елиминација и селекција већ завршени.“ Дедуктивна анализа и експертска процена обично су довољне да би се валидирао универзум задатака. Ипак, за проверу избора конкретних задатака, њихове формулације, избора дистрактора, потребно је да дедуктивна анализа буде допуњена емпиријском анализом одговора испитаника (Фајгељ, 2005).

Пофам (2003) оцењује да процена садржинске валидности није лак задатак и препоручује да се утврђује само у ситуацији припреме полугодишњег или годишњег теста. Такође, наставницима саветује да затраже још нечије мишљење, како си се добила што прецизнија процена садржинске валидности (Porham, 2003). Може се закључити да је за истраживаче важно да валидност теста процене са становишта критеријумске, конструктивне и

садржинске валидности, док је за наставнике довољно да испитају садржинску валидност теста.

Напоменимо на крају да се у класичној теорији тестирања валидност испитује применом одговарајућих статистичких поступака. Између осталих поступака, користе се линеарна корелација (r_{xx} , r_{xy} , при чему „ xx “ и „ xy “ означава стварне скорове делова теста), бисеријална и поинт-бисеријална корелација (r_{bis} , $r_{p.bis}$) (Буквић, 1996). У теорији ставског одговора, процена валидности теста се остварује израчунавањем показатеља фитовања модела мерења подацима који се добијају од испитаника. Ипак, докази о фитовању модела не искључују потребу за калсичном валидацијом теста (Фајгел, 2005).

Поузданост теста. Поузданост или релијабилност представља метријску карактеристику теста којом се упоређују вредности мерења истог својства, истим тестом у поновљеним ситуацијама мерења. Другим речима, поузданост указује на степен прецизности мерења варијабле помоћу одговарајућег теста. Варирање скорова може бити последица деловања постојаних фактора, што се може пратити правилним варирањем скорова у поновљеним мерењима. Када таква правилност у одступањима није уочљива, говори се о деловању случајних грешака. Чиниоци који изазивају систематске грешке, уколико се препознају, могу бити уклоњени. Премда није увек могуће препознати све систематске факторе, откривање и неутралисање случајних грешака представља још озбиљнији проблем (Буквић, 1996).

Извори грешака у мерењу тестом најчешће потичу од узорка испитаника, карактеристика самог теста и одређених начина реализације поступка тестирања (Буквић, 1996). Испитаници, на пример, могу бити мање или више мотивисани да решавају тест, проблем може бити и њихова психофизичка спремност, емоционално стање у току тестирања, умешност у решавању задатака у тесту и томе слично. Млађи испитаници обично журе да што пре попуне тест, док су одрасли испитаници опрезнији и спорије решавају тест. Извор грешака у мерењу може бити и склоност испитаника да погађају тачне одговоре у задацима са понуђеним алтернативама (Буквић, 1996).

Грешке мерења чији се извор крије у самом тесту најчешће настају услед пристрасног избора индикатора варијабле, а самим тим и узорка задатака. На пример, може се десити да у једном тесту преовладавају задаци којима се испитује ниво познавања чињеница, или ниво примене знања. Зависно од наклоности састављача теста према једној или другој врсти задатака, тестом се може добити погрешна слика о знању појединих испитаника у зависности од тога на који начин су учили градиво (Буквић, 1996). Извор грешака могу бити и лоше одабрани дистрактори у задацима са вишеструким одговорима (Rudner & Schafer, 2002).

Грешке које настају у примени теста везују се за неправилно поступање испитивача, односно других особа које су укључене у процес тестирања. Најчешће је реч о изненадним сметњама у току тестирања, прокорачењу или скраћењу времена за рад, нејасним упутствима, договарању испитаника током рада, погрешном тумачењу поставке задатка или упутства за рад и слично. Грешке се дешавају и када оцењивачима није јасно шта се може признати као тачан одговор, када постоје пристрасни системи бодовања и израчунавања укупног скорa на тесту (Буквић, 1996; Rudner & Schafer, 2002).

Релијабилност теста се може испитати на више начина. Први начин се остварује понављањем тестирања применом истог теста са истим испитаницима после одређеног временског периода. По завршетку другог тестирања израчунава се корелација између вредности скорова остварених на тестовима у две ситуације примене. Што је корелација већа, тест је поузданији. Коефицијент корелације који се у датим околностима израчунава, Мужих (1986) назива „коефицијентом стабилности“. Слично, Пофам (2003) описану врсту поузданости теста назива „стабилна поузданост“ (*stability reliability*).

Када је реч о временском периоду који је потребно да протекне између два тестирања, његова дужина одређена је узрастом испитаника, карактеристикама варијабле која је предмет мерења и карактеристикама самог теста (Буквић, 1996). Сувише кратак период олакшава преношење претходно стеченог искуства. Предуг временски размак између тестирања доноси ризик да у међувремену испитаници буду изложени чиниоцима који могу изменити

варијаблу која се мери. Такође, што више времена пролази, то се и узорак испитаника више осипа, па се може десити да у ситуацији задавања ретеста немамо довољан број испитаника. Имајући у виду претходно наведено, код млађих испитаника препоручује се да временски размак између два тестирања не буде велики. За тестове који су састављени од мањег броја задатака и тестове чије су ставке занимљиве и лаке за памћење, временско растојање између два тестирања би требало да буде дуже. У пракси, ретест се задаје у периоду од једног до шест месеци (Буквић, 1996).

Други начин је да се конструишу два паралелна теста и да се примене на истој групи испитаника. Као и у претходном случају, након тестирања испитује се корелација између вредности скорова остварених у две ситуације тестирања. Коефицијент корелације који се у датим околностима израчунава, Мужих (1986) назива „коефицијентом еквивалентности“. Слично, Пофам (2003) описану врсту поузданости теста назива поузданошћу алтернативних форми (*alternate-form reliability*).

Паралелне форме теста су сличне по свом садржају, сваки ајтем у једној форми теста мора имати одговарајући ајтем у другој форми теста. Паралелни ајтеми би требало довољно да личе једни на друге, али истовремено и да буду довољно различити. Самим тим, израда паралелних форми тестова захтева доста времена и ангажовања од стране истраживача. У односу на претходно описани тест-ретест поступак, рад са паралелним формама тестова има одређене предности. Најпре, временски интервал између задавања паралелних форми теста не мора да буде превише дуг, јер се ради о тестовима који нису идентични. Такође, овог пута нема бојазни да су испитаници могли да запамте задатке или да сазнају решења задатака који би се ретест поступком морали поновити (Буквић, 1996).

Трећи начин утврђивања поузданости теста је утврђивање конзистентности ајтема унутар једног теста. Наиме, тест се може поделити на две или више целина међу којима се утврђује корелација скорова које испитаници остварују. Коефицијент корелације који се у датим околностима израчунава, Мужих (1986) назива „коефицијентом хомогености“, док Пофам

(2003) описану врсту поузданости теста назива „унутрашња доследност“ (*internal consistency reliability*). У пракси је овај поступак утврђивања поузданости теста најзаступљенији, јер не захтева реализацију два одвојена тестирања.

За тестове знања код којих су задаци поређани по тежини, препоручује се да се подела теста на два дела оствари тако што се у једну половину сврстају задаци за парним, а у другу половину задаци са непарним бројем. Код тестова код којих задаци нису поређани по тежини, делови се сачињавају на основу квалитативне анализе задатака и на основу анализе њихових мерних карактеристика (Буквић, 1996).

Уместо дељења теста на два или неколико делова, тест се може поделити на онолико делова колико има ајтема. У том случају, сваки ајтем се посматра као „паралелна форма“ (Буквић, 1996: 123), а поузданост теста се може израчунати на основу Сприман-Брауновог обрасца. За тест који има 50 ајтема, поменути образац предвиђа израчунавање 1225 интеркорелација (према формули $n^2 - n / 2$, при чему „n“ означава број ставки у тесту). Како би се израчунавање поузданости упростило, изведени су алтернативни обрасци међу којима се највише користе Кудер-Ричардсон обрасци (K-R-20 и K-R-21)⁸ и Кромбахов алфа коефицијент (r_{xx}). Светски стандард степена корелације за тестове који имају врло високу поузданост требало би да се креће изнад 0.90, а за тестове високе поузданости изнад 0.80 (Буквић, 1996; Момировић, Wolf и Поповић, 1999; Rudner & Schafer, 2002).

У школској пракси, према становишту Пофама (2003) готово да нема потребе за проверавањем релијабилности теста. Тестови које наставници сами праве могу се показати непоузданим, али је наставник у ситуацији да брзо реагује и да одговарајућим мерама надомести неправилности до којих је дошло током тестирања (Rudner & Schafer, 2002). Са друге стране, тестови који се користе у истраживачке сврхе морају бити поуздани, јер скорови на непоузданом тесту неће омогућити да се изведу валидни закључци.

⁸ Називи образаца формиран су спајањем почетних слова Кромбаховог и Ригардсоновог презимена и редног броја једначине у одговарајућем математичком поступку (Буквић, 1996).

Објективност теста. Објективност се може дефинисати као степен сагласности различитих оцењивача у оцењивању исхода тестирања групе испитаника. Сагласност се утврђује израчунавањем коефицијента линеарне корелације између скорова различитих оцењивача. Када тестове оцењују више од два испитивача, неопходно је израчунати све могуће коефицијенте корелације између скорова, и потом утврдити њихову просечну вредност. Потпуно објективан тест требало би да има коефицијент корелације у вредности 1.00. Дата вредност коефицијента корелације означава да су оцењивачи тестове оценили на истоветан начин (Буквић, 1996). Објективност теста се може утврдити и помоћу анализе значајности разлика аритметичких средина скорова различитих оцењивача. Уколико се утврди да је разлика аритметичких средина статистички значајна, тест није објективан. Тест се може прихватити као потпуно објективан када су аритметичке средине скорова оцењивача истоветне (Буквић, 1996).

У приказаним приступима, објективност је схваћена као интерсубјективна сагласност оцењивача теста. Данас се на исти начин дефинише и поузданост оцењивача и све је присутнија тенденција да се сагласност оцењивача означи својством поузданости, а не објективности (Фајгељ, 2003).

Последично, у савременој методолошкој литератури, традиционално схваћена објективност више се не разматра у контексту метријских карактеристика теста (на пример: Popham, 2001; Popham, 2003). Данас када се говори о тесту знања, готово увек се мисли на тестове чији је поступак реализације, оцењивања и тумачења добијених резултата стандардизован (Turner *et al.*, 2001). Дата околност искључује било какву могућност да тест знања буде оцењен на више различитих начина. Другим речима, постоји само један начин, дефинисан кључем за оцењивање теста. Кључ се израђује унапред и обавезујући је за оцењивача теста. Када год је могуће, оцењивање теста спроводи се машински. Код електронских тестова на пример, оцењивање теста врши се применом рачунара. Када је реч о теорији ставског одговора, сматра се

да је објективност „уграђена“ у Рашов модел и да су сви подаци који се уклопе у теоријски модел аутоматски објективни (Фајгељ, 2003).

Осетљивост теста и дискриминативност ајтема/задатака. Осетљивост теста указује на то до које мере тест може направити разлику између испитаника у погледу поседовања мереног својства. У класичној теорији тестирања, осетљивост теста детерминисана је бројем и квалитетом ајтема у тесту, као и начином оцењивања исхода тестирања. Сваки од наведених елемената доприноси осетљивости теста у одређеном степену, али најчешће није могуће утврдити који елемент доприноси више од осталих у конкретној ситуацији тестирања. У начелу, методолози су сагласни да осетљивост теста расте са повећањем броја ајтема у тесту (Буквић, 1996).

Ипак, ајтеми у тесту не утичу подједнако на осетљивост теста, већ у зависности од сопствене дискриминативности. На пример, у тесту знања може се појавити задатак који успешно решавају сви испитаници, или се може појавити задатак који ниједан испитаник није успео да реши. Примери наведених задатака не доприносе осетљивости теста зато што не омогућавају разликовање испитаника који поседују одговарајући индикатор варијабле од оних испитаника који га не поседују (Буквић, 1996). У том смислу, важно је нагласити да само повећање броја ајтема у тесту независно од њиховог квалитета, неће нужно допринети повећању осетљивости теста.

Огледало осетљивости теста представља расподела скорова. Осетљивији тест карактерише веће распршење резултата теста, у односу на мање осетљив тест (Буквић, 1996). Управо из тог разлога, у испитивањима осетљивости теста користе се статистички поступци којима се утврђује одступање добијене расподеле скорова од одговарајуће теоријске расподеле (Буквић, 1996). То су: хи-квадрат тест (χ^2), Смирнов-Колмогорљев тест и метод ДеКарла (Фајгељ, 2003). Поставља се питање, шта се може учинити са тестом чија расподела скорова одступа од нормалне расподеле. Уколико је одступање заиста велико, закључује се да је узрок у самом тесту, па је неопходно преиспитати поступак састављања и одабирања ајтема. Код мањих одступања приступа се „нормализацији“ дистрибуције или баждарењу теста (Буквић, 1996).

У савременој методолошкој литератури, а посебно у теорији ставског одговора, фокус је померен са испитивања осетљивости теста на утврђивање дискриминативности ајтема, и додатно, тежине ајтема.

У класичној теорији тестирања, индекс тежине једног задатка добија се дељењем броја испитаника који су тачно одговорили на задатак (Т) са укупним бројем испитаника у узорку (N) ($p=T/N$). Што је вредност „p“ мања, утолико је задатак тежи. На основу индекса тежине задатака задаци се могу поређати у један праволинијски низ. Пошто нормална дистрибуција скорова на тесту није праволинијска, индекс тежине потребно је изразити у јединицама нормалне расподеле скорова („z“ вредности). Избор задатака за коначну верзију теста по критеријуму тежине могуће је остварити тек када се „p“ вредности изразе у „z“ величинама нормалне криве (Буквић, 1996).

У класичним моделима тестирања, тестови имају најповољније метријске карактеристике када је просечна тежина задатака 50%. Из тог разлога, у коначној верзији теста знања требало би да доминирају задаци средње тежине (Буквић, 1996).

У IRT моделима мерења тежина ајтема утврђује се калибрацијом теста коју је први пут представио Бирнбаум, 1968. године. Убрзо по представљању, техника калибрације је имплементирана у различите компјутерске програме, тако да се читав поступак изводи уз помоћ рачунара (Вакер, 2001). Калибрација теста се остварује тако што се тест зада одређеној групи испитаника и дихотомно (тачно-нетачно) се скорују одговори које они дају. Утврђује се дистрибуција успешности у решавању задатака у тесту. Након тога се примењују одговарајући математички поступци како би се креирала скала мереног својства сходно одговорима испитаника. Потом се вредности ајтем-параметара и заступљеност мереног својства код испитаника изражавају јединицама исте мерне скале. Конкретно, уколико испитаник успешно решава само лакше задатке у тесту знања, процена је да поседује слабије знање у односу на испитаника који решава и задатке веће тежине.

Једнодимензионалност теста. И класична теорија тестирања и теорија ставског одговора на својим почецима, полазиле су од претпоставке о

једнодимензионалности теста. Сматрало се да скор на тесту може имати смисла једино уколико ајтеми у тесту мере само једно и увек исто својство. Уколико би било другачије, односно, уколико би различити ајтеми мерили различита својства, тешко би било интерпретирати скор на тесту, или упоредити резултате различитих испитаника (Hattie, Jaeger & Bond, 1999).

Једнодимензионалност теста је угрожена уколико је тест пристрасан. Пристрасност најчешће није у вези са конструктом који се испитује. На пример, уколико се тестом испитује знање из математике, важно је да формулација задатака у тесту буде таква да је подједнако могу разумети сви испитаници, независно од тога са ког говорног или дијалекатског подручја потичу. Уколико би формулација збуњивала одређене испитанике, њихова успешност у решавању задатка била би умањена, а тест би се могао означити пристрасним (Hattie, Jaeger & Bond, 1999).

Пристрасан тест, дакле, не даје испитаницима једнаке шансе да остваре добар резултат. Након решавања пристрасног теста, разлике у постигнућу се могу очитати међу испитаницима различитог пола, расе, етничке, религијске припадности или социоекономског статуса (Popham, 2003). Разлике настају услед деловања одређених чинилаца на начин на који испитаници одговарају на задатке у тесту. Међу такве чиниоце сврставају се језички стил, дијалекат, културне вредности (Green, 1976). На пример, у истраживању Грина (Green, 1976), установљено је да се разлике у одговорима припадника различитих друштвених група крећу у распону од 20% до 50%.

Последњих година, у реализацији обимнијих испитивања тестовима знања, формирају се тимови стручњака који процењују пристрасност ајтема у тесту. У Америци, такви тимови броје од 15 до 25 чланова, припадника мањинских група. Сваки члан тима процењује појединачне ајтеме у тесту, а из даљег испитивања се искључује сваки ајтем за који 5 или више процената испитаника утврди да је пристрасан. Процењује се да је пристрасност много више заступљена у тестовима које наставници самостално конструишу (Popham, 2003).

Још крајем прошлог века развијени су статистички поступци путем којих је било могуће испитати да ли је тест заиста једнодимензионалан (Bejar, 1980). Међу њима су *MH* стратегија (*Mantel-Haenszel strategy*), *SIB* стратегија (*Shealy & Stout's SIB*), *SIBTEST* (*Simultaneous Item Bias Test*) и *LR* (*Logistic Regression*), при чему су *SIBTEST* и *LR* осетљивији на потенцијалну пристрасност теста (Gierl, Khaliq & Boughton, 1999). Екерман (Ackerman, 1992) упозорава да и када је спроведена калибрација теста, тиме се још увек не гарантује једнодимензионалност теста и потребно је приступити додатним проверама.

Скорашњи развој теорије ставског одговора резултирао је конструисањем мултидимензионалних модела теста (*MIRT* модели) који омогућавају да се више својстава испитује једним истим тестом са задовољавајућим метријским карактеристикама (Фајгељ, 2003).

Обликовање коначне верзије теста знања

Обликовање теста је важно методолошко питање којим се обухватају питања као што су: одређење дужине теста, утврђивање финалног редоследа задатака у тесту, обликовање упутства за испитанике и реализаторе тестирања.

Утврђивање дужине теста може се састојати у одређењу броја задатака које би тест требало да садржи, али и у прецизирању времена које је потребно за решавање теста у целини. Реч је о два критеријума који су међузависни, при чему се један од критеријума обично одређује као полазни.

Уколико се постави питање који је оптималан број задатака у тесту, одговор који на то питање пружа теорија мерења је да је тест потребно употпуњавати задацима све док не стекне задовољавајућа метријска својства (Фајгељ, 2003). У полазним фазама конструкције теста, саставља се значајно више задатака него што ће бити потребно, јер ће након провере метријских карактеристика задатака одређени број бити искључен из даље обраде. У начелу, код класичних тестова поузданост и ваљаност су пропорционалне

дужини теста, док тај услов престаје да важи у рачунарском адаптивном тестирању.

Уколико је предвиђено да се на основу резултата теста знања донесу одлуке које су важне за испитаника, тест би требало да садржи већи број задатака, како би био што поузданији. Такође, ако су задаци теста хомогени, односно вреднују једну област знања или само један облик понашања, њихов број може бити мањи него код хетерогених задатака када се вреднује више наставних области и више облика понашања (Богавац и сар, 1972).

У педагошкој пракси, препоручује се да се за тестирање планира време које је довољно да 90% ученика стигне до последњег задатка (Мужић, 1968). Ипак, прецизна процена колико је времена ученицима потребно за решавање појединих задатака, односно теста у целини, може се добити тек након реализације пробног испитивања (Богавац и сар, 1972).

Што је обради градива посвећено више времена у настави, то ће тест садржати више задатака и тестирање ће дуже трајати. Када је реч о тежини и нивоу знања који се задацима мере, одговор у вези са дужином решавања задатака није могуће дати независно од форме самог задатака. На пример, очекује се да решавање низа задатака на нивоу познавања чињеница неће захтевати онолико времена колико решавање низа задатака на нивоу примене знања. Ипак, то не мора нужно бити тако. На пример, задатак на нивоу познавања чињеница може бити изражен у форми структурираног задатка који ће захтевати дуже време решавања него задатак вишеструког избора који испитује примену знања.

Истраживање Келоу и Келоу (Kellough & Kellough, 2003; према: Walker, 2006), показало је колико је времена потребно ученику просечног постигнућа да одговори на задатак одређене форме. Добијени су следећи резултати (табела 5):

Табела 5. Време решавања задатака одређене форме

Форма задатка	Време потребно за решавање
задаци вишеструког избора	1 минут
задаци допуњавања	30 секунди
задаци са кратким одговорима	2-3 минуте
есејска питања	10 и више минута
задаци сређивања	2-3 минуте
задаци спаривања	30 секунди по пару

Дакле, задаци који се најбрже решавају јесу задаци допуњавања и задаци са вишеструким одговорима. Њихова конструкција, напротив није једноставна. Ученицима је највише времена потребно за решавање задатака есејског типа, чија конструкција, пак, обично не траје дуго. Остале врсте задатака се решавају у приближно истом временском распону, са потенцијалним изузетком задатака спаривања код којих је време решавања одређено бројем парова које је потребно формирати.

Пошто се селекују задаци који према својим метријским карактеристикама могу представљати саставни део теста знања, потребно је утврдити њихов редослед (Vaughn, 1981). Методолошки захтев за утврђивањем редоследа задатака претпоставља да се задаци међусобно разликују. Разлике могу да се огледају у садржају, тежини и форми задатака. У зависности од наведених карактеристика, одређени задаци требало би да претходе другим задацима у тесту или да се са сличним задацима групишу у одговарајуће субтестове.

Са становишта стандардизације услова тестирања, важно је неутралисати деловање спољних чинилаца на постигнуће ученика на тесту, а неадекватан редослед задатака се може посматрати као реметилачки чинилац. На који начин? Уколико се редоследу задатака у тесту не би посвећивала пажња могло би да се догоди да се испитаник већ на почетку теста суочи са задатком који не може да реши. Последично, испитаник може да изгуби

самопоуздање и мотивацију у даљем процесу рада и да оствари слабији резултат него што би то био случај да је најпре решавао лакше задатке. Такође, може се десити да услед насумичног смењивања различитих форми задатака испитаник потроши драгоцену време тумачећи у којој форми да саопшти одговор, на штету времена потребног за решавање задатка.

Задаци на нивоу теста могу бити уређени према редоследу градива и према форми задатака (Taylor, Greer & Mussio, 1978). Оба решења имају своје предности и ограничења. Уређивање задатака према редоследу градива одликују следеће предности: решавање задатака је лакше када су садржаји логички уређени, прегледом теста лакше се може пратити савладаност одређених делова градива; задаци унутар једне целине могу се распоредити према тежини – од лакших ка тежим, што погодује јачању самопоуздања код испитаника током решавања теста (Taylor, Greer & Mussio, 1978). Истовремено, наведени начин уређивања задатака у тесту има следеће слабости: погодан је за испитанике који памте градиво као низове информација и захтева додатно мисаоно ангажовање због промене упутства за рад у току решавања теста (Савовић и сар, 2006).

Уређивање према облику задатака има следеће предности: упутство за решавање групе задатака је заједничко и штеди се време потребно за читање упутства за решавање групе задатака; пошто се не мења начин одговарања, смањује се и могућност грешке приликом решавања задатака; испитаник приступа информацијама као релативно независним и може да их комбинује на нови начин; рационалније је техничко уређивање теста и олакшано бодовање одговора. Недостаци овог модела огледају се у отежаном праћењу савладаности одређених делова градива приликом прегледања теста и повећаном мисаоном ангажовању испитаника при прелажењу са једног садржаја на други (Савовић и сар, 2006).

У погледу уређивања задатака на нивоу странице теста, наглашена је потреба да се испитанику читање и анализа садржаја задатка учине што једноставнијим. Самим тим, потребно је водити рачуна да се садржај једног задатка не дели на две странице, да графикони и слике претходе тексту који се

на њих односи, како би се обезбедило да се понуђене алтернативе надовезују на стабло задатка. Такође, код задатака са понуђеним одговорима препоручује се да се одговори понуде у колони (један испод другог), радије него у продужетку реченице (Taylor, Greer & Mussio, 1978).

Пошто се одреди дужина теста и редослед задатака у тесту, приступа се финалном обликовању упутства за решавање теста и реализацију поступка тестирања. Стандардизација услова тестирања подразумева да се упутства припреме како за испитанике, тако и за реализаторе тестирања, што је посебно наглашено у околностима када реализатори тестирања нису истовремено и конструктори теста.

У зависности од степена образовања испитаника, врсте теста и сложености поступка решавања појединачних типова задатака, упутства о начину рада се могу саопштити усмено, одштампати на самом тесту или презентовати на демонстрационој табли (Буквић, 2007). Код тестова намењених млађим испитаницима, пожељно је да иза упутства следи и конкретан пример како се решавају задаци сложеније форме, као што су задаци спаривања или сређивања (Taylor, Greer & Mussio, 1978).

Указивање на временски оквир тестирања важно је како би испитаници проценили колико времена могу да планирају за решавање конкретних задатака, у зависности од њихове тежине и форме. Пожељно је да испитаник пре решавања прегледа све задатке, како би препознао задатке који су лакши и чије решавање неће захтевати пуно времена, у односу на задатке који су тежи, или чија форма подразумева дужи процес решавања.

Упутства за реализаторе тестирања предвиђају све активности о којима је потребно да истраживач води рачуна, као и указивање на недозвољене активности. Активности које наведена упутства најчешће предвиђају су: проверити да ли испитаници имају сав потребан материјал за рад; поделити тестове; издвојити време да испитаници упишу своја имена и остале потребне личне податке; прочитати упутство за рад; дати знак за почетак рада и завршетак рада; пратити рад испитаника; сакупити тестове (Мужић, 1968).

ПРИМЕНА ТЕСТА ЗНАЊА У ШКОЛСКОЈ ПРАКСИ

Специфичности процеса конструкције тестова знања које наставници самостално израђују

Досадашња истраживања у домену примене тестова знања у образовној пракси, углавном су се односила на стандардизоване тестове знања. Врло мало пажње посвећено је испитивању природе и квалитета тестова које наставници самостално израђују. Разлог томе је дуга традиција научних расправа о тестовима знања и примена психометријских модела мерења у домену васпитања и образовања (Stiggins & Bridgeford, 1985). Таква традиција води ка опасности да се не сагледају исправно потребе наставника када је у питању конструкција тестова. У којој мери је важно да се таква пракса промени, сведоче испитивања која показују да су тестови које наставници самостално израђују врста тестова коју наставници најчешће користе (Бјекић и Папић, 2005).

Како испитивања показују (Kellough & Kellough, 2003; према: Walker, 2006) када приступају конструкцији тестова знања, наставници теже да: тестом обухвате различите типове задатака; језички коректно обликују тест како би ученици разумели захтеве; варирају захтеве на нивоу задатка према сложености; у техничком смислу, тест коректно припреме (без граматичких и словних грешака, са довољно остављеног простора за уписивање одговора). Иако су наведена стремљења важна за квалитет теста, она не исцрпљују листу свих захтева о којима је потребно водити рачуна када је реч о конструкцији теста. Конструкција теста која би била утемељена на актуелним сазнањима методологије педагошких истраживања захтевала би да се размотре и друге фазе о којима је нешто више речи било на претходним страницама: утврђивање циља који се жели постићи применом теста знања, одређивање садржаја теста, одређивање дужине теста, утврђивање редоследа задатака, израда кључа за оцењивање теста знања, утврђивање начина оцењивања теста,

састављање упутства за решавање теста, пробно испитивање и формирање, чување и развијање базе задатака. Неоспорно је да би методолошке захтеве требало „пренети“ из теорије у праксу, али то није могуће учинити без остатка. Конструкцију тестова који налазе примену у наставној пракси одликује низ специфичности на које ћемо указати у наставку текста.

Као што смо претходно видели, израда теста почиње одређењем циља примене теста знања. У школској пракси, циљ теста може бити оцењивање ученика, утврђивање претходног знања ученика, планирање и предузимање мера подршке ученицима у учењу, планирање рада у настави и слично. Наставницима је битно да циљ теста дефинишу довољно прецизно како би се са што мање потешкоћа одговорило на захтеве осталих фаза конструкције теста. Проблем који се такође доводи у везу са одређењем циља теста је да ли тест може истовремено имати више циљева? На пример, наставник може да дефинише циљ теста истовремено као оцењивање ученика, и као идентификовање области у којима слабији ученици имају највише потешкоћа у учењу. Да ли је то адекватно? Ако се исправно конструише тест намењен оцењивању ученика, у њему ће доминирати задаци средње тежине. Поуздано утврђивање области у којима слабији ученици имају највише тешкоћа, захтевало би да тест буде другачије конструисан, односно да у њему буде више лакших задатака (Crocker & Algina, 1986). Самим тим, највероватније неће бити остварена оба постављена циља. Уколико се поставља више циљева истовремено, важно је проценити колико су циљеви међусобно усаглашени, и да ли се неком од њих може дати предност.

У правцу остваривања наредних фаза конструкције теста, којима се утврђује садржај теста, ниво знања који се задатком мери, тежина и форма тестовских задатака, али и прелиминарна верзија кључа за оцењивање теста знања, у литератури се препоручује ослањање на *табеле спецификације*. Табеле спецификације обично садрже више колона и редова. Најчешће, у редовима се прецизирају наставне области које се тестом испитују (Богавец и сар, 1972; Taylor, Greer & Mussio, 1978; Crocker & Algina, 1986; Izard, 2005; Fives & DiDonato-Barnes, 2013). У колонама табеле прецизира се ниво знања који се у

вези са одговарајућим садржајем испитује (Богавец и сар, 1972; Crocker & Algina, 1986; Izard, 2005) (табела 6).

Табела 6. Табела спецификације⁹

Садржај	Ниво знања			
	Познавање чињеница	Поступци израчунавања	Разумевање	Укупно задатака
дистрибуција фреквенција	2 задатка	/	4 задатка	6
аритметичка средина	2 задатка	4 задатка	2 задатка	8
варијанса	2 задатка	4 задатка	2 задатка	8
корелација	4 задатка	4 задатка	12 задатака	20
укупно	10 задатака	12 задатака	20 задатка	42

Поједини аутори предвиђају да се у посебним колонама прецизира време посвећено обради одговарајућег садржаја у настави. Сходно времену повећеном обради одговарајућег садржаја у настави, одређује се заступљеност одговарајућег садржаја на нивоу теста (табела 7). На пример, уколико је одређеном наставном садржају посвећено 10% времена у настави, онда је у тесту са 10 задатака потребно предвидети један задатак који се односи на ту тему (Fives & DiDonato-Barnes, 2013).

Табела 7. Део табеле спецификације¹⁰

Наставни садржај	Време посвећено датом садржају у настави (у минутима)	Процент времена посвећеног датом садржају у настави у току полугодишта	Број задатака	Нижи нивои знања (познавање чињеница и разумевање)	Виши нивои знања (примена, анализа, синтеза, евалуација)
садржај 1	15	10%	1	задатак са кратким одговором	/
садржај 2	10	6,7%	0.7	ајтем вишеструког бирања одговора	/
садржај 3	20	13,3%	1.3	/	задатак и ајтем виш. бир. одговора

⁹ Адаптирано према: Izard, 2005: 34.

¹⁰ Адаптирано према: Fives & DiDonato-Barnes, 2013: 5.

Како Пофам (2003) истиче, у изради табеле спецификације, потребно је водити рачуна и да тежина задатака буде адекватно варирана. Уколико се у табели спецификације разматрају и питања тежине задатака, то заправо значи да се већ на почетку конструкције теста може размишљати и о броју поена који задатак носи. Самим тим, у току израде табеле спецификације може се утврдити и прелиминарна расподела поена на нивоу задатака.

Пожељно је да табеле спецификације садрже податке о форми задатка (Fives & DiDonato-Barnes, 2013). У процесу доношења одлуке да ли да се тестом обухвате превасходно задаци бирања или задаци уписивања одговора, потребно је имати у виду да обе врсте задатака имају своје предности и недостатке, а ком типу ће наставник дати предност, зависи од више чинилаца. Први критеријум је циљ примене, односно ниво знања који се жели испитати и део градива на који би задатак требало да се односи. На пример, уколико је циљ наставника је да испита способност селекције код ученика препоручује се избор задатака са бирањем одговора. Уколико се испитује способност конструкције одговора или решења, препоручује се избор задатака са уписивањем одговора. Такође, важан чинилац је и искуство наставника у изради задатака одређеног типа. Иако није уобичајено, одређени аутори критеријум искуства укрштају са критеријумом времена које наставник има на располагању за оцењивање теста (Reiner *et al.*, 2002). Укрштањем датих критеријума, изводе се следеће препоруке: а) уколико наставник нема довољно искуства у састављању задатака бирања одговора, али има довољно времена за оцењивање тестова, предлаже се избор задатака са уписивањем одговора; б) уколико наставник има довољно искуства у састављању задатака бирања одговора, али нема довољно времена за оцењивање тестова, препоручује се избор задатака бирања одговора (Reiner *et al.*, 2002). Како би се умањила вероватноћа да ученик успешно одговори на задатак када не поседује потребно знање, препоручује се да примери на које се поставка задатка односи буду другачији од примера који постоје у уџбенику (Frey *et al.*, 2005).

Независно од типа задатка који се саставља, потребно је водити рачуна да формулација захтева буде јасна и једнозначна. Како бисмо били сигурни да

смо успели у тој намери, можемо покушати да захтев тумачимо из позиције неког од слабијих ученика. Тек када смо сигурни да би и слабији ученик могао да разуме захтев, задатак се означава јасним/недосмисленим. У ситуацији избора између сложене и једноставније формулације захтева, предност се увек може дати једноставнијој формулацији (Frey *et al.*, 2005).

У домену техничких захтева, препоручљиво је нумерисати сваки задатак и водити рачуна да задатак у целини припада једној страници у тесту (да се не наставља на следећој страници). Тест би требало откуцати на компјутеру и умножити тако да све странице буду јасне, а текст задатка читљив без напора (Frey *et al.*, 2005).

Ма колико да наведени захтеви звуче логично, њихова реализација у пракси није једноставна. Притиснут бројним административним обавезама, наставник обично нема довољно времена за састављање теста према свим методолошким захтевима. У датим околностима, наставнику се лако може догодити да прескочи полазне фазе конструкције (одређење циља и садржаја теста) и да одмах приступи састављању задатака. Јасно је да је такав поступак ризичан, и са становишта добре операционализације садржаја теста и са становишта циља који се тестирањем жели постићи. Наставник постаје свестан проблема тек када се ученици пожале да је тест био превише тежак/превише лак, или су одређени садржаји у тесту више наглашени од оних којима је у настави посвећена пажња. Уколико је циљ теста било оцењивање ученика, у датим околностима може се очекивати да оцене не буду реална процена знања које је ученик стекао на настави.

Детаљно разрађена табела спецификације омогућава да се утврди и дужина теста, изражена укупним бројем задатака које тест обухвата. У нашим школама, међутим, дужина теста најчешће се усаглашава са трајањем школског часа. Унутар датог временског оквира процењује се колико би задатака тест требало да садржи с обзиром на: циљ тестирања, обим градива који се тестом испитује, времена које је посвећено усвајању тог дела градива, тежине задатака и нивоа знања који испитују, форме примењених задатака у тесту и узраста ученика (Бјекић и Папић, 2006).

Поред одређења дужине, важно је утврдити и редослед задатака у тесту. Уређење задатака одговарајуће форме у субтестове омогућава да процес прегледања и оцењивања одговора буде ефикаснији. Уређење задатака према редоследу градива, наставнику омогућава да лакше установи којим деловима градива ученик слабије влада, односно које нивое знања поседује у вези са конкретним наставним садржајем.

Када је реч о вредновању ученичких одговора, уколико је израђен детаљан кључ за вредновање одговора, вредновање одговора код задатака бирања одговора за већину наставника не представља тешкоћу. Вредновање одговора код задатака уписивања одговора представља већи изазов. Као што је претходно указано, у вредновању одговора код задатака уписивања одговора, пожељно је конструисати рубрике. У процесу конструисања рубрика у школској пракси, потребно је имати у виду следеће препоруке: број критеријума вредновања не би требало да буде велики – много је важније да се издвојени критеријуми односе на кључне вештине; пожељно је да сваки критеријум вредновања буде дефинисан што сажетије, али и што јасније; критеријуми вредновања би требало да се односе на вештине које ученици могу стећи током наставног процеса (којима их наставник подучава) (Porham, 2003). Наставник који самостално конструише рубрике може бити у дилеми да ли је направио добар избор критеријума и да ли је успешно препознао вештине које су кључне за одређени узраст ученика. У том процесу, наставнику може помоћи ослањање на исходе и стандарде изучавања одговарајућег наставног предмета за одређени узраст ученика, консултовање са искуснијим колегама или стручним сарадницима школе.

Израчунавање скорa на тесту обично је последњи корак у поступку вредновања ученичких одговора. Уколико је циљ теста оцењивање ученика, добијени скор неопходно је превести у одговарајућу оцену. На пример, наставник може одредити да се ученик према постигнућу „истиче“ уколико је на тесту решио 90% или више задатака (и бити оцењен највишом оценом); односно да „не задовољава“ уколико је решио мање од 50% задатака (и бити оцењен негативном оценом). Границе категорија одређене су тежином и

садржајем задатака, а одлуку о томе доноси наставник (Мужић, 1968). Последично, „благих“ наставници доњу границу довољне оцене могу поставити доста ниско (на пример, испод 30% тачних одговора), док „строги“ наставници границу постављају значајно више (на пример, преко 60% тачних одговора). Исти принцип односио би се и на остале оцене на скали школских оцена. Поред разлика у висини прага за остваривање одговарајуће оцене, међу наставницима могу постојати разлике и у распону поена који се односе на одговарајућу оцену. Распони такође могу варирати од оцене до оцене (Мужић, 1968).

Повремено, исправљање задатака у тесту може се поверити ученицима. На тај начин, ученици уочавају који су то делови градива којима слабије владају. Проналажење властитих грешака и дискусија о њима, помоћи ће ученику да трајније запамти исправно решење. Одабир оваквог решења за прегледање тестова знања, наставницима значи уштеду времена, али је важан и морални учинак – наставник таквом одлуком исказује поверење у ученика и правичност његовог поступка (Мужић, 1968).

У савремено доба, рачунари се користе и за вредновање ученичких одговора. То, ипак, не значи да рачунар може да „чита“ одговоре ученика. Поступак вредновања се састоји у идентификовању одговарајућих структуралних и семантичких облика за које је предвиђено да сачињавају део тачног одговора. Последњих година, развијени су различити програми за оцењивање писаног текста и графичких записа. Најпознатији су *e-rater*, *c-rater* и *m-rater*. *E-rater* располаже узорцима готових есеја на задату тему. Узорци су различитог квалитета и унапред им је одређен одговарајући број поена. Програм *e-rater* омогућава да се у одговорима ученика идентификују елементи који су слични елементима у узорцима есеја и да се препоручи оцењивање одговарајућом оценом. *C-rater* омогућава претраживање текста и идентификовање тврдњи одређеног значења и њихово аутоматско вредновање одређеним бројем бодова. *M-rater* се примењује у оцењивању алегбарских записа као што су формуле, графици и геометријске фигуре (Livingston, 2009).

Када израђује тест за примену у редовном наставном раду, наставник обично не организује поступак провере квалитета теста на начин како се то чини у припреми научног испитивања. Ипак, када год је то могуће, пожељно је да се пре примене теста реализује барем основна провера како би се отклонили озбиљнији недостаци теста.

Када је реч о провери квалитета теста кроз консултације са експертима (ревизија), она може бити садржинска и методолошка. Садржинску ревизију остварују лица која су експерти у домену садржаја за који се тест израђује. У школској пракси то се остварује тако што наставник одређеног предмета у домену провере квалитета теста консултује колеге које предају исти предмет (Rudner & Schafer, 2002; Павловић-Бабић и сар, 2003). У домену методолошке ревизије теста наставнику могу да помогну стручни сарадници школе.

Тест се може проверити и применом у другој школи, уз помоћ наставника који предају исти предмет, по истом наставном програму и користећи исти уџбеник (Богавац и сар, 1972). На нивоу школе, наставници могу реализовати претходну проверу теста на малом узорку ученика којима предају (Богавац и сар, 1972).

Аспект провере квалитета теста јесте и анализа квалитета задатака која се спроводи након реализације тестирања. Реч је о анализи која подразумева утврђивање индекса тежине и индекса дискриминативности сваког појединачног задатка. Оба индекса се израчунавају применом одговарајућих математичких формула (Taylor, Greer & Mussio, 1978).

Пре задавања теста, наставници могу само приближно да процене тежину задатака. У процени, наставник се ослања на садржај задатака и дотадашње постигнуће ученика у одељењу (Џордан, 1966). Задатак наставника је да препозна који задаци могу имати индекс тежине између 0.3 и 0.7. Управо такви задаци би требало да доминирају у финалној верзији теста. Индекс тежине задатка, дакле, нема апсолутну вредност за ученике истог узраста, већ се његова вредност утврђује за одговарајућу групу ученика у одговарајућем тренутку (Taylor, Greer & Mussio, 1978).

Индекс дискриминативности даје информацију о томе да ли су ученици који су били успешнији на тесту генерално, чешће одговарали исправно на дати задатак од ученика који су били мање успешни у решавању теста у целини. Индекс дискриминативности може имати вредност од -1.0 до +1.0. Што је вредност индекса већа, то је задатак дискриминативнији (Taylor, Greer & Mussio, 1978). Висока вредност индекса дискриминативности једног задатка сугерише да задатак адекватно мери својство за чије испитивање је тест конструисан. Ниска вредност датог индекса још увек није поуздан показатељ да задатак није добар, док негативна вредност индекса дискриминативности указује да извршен проблем са задатком постоји, јер га успешни ученици лошије решавају од мање успешних ученика. На пример, задатак може имати више тачних одговора или је предвиђени начин бодовања неодговарајући. Уколико се констатује да поједини задаци нису били адекватно састављени, ревизија задатака може да допринесе њиховом побољшању и укључивању ревидиране форме у неки од тестова знања који ће се користити са другом групом ученика или следеће школске године (Taylor, Greer & Mussio, 1978).

Једном примењени задаци могу се поново употребити у некој наредној ситуацији тестирања. У тежњи да сваки тест буде другачији од претходног, наставници могу правити различите комбинације задатака. Да би такав начин рада био могућ, неопходно је да наставници формирају базе задатака, у папирној или електронској форми. Базе се могу употпуњавати, задаци који већ постоје могу се дорађивати или заменити новим задацима. База задатака се не мора нужно састојати из задатака које је наставник лично саставио. У бази се могу наћи и задаци, или тестови у целини, који су наставнику постали доступни путем стручне литературе, интернета или разменом са колегама. Коришћење „готових“ тестова наставнику олакшава рад и омогућава боље економисање временом.

У Америци, наставници се могу одредити за коришћење комерцијалних тестова знања које израђују одговарајуће компаније. Испитивање Пофама (2001) потврдило је да се већина наставника радије одређује за ту могућност него за самосталну израду тестова. Комерцијални

тестови су доступни заједно са табелом спецификације и упутствима за примену и тумачење скорова, што их чини атрактивном алтернативом самосталној изради теста. У нашој земљи још увек нема комерцијалних тестова, али се израдом тестова све више баве издавачи уџбеника. На пример, неки од издавача који израђују тестове знања за наставу српског језика су: *Клет*, *Нови Логос*, *Завод за уџбенике*, *Креативни центар*, *Едука*. Издавачи који израђују тестове за наставу математике су: *Клет*, *Креативни центар*, *Завод за уџбенике*, *Круг*, *Математископ*, *Архимедес* и тако даље.

Формирање базе задатака олакшано је применом рачунара у свакодневном животу. Поред уређења база задатака на персоналном рачунару, наставници могу приступити различитим базама задатака путем интернета. Некада су у питању базе комерцијалне природе. Врло често је реч о базама које формирају наставници за своје личне потребе и које су вољни да поделе и унапређују кроз сарадњу са другим корисницима.

Упркос малобројности, до сада реализована испитивања у вези са тестовима које наставници самостално израђују указују на нека важна питања и проблеме које би било значајно даље истраживати. Када је реч о иностраним испитивањима, истраживање америчких аутора (Stiggins & Bridgeford, 1985) показало је да у процесу израде и примене тестова наставници ретко преиспитују своје компетенције и ретко сарађују са другим колегама, мада, генерално, изражавају спремност да се посвете унапређивању квалитета тестова које праве. Као основне препреке на том путу наводе недостатак времена, средстава и понекад мотивације да ревидирају своју праксу. Исто испитивање (Stiggins & Bridgeford, 1985) је показало да наставници математике и природних наука чешће израђују тестове знања у односу на наставнике језика. Наставници језика се више ослањају на друге форме утврђивања постигнућа ученика. Комерцијалне тестове знања користи око половине испитаних наставника. Примена комерцијалних тестова је учесталија у раду са ученицима нижих разреда и нешто више у математици него у другим предметима. У раду са старијим ученицима, наставници више поверења поклањају тестовима које самостално припремају.

Када је реч о истраживањима на нашим просторима, у недостатку новијих испитивања, осврнућемо се на резултате истраживања Богавца и сарадника (1972). У наведеном испитивању наставници су оценили да су у изради тестова најефикаснији када им помажу стручни сарадници (42,7% одговора), мада је значајан број наставника самосталан у изради тестова (30,3%). Укупно 21% наставника изјавио је да се опредељује за примену већ готових тестова знања, а 6% наставника сматра да је најбоље решење да се израда и реализација тестирања повери стручним лицима. Уколико процене да им је у изради тестова потребна помоћ, наставници се најчешће опредељују за консултовање приручне литературе (58,8% одговора) или за похађање стручних предавања и семинара (16,8%). Занимљиво је да је 16,8% наставника оценило да им није потребна никаква помоћ у домену израде тестова знања.

Ипак, друга испитивања показују да тестови које наставници конструишу имају недостатака (Ebel, 1979). Објективнија слика о карактеристикама тестова може се добити на основу извештавања ученика који тестове решавају или на основу извештавања стручних лица којима се наставници обраћају за помоћ у вези са конструкцијом теста. Ебел (1979) запажа да су наставници склони да тестове састављају „на брзину“, а приментна је и тенденција састављања релативно кратких тестова којима се не може обухватити сво градиво. Известан број наставника не води рачуна о различитим нивоима знања који се тестом могу мерити, као ни о формулацији задатака. Додатно, мали број наставника анализира постигнуће које ученици остварују на тесту. Самим тим, неповољна дистрибуција скорова на тесту пре се тумачи као резултат стварног стања ствари у вези са знањем ученика, него као последица квалитета самог теста (Ebel, 1979).

Резултати испитивања показују да је учесталост примене тестова различита у настави различитих школских предмета. Такође, самосталност наставника у изради тестова варира од потпуне до одлуке да се израда теста повери неком другом. Испитивања показују да постоји простор да се квалитет тестова унапреди, као и да се шанса може тражити у домену стручног усавршавања.

Особености тестова знања који се користе у настави за праћење и вредновање рада ученика

Поред тестова знања, уобичајени начини проверавања ученичких знања у нашим школама су усмено испитивање и израда писаних вежби. За усмено испитивање је карактеристично то да трајање испитивања, обухватаност, сложеност и број питања који се могу поставити, нису унапред дефинисани, већ их наставник прилагођава конкретном ученику. Писане вежбе подразумевају задавање мањег броја питања или задатака свим ученицима једног одељења. Ученици раде у истоветним просторно-временским условима, а одговоре дају у писаној форми: најчешће састављајући есеј, описујући решење проблемске ситуације, приказујући поступак решавања рачунског проблема или конструишући решење задатка.

У раду са ученицима који нису савладали вештину писања, усмени испит је једини начин испитивања који се може спровести, независно од предметне области. Када имамо у виду старије ученике, у настави одређених школских предмета усмено испитивање више је заступљено од осталих начина испитивања ученика. На пример, у настави језика усмено изражавање је уобичајени начин да се провери акцентовање речи, обликовање реченица или исправан говор на матерњем или страном језику (Трој, 1957). Код усменог одговарања ученик не може предуго да се задржи на неком питању које му је тешко, јер наставник одговарајућим потпитањима може да преусмери излагање ученика. Усмено испитивање, ако је вешто вођено, може да помогне наставнику да препозна тешкоће које ученик има у разумевању одговарајућег садржаја или у разумевању одговарајућег садржаја (Рајстон, Цастман и Робинс, 1966). Такође, од свих начина испитивања ученика, усмено испитивање је најмање формалан начин (*Guide to Assessment*, 2015).

Ипак, усмено испитивање сматра се најмање поузданим начином проверавања ученичких знања. Аргументи у прилог таквом становишту изводе се из природе самог поступка провере који карактеришу: неједнаки

захтеви за различите ученике; ограничен обухват наставног садржаја; дуготрајност која води смањењу објективности испитивања услед замора код ученика и наставника; могућност да наставник својим интервенцијама утиче на квалитет одговора ученика; непостојање објективног трага о учениковом одговору (ученик увек може тврдити да је знао за бољу оцену) (Фурлан, 1964; *Guide to Assessment*, 2015).

Када је реч о писаним вежбама (у које сврставамо и писмене вежбе) њихове основне предности су: истоветност захтева за све ученике; лакоћа састављања задатака и реализације поступка испитивања; пружање могућности испитанику да образложи свој одговор; примењивост у већини школских предмета; поступак испитивања није ограничен на кратко време и није обликован иницијативом наставника или реакцијама других ученика (Богавац и сар, 1972). Недостаци се односе на мањи обухват наставног градива, дуг временски период потребан за прегледање и деловање субјективног фактора приликом оцењивања (Богавац и сар, 1972).

Деловање субјективних чинилаца карактеристично је за усмено испитивање, али у одређеној мери и за писане вежбе. Субјективности испитивања доприноси понашање и карактеристике самог оцењивача као што су (Фурлан, 1964):

- индивидуално схватање важности појединих делова градива
- лична једначина наставника у оцењивању (претерана строгост или благост у оцењивању)
- тренутно расположење наставника
- подложност деловању закона контраста (деловање утиска о квалитету рада претходног ученика на процену квалитета рада следећег ученика).

Међутим, мањој објективности испитивања доприносе и чиниоци који потичу од испитаника (Фурлан, 1964), а који утичу на наставникову процену:

- говорна вештина или вештина писаног изражавања испитаника
- претходно постигнуће ученика у одређеној области знања
- психо-физичко стање испитаника у тренутку испитивања
- изглед и уредност испитаника.

У литератури се усмено и писано испитивање сврставају у облике субјективног проверавања ученичких знања, за разлику од тестова знања који се обично означавају као вид објективног проверавања ученичких знања (Трој, 1957; Спахиу, 1991; Хавелка, 2000). Најзначајније предности тестова знања огледају се у (Хавелка, 2000; Бјекић и Голубовић, 2003; *Guide to Assessment*, 2015):

- широком обухвату наставног градива
- истоветности захтева који се ученицима постављају
- објективности поступка оцењивања
- економичности и
- флексибилности (могућност поделе теста на субтестове и њихово независно коришћење).

Обухват наставног градива, објективност поступка оцењивања и флексибилност (у наведеном значењу) могу се сматрати предношћу тестова знања у односу, како на усмено испитивање, тако и у односу на испитивање ученика путем писаних вежби (Ebel, 1979). Истоветност захтева и економичност могу се сматрати предношћу тестова знања у односу на усмено испитивање ученика.

Ипак, примена теста није једноставна. Најпре, реализација испитивања тестом, као и писаном вежбом, захтева одговарајућу просторно-временску организације испитивања. Такође, решавање теста знања и израду писаних вежби увек прати могућност „варања“ од стране ученика, која је мање изражена у околностима усменог испитивања. Истовремено, тестови знања представљају поступак провере знања чија припрема је сложенија од припреме усменог испитивања и припреме писаних вежби, посебно ако је циљ састављање задатака који испитују различите нивое знања, односно састављање задатака који су различити према форми (Ebel, 1979).

У односу на електронске тестове знања, ниједна од разматраних карактеристика не би се могла сматрати предношћу класичних тестова знања. Можда би се, чак, могло говорити о томе да класичан тест у односу на електронски има одређена ограничења: електронски тестови се брже решавају, аутоматски се скорују, обезбеђују испитанику повратну информацију

непосредно по решавању теста, и не захтевају умножавање и физичку дистрибуцију. Међутим, када је реч о „варању“ на тесту, у условима недовољне стандардизације тестирања, електронски тестови могли би се суочити са истим потешкоћама као и класични тестови. Уколико наставници нису адекватно обучени за примену електронских тестова, класични тестови знања имају предност над електронским у погледу могућности адекватне примене.

Може се закључити да тестови знања и њихове рачунарске верзије имају више предности него недостатака, када се упореде са усменим испитивањем и писаним вежбама. Иако имају своје предности, неоспорно је да тестови не смеју бити једина форма испитивања ученичког постигнућа. Тестови знања су ипак само један од начина испитивања ученичких знања, који је у редовном наставном раду пожељно комбиновати са другим начинима испитивања (Kyriacou, 2001). Аналогно томе, не би било добро да се било која форма испитивања учини искључивом или важнијом од других. Уколико се постигнуће ученика процењује на више различитих начина, односно ученик има могућност да пружи и усмени и писани одговор, и да покаже овладаност одговарајућим вештинама, тада је вредновање резултата праведније. Другим речима, умањује се вероватноћа да ће процена постигнућа бити ослоњена само на облик испитивања у коме ученик генерално показује слабије резултате (Gipps, 2003).

Упркос предностима, критика тестова знања развијала се у више праваца. Историјски најранију примену тестова знања пратила је критика ограничавања васпитног деловања наставника на ученика. Наиме, сматрало се да је решавање теста ситуација у којој наставник не може помоћи ученику у току одговарања, не може га охрабрити, мотивисати, поставити му потпитање, преусмерити излагање (Гргин, 2001; Бјекић и Голубовић, 2003). Истицало се да тест не пружа могућност развијања критичког мишљења, интелектуалне радозналости и оригиналности. Наведене критике праћене су контрааргументима да истакнути недостаци тестова знања одликују искључиво неадекватну примену тестова знања у којој се превиђа могућност

васпитног деловања наставника *након* примене теста, пружањем индивидуализованих и конструктивних повратних информација (Гргин, 2001).

Критике које су уследиле касније, а које се могу сматрати и даље актуелним, односе се на немогућност прилагођавања захтева у тесту индивидуалним карактеристикама ученика (Ebel, 1979; Гргин, 2001). Ученици се разликују не само према квалитету знања које поседују, већ и у погледу могућности да своје знање саопште писаним путем. У критици тестова се наглашава да се у тестовима знања траже кратки и прецизни одговори који не пружају праву слику о знању ученика – ученик може дати исправан одговор „погађањем“ или, на супрот томе, бити непрецизан у изражавању знања које је заиста стекао. Максимална оцена на тесту знања не значи да је ученик у потпуности савладао све значајне теме из испитиване области (Бјекић и Голубовић, 2003). Такође, не може се тврдити ни да ученик који добије слабу оцену заиста не поседује никакво знање из испитиване области.

Израда квалитетних тестова је сложен и дуготрајан процес, посебно у околностима када се желе испитати различити нивои знања (Баковљев, 1997). Неискусни и недовољно вешти састављачи теста, изложени су опасности да конструишу тест који мери искључиво ниво познавања чињеница, и онда када је намера била да се испитају и виши нивои усвојености знања (Мужић, 1986). Међутим, испитивања показују да чак и искуснији испитивачи подлежу истој врсти грешке (Gipps, 2003). Таква пракса довела је до тога да су тестови које срећемо у школској пракси често критиковани због испитивања превасходно елементарних и фактографских знања (Гргин, 2001; Баковљев, 1997; Хавелка, 2000; Бјекић и Голубовић, 2003).

Недостаци тестова знања сагледавају се и у немогућности да се тестом испитају одређене активности као што су спровођење експеримената, решавање проблемских ситуација (Ediger, 1999; према: Ediger, 2000), креативно мишљење (*Guide to Assessment*, 2015), али и труд, радозналост, иницијатива, упорност (Кузмановић и Павловић-Бабић, 2011). Примена тестова у датим околностима није погодна, јер се тестом утврђује стање неке појаве и резултати

образовно-васпитног рада (Банђур и Поткоњак, 1999), али не и процес којим се стиже до тих резултата (Anastasi, 1984; Баковљев, 1997).

Последњих деценија говори се о два доминантна приступа процењивању образовних постигнућа: психометријском и аутентичном процењивању (Кузмановић и Павловић-Бабић, 2011). Наведни приступи се сматрају међусобно супротстављеним, с обзиром на то да другачије приступају процесу школског учења. Психометријско процењивање је утемељено на бихевиоризму, не наглашавајући важност контекста у коме се учење дешава. Оријентисано је на примену нормативних тестова знања којима се процењује постигнуће ученика у савладавању образовног садржаја истоветног за све ученике одређеног узраста. Аутентично процењивање је утемељено на когнитивистичком, конструктивистичком и ситуационом приступи учењу. Залаже се за проучавање особе у њеном природном окружењу, и оријентисано је на праћење напредовања ученика у достизању индивидуализованих циљева.

Најозбиљнија критика тестова знања дошла је из редова присталица аутентичног процењивања. Поступци процењивања постигнућа који се у оквиру овог приступа означавају као добра алтернатива тестовима знања су процењивање делатности (*performance assessment*) и вредновање постигнућа на основу ученичких досијеа – портфолија. Процењивање делатности у најширем значењу подразумева сваку форму утврђивања постигнућа која се не ослања на примену задатака са вишеструким понуђеним одговорима (Algarabel & Dasi, 2001; Gipps, 2003). У ужем значењу, реч је о испитивању у коме на постављени задатак ученик одговара конструишући свој одговор или изводећи одговарајућу радњу (не правећи избор међу већ понуђеним одговорима). Ипак, наглашавање форме испитивања није оно што у потпуности одређује значење процењивања делатности. Реч је о начину испитивања који је увек усмерен ка сложеним знањима и вештинама, као што су вештине вербалне комуникације, вештина писања есеја, активности решавања проблема (Popham, 2003).

Полазно интересовање за процењивање делатности развило се у оквиру разматрања проблематике критеријумских тестова знања (Algarabel & Dasi, 2001; Gipps, 2003). Критеријумски тестови вреднују постигнуће ученика у

односу на дефинисани низ циљева и стандарда, а такви циљеви и стандарди се неретко односе на видљиво понашање, односно извођење одговарајућих радњи (Hariss, Alkin & Popham, 1974; према: Frey & Schmitt, 2007). Предности процењивања делатности су: фокусирање на процес, не само на резултат одређене активности; испитивање постигнућа у динамичном окружењу; виша мотивисаност испитаника да учествују у тако коципираном испитивању него у класичним моделима испитивања (усмено, писане вежбе и тестови знања) (*Guide to Assessment*, 2015). Основно ограничење у примени процењивања делатности је неекономичност, с обзиром на дуготрајност испитивања појединачног испитаника. Други озбиљан недостатак је поузданост процена које се на описани начин постижу (Popham, 2003; *Guide to Assessment*, 2015).

Ученички досијеи или портфолија подразумевају колекције продуката учениковог рада којима се потврђује да су остварени очекивани захтеви у настави. Нагласак је на сакупљању резултата рада и њиховој евалуацији (Popham, 2003). Ученик има задат циљ пред собом, а на путу до циља вреднује резултате свог рада (самостално или уз помоћ других ученика) како би утврдио да ли напредује у стицању знања или овладава вештинама које су постављене као циљ учења. Предности оцењивања постигнућа на основу досијеа огледају се у могућности да се напредовање ученика прати континуирано, из дана у дан, као и да добијене информације о напретку буду контекстуално утемељене (Ediger, 2000; Ali Salmani Nodoushan, 2010). Порфолија омогућавају да се створи интегрисана слика о постигнућу ученика у различитим аспектима рада у настави, која је притом и високо персонализована. Недостатак портфолија је њихова слаба упоредивост. Тешко је применити јединствен критеријум у процени ученичког рада на основу досијеа који могу бити сачињени од врло различитих елемената. Портфолија се конституишу током дужег временског периода, што се такође сматра њиховим ограничењем (Popham, 2003; *Guide to Assessment*, 2015).

Нови начини испитивања инспирисани су теоријом Виготског и учењем о зони наредног развитака. Тестови знања постају недовољни у мери у којој испитују само оно што ученик зна – наставнику постаје важно којим још

знањима и вештинама ученик може да овлада уз одговарајућу подршку. У духу критичке теорије, студија случаја замењује узорковање, а квалитет извођења аутентичних радњи замењује поновљивост и генерализацију (Gipps, 2003). Процењивање делатности и примена портфолија, испитивање постигнућа чине мање економичним, мање поузданим и објективним, али су више оријентисани на конкретног ученика.

Примена стандардизованих тестова знања у испитивању образовних постигнућа ученика

У нашем образовном систему, примена стандардизованих тестова карактеристична је за завршни испит на крају основног образовања (Мала матура), за реализацију националних и међународних тестирања. Завршни испит је обавезан за све ученике осмог разреда и спроводи се на крају школске године. Реализација националних и међународних тестирања спроводи се у одређеним временским циклусима, обухвата ученике одређеног узраста (четвртог и/или осмог разреда основне школе, или првог разреда средње школе) и поједине школе образовног система.

Полагање завршних испита присутно је у образовним системима широм света. Тренутно, у руском образовном систему завршни испити се полагају на крају средње школе (Tumeneva, 2013). У Великој Британији обавезна тестирања почињу већ у нижим разредима основне школе и реализују се у више наврата током основног и средњег образовања. На основу резултата које ученици остварују на тестовима знања формира се ранг-листа престижних школа, како основних тако и средњих (Rotberg, 2006). Слично као у Великој Британији, у Немачкој обавезна тестирања почињу рано, око десете године ученика. Циљ са којим се тестирања спровode јесте праћење напредовања ученика и помоћ у избору средње школе (Rotberg, 2006). У Сингапуру, циљ тестирања је усмеравање ученика у процесу даљег школовања, са нагласком на селекцију ученика који могу да похађају престижне високе школе. На основу резултата тестирања не вреднује се рад наставника, али се формира ранг листа

школа (Rotberg, 2006). У Кини су тестирања учестала од најранијих узраста, а наставници пуно времена посвећују припреми ученика за тестирање (постоје чак и приватне школе које се отварају са том наменом) (Ravitch, 2014).

Завршни испит у форми теста знања у наш образовни систем уведен је школске 2010/2011. године. Базиран је на усвојеним стандардима за крај обавезног образовања и састоји се од три теста знања: теста из српског/матерњег језика, математике и комбинованог теста (који обухвата знања из биологије, географије, историје, физике и хемије). Њиме се процењује степен савладаности прописаних васпитно-образовних циљева и задатака, као и општих стандарда знања на крају основне школе. Тестови које ученици решавају на завршном испиту састоје се од задатака којима се испитују компетенције на сва три нивоа образовних стандарда. Садрже од 20 до 40 задатака у зависности од процењене тежине и типа коришћених задатака. Ученик може да оствари највише 10 бодова на тесту из српског/матерњег језика, 10 бодова на тесту из математике и 10 бодова на комбинованом тесту. Преостале бодове сачињава остварени школски успех током обавезног образовања. Максималан број бодова који ученик може остварити на завршном испиту је 100 бодова. Од 2014. године завршни испит је састављен искључиво од ученицима непознатих задатака (*Правилник о програму завршног испита у основном образовању и васпитању*, 2014: члан 3).

Завршни испит омогућава да се утврде образовна постигнућа ученика као појединаца. У реализацији завршног испита тежи се високом степену стандардизације услова тестирања и критеријума вредновања. Примарна тежња је да сви ученици имају исти третман у току испитивања и у примени критеријума вредновања. На основу положеног испита, ученик добија сведочанство о етапи и врсти школовања које је завршио и процену успешности у процесу школовања (Ковач-Церовић и Левков, 2002).

Национална тестирања почињу спорадично да се примењују у другој половини XX века, да би њихова заступљеност у Европи нагло порасла крајем прошлог века. Национална тестирања могу се поделити у три категорије (*National Testing of Pupils in Europe: Objectives, Organizations and Use of Results*, 2009):

- тестирања постигнућа ученика на крају одређеног образовног стадијума, а која имају значајан утицај на даље образовање ученика
- тестирања која се спровode са циљем мониторинга или вредновања рада школа и/или школског система у целини
- тестирања чији је основни циљ унапређивање процеса учења код ученика, идентификовањем њихових специфичних потреба и прилагођавањем наставног рада потребама ученика.

Већина земаља у Европи усвојила је модел реализације националних тестирања два до три пута током обавезног образовања. Постоје државе у којима је тестирање учестало, као на пример у Великој Британији и Француској где се спроводи седам, односно шест националних тестирања током обавезног школовања. Наспротив таквој пракси, у државама као што су Немачка, Шпанија, Холандија и Словачка национална тестирања се организују само на завршној години основног или нижег средњег образовања (*National Testing of Pupils in Europe: Objectives, Organizations and Use of Results*, 2009).

У случају Велике Британије (*Testing and Assessment*, 2008) и Ирске (*Supporting Assessment in Schools – 3*, 2005), учестала тестирања се доводе у везу са вишеструким циљевима којима се тежи реализацијом националних тестирања. У зависности од тога који се циљ одреди као приоритетан, одређује се динамика тестирања, начин на који ће се обликовати извештај о испитивању и утврђује ко све може имати приступ подацима добијеним испитивањем (*Supporting Assessment in Schools – 3*, 2005).

Компетенције које се углавном испитују националним тестовима широм Европе су комуникација на матерњем језику, страним језицима, математичке компетенције и базична знања у области науке и технологије. Испитивања тестом знања у оквиру националних тестирања, најчешће су праћена анкетним испитивањем директора школа, наставника, родитеља ученика и самих ученика. Упитници омогућавају да се спроведе анализа контекста у коме се постижу одређени резултати на тесту, и да се идентификују чиниоци који могу допринети одговарајућем постигнућу ученика (Бауцал и сар, 2006; *National Testing of Pupils in Europe: Objectives, Organizations and Use of Results*, 2009).

Две трећине европских земаља резултате националних тестирања прикупља на нивоу појединачних школа и државе у целини (*National Testing of Pupils in Europe: Objectives, Organizations and Use of Results*, 2009). Међутим, већина земаља резултате појединачних школа не објављује јавно. Обично се само национални резултати објављују у форми посебних извештаја уз анализу контекстуалних варијабли, популационих карактеристика и школске инфраструктуре. За разлику од праксе у Америци и Канади, европске земље ретко примењују праксу награђивања односно кажњавања школа сходно резултатима које њихови ученици остваре на националним тестирањима.

Прво национално тестирање образовних постигнућа у Србији спроведено је школске 2003-2004. године. Тестирањем је обухваћено градиво математике и српског језика и односило се на ученике трећег разреда (Бауцал и сар, 2006). Циљеви националног тестирања били су да се: обезбеде подаци о томе којим знањима и вештинама из српског језика и математике владају ученици након три године школовања; омогући дефинисање стандарда знања из српског језика и математике за одговарајући узраст; утврде фактори који детерминишу образовна постигнућа ученика, и започне праћење квалитета разредне наставе у Србији периодичним понављањем националних тестирања. Национално испитивање ученика четвртог разреда реализовано је 2006. године. Циљеви националног тестирања били су да се: утврде знања, умења и вештине којима ученици владају из српског језика и математике на крају првог циклуса образовања; материјални и људски фактори који детерминишу образовна постигнућа ученика; сагледају услови у којима раде школе и процени стручност образовног кадра (Чапрић и сар, 2007). Након 2006. године национална тестирања у нашој земљи нису реализована.

Када је реч о међународним тестирањима, наша земља је до 2015. године учествовала у пројектима PISA - *Programme for International Student Assessment* (2003, 2006, 2009, 2012) и TIMSS - *Trends in International Mathematics and Science Study* (2003, 2007, 2011).

PISA студија испитује у којој мери су ученици успели да развију компетенције (читалачка, математичка и научна писменост) које су им

неопходне да би могли успешно да реше различите животне ситуације. За разлику од таквог полазишта, где се наставни програм посматра као инструмент у циљу остваривања постављених исхода, TIMSS студија се фокусира на наставне програме. Полази се од становишта да је за образовну политику најважије да се добију одговори на питања шта је садржај образовних програма (предвиђени курикулум), који део програма је реализован у пракси (примењени курикулум), а који део реализованог програма је и фактички усвојен од стране ученика (постигнути курикулум) (Милановић-Наход и Јањетовић, 2004). Предвиђени курикулум обухвата национални, друштвени и образовни контекст у коме се реализује наставни процес; примењени курикулум обухвата контекст одељења, школе и наставника, а постигнути курикулум обухвата ученичка постигнућа и обележја (Гашић-Павишић, Станковић и Малинић, 2011). У нашој земљи, до сада реализована PISA истраживања обухватала су петнаестогодишње ученике, док је реализација TIMSS 2003 и 2007 испитивања обухватила ученике осмог разреда, а испитивање TIMSS 2011 ученике четвртог разреда основне школе.

И PISA и TIMSS студија испитују одређене аспекте контекста у коме се постигнуће остварује. У ту сврху, спроводе се анкетна испитивања директора школа, наставника и самих ученика. Подаци анкетних испитивања користе се у интерпретацији података добијених тестирањем. Извештаји који прате реализацију истраживања су садржајни и комплексни. Међународни тим истраживача израђује сумарни извештај о резултатима истраживања, док све земље учеснице израђују националне извештаје.

PISA и TIMSS су компаративне студије које омогућавају земљама учесницама да упореде своје образовне системе и да сагледају специфичности образовног контекста у ком се школују најуспешнији ученици. Изведени закључци могу водити ка променама у домену образовне политике. На пример, резултати PISA и TIMSS испитивањима указали су на то да најуспешније земље имају развијене системе за праћење образовних постигнућа ученика засноване на јасном дефинисању стандарда знања за поједине разреде и предмете, међупредметних стандарда, као и стандарда

квалитета образовних институција (Гашић-Павишић, 2011). Такав податак пружа подстицај и другим земљама учесницама да дефинишу, размотре или унапреде своје системе за праћење образовних постигнућа ученика.

Припрема и реализација тестирања

Начин реализације тестирања је важно методолошко питање, посебно када се имају у виду два паралелна стремљења: а) да се обезбеде стандардизовани услови за сповођење тестирања и б) да се тестирањем добију подаци који су аутентични за сваког испитаника. У том смислу, стандардизација услова тестирања у вези је са питањима валидности теста, могућностима компарације и обезбеђивања праведности испитивања (Hattie, Jaeger & Bond, 1999).

Одређени аспекти стандардизације услова тестирања остварују се кроз припрему за тестирање. Суштина припреме за тестирање јесте упознавање реализатора тестирања и испитаника са процедуром тестирања и са садржајем теста у одређеној мери. Упознатост са различитим аспектима ситуације тестирања омогућава да се умањи ефекат „страха од непознатог“, предвиде одређене околности и испланирају неопходни кораци у циљу превазилажења потенцијалних проблема.

Припрема за тестирање обично подразумева низ корака које испитивач предузима у ситуацији када планира реализацију тестирања. Три основна корака су: припрема испитивача, припрема тестовног материјала и припрема испитаника.

У зависности од тога о којој врсти тестова је реч, припрема испитивача је мање или више сложена и креће се од овладавања садржајем упутства за реализацију теста и поступање у складу са захтевима упутства, до стицања формалних сертификата за реализацију тестирања (у случају, на пример, тестова интелигенције) (Гојков, 2012). Припрема материјала за тестирање подразумева припрему тестова (евентуално и листова за одговоре), упутства за тестирање и протокола тестирања у који истраживач може да убележи све

неправилности које су се догодиле у току тестирања. Уколико је могуће, истраживач може обезбедити резервни прибор за писање, а према потреби и друга неопходна средства (празне листове папира, сат, слике, аудио или видео записе) (Муџић, 1986; Баковљевић, 1997; Фајгел, 2003).

Најсложенији аспект припреме за тестирање свакако је припрема испитаника. Упркос сложености, припрема испитаника је од изузетне важности, јер испитаници који се припремају за тестирање постижу боље резултате на тесту знања у односу на испитанике који се не припремају (Бјекић и Папић, 2006).

Први корак у припреми испитаника за тестирање подразумева благовремено информисање о томе да ће се тестирање обавити, шта је циљ тестирања и какви типови задатака ће бити заступљени на тесту. Када је реч о нашој школској пракси, одредбом Правилника о оцењивању (2013), наставник је дужан да обавести ученике о наставним садржајима који ће се писмено проверавати најкасније 5 дана пре провере. Поред основног информисања о планираном тестирању, наставник би требало да допринесе формирању позитивних ставова према испитивању и да мотивише ученике да се што боље припреме за тест.

Успех у решавању теста детерминисан је не само знањем које испитаник поседује, већ и овладаношћу стратегијама решавања теста. У литератури је прихваћено становиште да се стратегије могу усвојити правилно организованим вежбањем (Hambelton & Jirka, 2004). Уобичајен начин подучавања стратегијама решавања задатака подразумева да наставници презентују ученицима задатке на којима се најчешће грешти (Rudner & Schafer, 2002). Потом се ученицима остави време да својим речима препричају захтев и да самостално реше задатак. Након тога, приступа се поређењу интерпретација и решења задатка, указивању на исправна тумачења захтева и исправно решење задатка. Искуства наставника показују да ученици често лоше интерпретирају захтев у задатку, и међу понуђеним одговорима просто одаберу онај који се језички поклапа са неким делом поставке задатка.

У вези са одговарањем на есејска питања, у припреми за тестирање ученицима се може саветовати да најпре израде структуру свог одговора. На тај начин, лакше се организују мисли, мање су шансе да се изостави важан део одговора и писање иде брже. Такође, ученицима се саветује да одговарају прецизно на захтев питања – уколико је потребно да наведу пример, нема потребе да наводе дефиниције одговарајућих појмова. Уколико преостане времена, пожељно је преконтролисати граматичку и правописну исправност својих одговора (Rudner & Schafer, 2002).

У школској пракси, указивање на начин решавања конкретних задатака обично је кратко или потпуно изостаје, јер се полази од претпоставке да су ученици током припрема за тест већ упознати са основним захтевима у том погледу. Очекује се да су операције које је потребно да ученик примени у решавању задатака претходно увежбане на часу (Џордан, 1966). Уколико се не поступа на тај начин, лоше постигнуће на тесту није могуће једнозначно довести у везу са неовладаношћу градивом. Напросто, ученицима може бити нејасан поступак решавања задатка, а да сасвим солидно владају градивом.

Обезбеђивање одговарајућих просторних услова један је од основних услова успешне реализације тестирања. У просторији у којој се спроводи тестирање неопходно је сваком испитанику обезбедити место за рад, довољно удаљено од места предвиђених за друге испитанике. Просторија за рад би требало да буде светла и прозрачна, заштићена од буке, како се не би реметила концентрација ученика (Буквић, 1996). Када је реч о временском оквиру, препоручује се да се ученици испитују тестом знања у време наставе и да се строго поштује сатница тестирања (Буквић, 1996). Организовање тестирања у време које није предвиђено за наставу чини ситуацију тестирања више формалном, чиме се појачава психолошки притисак на ученике. Код одређених ученика то се може лоше одразити на постигнуће које ће остварити на тесту. Инсистирање на прецизном временском оквиру важно је како би се гарантовало да су сви ученици имали подједнако време за рад. Уколико се временски оквир не би поштовао, већ би се појединим ученицима дозвољавало

да раде дуже (да раније почну или касније заврше са радом), јасно је да не би сви имали једнаке шансе за добро постигнуће.

Разумљиво је да ученици пред само тестирање осећају нелагоду и напетост. Повремено, ниво тензије може бити толико висок да се негативно одрази на резултате тестирања. Што су резултати тестова значајнији, то је и емоционална нелагодност већа. Задатак испитивача је да почетну напетост ублажи пригодним обраћањем испитаницима, указивањем на сврху тестирања, саветом да пажљиво прочитају задатак у тесту, размисле о захтеву и тек након тога упишу одговор (Буквић, 1996). У датој ситуацији, ученике је пожељно подсетити да свака оцена може да се поправи и да је резултат теста само један од резултата њиховог рада. Као што није пожељно да наставник својим понашањем индукује страх и нервозу код ученика, није пожељно ни да ствара превише неформалну атмосферу која би подстакла ученике да користе недозвољена средства у раду (Мужић, 2004).

„Добро“ конструисан тест знања пратиће врло мало питања од стране ученика. Уколико ученици имају потребу да затраже објашњење, препоручује се да то учине пре почетка рада на тесту. Обраћање групи ученика, или целом одељењу није препоручљиво након што започне решавање теста (Taylor, Greer & Mussio, 1978; Мужич, 1986). У току решавања теста, није препоручљиво да се наставник интензивно креће кроз учионицу, као ни да омета рад ученика на било који начин (разговором, посматрањем учениковог рада „преко рамена“) (Мужић, 1986).

Пожељно је да наставник пре почетка рада инструкује ученике да ли тест предају оног тренутка када заврше са радом, или је потребно да сачекају крај часа. Ученике који заврше са радом пре истека укупног времена за рад, наставници обично саветују да преконтролишу своја решења. Међутим, ученици то нерадо чине (Rudner & Schafer, 2002). Основни разлог је тај што за већину ученика провера решења значи да поново „проживљавају стрес“ решавања теста од почетка до краја. Ипак, то не мора да буде једини начин. Препоручује се да наставници инструкују ученике да провере само задатке у вези са којима су имали дилеме. Такође, ученицима се може саветовати да

направе кратку паузу пре него што преконтролишу своја решења – да исправе леђа, освеже се водом, како би „бистре главе“ могли да наставе са радом (Rudner & Schafer, 2002: 101).

У школској пракси, уколико наставници нису самостално саставили тест, неопходно је да се руководе захтевима за реализацију тестирања које припремају конструктори теста. Поступање у складу са захтевима упутства обезбеђује бољу контролу услова у којима се тестирање спроводи. Тестове које наставници самостално израђују обично не прате структурирана упутства за реализацију тестирања, али је контрола услова у којима се тестирање спроводи и даље потребна.

Структуру упутства за ученике сачињавају следећи елементи: указивање на циљ тестирања, указивање на начин решавања конкретних задатака, информације у вези са почетком, трајањем и завршетком испитивања (Мужић, 1968). Упознавање ученика са циљем тестирања одредиће њихово понашање током тестирања. Што је циљ тестирања важнији за наставак школовања, то ће мотивација ученика за решавање теста бити већа, али и притисак ка остваривању што бољег постигнућа (чиме се повећава и вероватноћа варања на тесту). Уколико, пак, циљ тестирања није оцењивање, ученици ће бити више слободни да погађају одговоре, чак и када су предвиђени негативни поени. Истовремено, један број ученика ће у датим околностима мање студиозно приступити решавању задатака (Konrad, 1981).

Пожељно је да сваки примерак теста буде нумерисан и да реализатор тестирања потпише реверс о пријему инструмента. Наведена препорука спречава могућност да испитаник „прокријумчари“ примерак теста. Понекад се на крају тестирања предвиђа и провера враћених примерака теста прелиставањем, како би се утврдило да ли је враћен тест у целини (Konrad, 1981).

Могућности коришћења резултата добијених применом теста знања за унапређивање васпитно-образовног рада школе

Тестирање знања се спроводи како би се резултати теста на неки начин могли употребити (Kean, 1980; Hattie, Jaeger & Bond, 1999). Деценијама уназад у школској пракси је преовладавала примена тестова знања у циљу оцењивања ученика (Porham, 2001). Коришћење тестова у циљу оцењивања ученика у литератури се доводи у везу са контролном функцијом тестова (Хавелка, 2000), односно, пружањем информације ученику о оствареном резултату на тесту.

Крајем прошлог века, најпре у литератури, а касније и у домену праксе почиње да преовладава становиште да је основна функција примене тестова знања унапређивање процеса учења (Ebel, 1979; Cooper & Leiter, 1980; Buser, 1996). Наглашава се да је потребно да интересовање наставника за резултате остварене учењем увек прати и интересовање за процес учења. Оријентисање на процес учења који доводи до одређеног резултата може се довести у везу са инструктивном функцијом тестирања у настави. Инструктивна функција се односи на усмеравање процеса учења, развијање стратегија учења са разумевањем, вештине селектовања битних информација и њиховог уређивања у хијерархијску структуру (Хавелка, 2000).

Да би се наведена очекивања наставника могла испунити, неопходно је да тестови знања буду ослоњени на „стварни“ курикулум, односно да одговарају васпитно-образовном процесу који се реализује у учионици (Cooper & Leiter, 1980). Када се тако поставе ствари, наставник на основу резултата теста може да одлучи да битно промени садржај или начин свог рада, уколико се покаже да су резултати теста испод очекивања, или да настави да ради неизмењено уколико резултати буду задовољавајући. Тестови знања могу истовремено бити примењени и са циљем оцењивања ученика и у циљу предлагања мера за унапређивање процеса учења. Наставник може најпре утврдити резултате које ученици остварују на тесту, а потом донети одлуке у вези са даљим радом у настави.

Да би се искористио потенцијал који тестови носе на плану унапређивања процеса учења, потребно је да ученици доживљавају тестове као помоћ у учењу, и да им је обезбеђена благовремена повратна информација о постигнућу (Гргин, 2001; Dihoff, Brosvic & Epstein, 2003).

Може се рећи да је пружање повратне информације од пресудне важности у контексту остваривања инструктивне функције тестирања. Експериментално испитивање Војдановске, Кранија и Њула (Vojdanovska, Cranney & Newell, 2010) показало је да повратна информација о тачним решењима задатака у иницијалном тесту унапређује постигнуће ученика на финалном тестирању. Постигнути ефекат објашњава се као ефекат формативне евалуације. Током последње две деценије све се више наглашава важност формативног процењивања или процењивања у функцији учења (*assessment for learning*) (Кузмановић и Павловић-Бабић, 2011). Предност формативне над сумативном евалуацијом са становишта резултата учења, мотивације и самоефикасности ученика доказују и друга истраживања у области образовања (на пример: Lipnevich & Smith, 2008; Zimmerman & DiBenedetto, 2008). Корак даље од процењивања постигнућа оријентисаног на учење, приступна је и оријентација на „процењивање као учење“ (*assessment as learning*) које има за циљ постизање индивидуализованих циљева учења кроз процес учениковог самовредновања (Earl, 2003).

Занимљиво, испитивање Липњевић и Смита (Lipnevich & Smith, 2008) показало је да се најбољи ефекат постиже давањем искључиво коментара постигнућа, без извођења оцене. Садржински адекватна повратна информација има позитивне ефекте на рад ученика независно од тога да ли потиче од наставника или је компјутерски генерисана.

Пошто се тестирање реализује, наставник може спровести анализу резултата остварених на тесту. Анализа успешности у решавању појединачних задатака на тесту може бити корисна за наставника, како са становишта утврђивања којим деловима градива су ученици боље овладали, а којим слабије, тако и са становишта повратне информације о квалитету задатака у тесту. Таква сазнања могу помоћи наставнику да успешније планира рад у

допунској и додатној настави, као и да информише родитеље и разредне старешине о напредовању ученика у учењу.

Наставник се може одредити за квалитативну, квантитативну обраду података добијених тестирањем, или за комбинацију оба модела. У домену квалитативне анализе резултата теста, евидентира се и описује успешност у решавању сваког појединачног задатка, за сваког од ученика у одељењу. Овакав вид бележења ученичког постигнућа омогућава да се прецизно утврди којим делом градива ученици нису овладели, да се препознају ученици којима је потребна додатна помоћ у раду, односно ученици који се посебно истичу у погледу постигнућа (Hamzah & Abdullah, 2011).

Резултате остварене на тесту, наставник може изразити и квантитативно. Може се одредити за табеларни или графички приказ резултата, означавајући фреквенцију одговарајућих скорова на нивоу одељења. Свакако, и финална статистичка анализа може наћи своје место у приказу резултата тестирања (Taylor, Greer & Mussio, 1978). Испитивања показују да се анализа резултата најчешће своди на израчунавање постотка ученика који су на тесту добили пролазну оцену, односно постотка ученика који у томе нису успели. Врло ретко, израчунава се просечни скор, медијана, мод, стандардна девијација, индекс тежине и дискриминативности задатка (Al-Shara'h, 2013).

За разлику од тестова које наставници конструишу, циљеви израде стандардизованих тестова у склопу завршног испита, националних и међународних тестирања, превазилазе оквире конкретне учионице и школе, па и образовног система једне државе. Могућности коришћења података добијених применом стандардизованих тестова су велике, с обзиром на сложену концепцију истраживања, високу стандардизацију услова тестирања и величину узорка на којима се тестирања спроводе.

Стратегијом развоја образовања до 2020. године у нашој земљи предвиђено је да се резултати Мале матуре користе: а) како би се проценили ефекти основног образовања и васпитања, б) да се повратно делује на обликовање процеса наставе и учења у основној школи и в) да се омогући диференцирање ученика при упису у средњу школу. Додатно, очекује се да

завршни испит као форма екстерног испита обезбеди контролу реализације курикулума у школи, омогући праћење остварености образовних стандарда, али и мотивише наставнике и ученика на рад (*Квалитетно образовање за све – пут ка развијеном друштву*, 2002).

Приступ резултатима завршних испита као мерилу успешности образовног система, има одређене последице по наставни процес. Добра илустрација за то јесте стање у америчком школском систему. Последњих деценија XX века у Америци резултати тестова знања на завршним испитима посматрали су се као мерило успешности образовног система, школе и рада појединачних наставника. Последице таквог приступа биле су вишеструко неповољне: школе чији ученици не би остварили очекивано добре резултате на тестовима биле би у опасности да им се драстично смањи финансирање, наставницима умањи број плаћених сати рада, а родитељима дозволи да испишу децу из „лоше“ школе (Popham, 2003; Rivera, 2007). Наставници схватају да од резултата тестова почиње да зависи рејтинг школе и њихова материјална ситуација, те се у датим околностима фокусирају на што бољу припрему ученика за финалне тестове. Неминовно, сваки садржај који неће бити обухваћен тестом, престаје да буде важан и изоставља се. Дешава се масовно редуковање школског курикулума, а настава се своди на подучавање ученика тачним одговорима за финални тест (Popham, 2001; Gipps, 2003; Popham, 2003; Fluckiger, 2004). Истраживање Хермана и Голан (Herman & Golan, 1991) показало је да су наставници припреми ученика за тестирање посвећивали, у просеку, од једне до четири недеље у току школске године. У данима пред тест, са ученицима су решавали неке од комерцијалних¹¹ тестова или прошлогодишње тестове. Највише времена у настави посвећивано је управо оним садржајима који ће се наћи на тесту. Испитивање је показало да

¹¹ Од прве половине двадесетог века, па до данас, у Сједињеним Америчким Државама појављују се специјализоване компаније које конструишу тестове. Тестови које компаније израђују међусобно се разликују, с обзиром на то да америчке државе могу имати посебне системе образовања и посебан систем стандарда знања које дефинишу. Занимљив је податак да је продаја тестова 1955. године компанијама донела профит од близу 7 милиона долара. У 1997. години та цифра је износила преко 263 милиона долара (Clarke *et al*, 2000), а процењује се да је 2008. године профит износио 1,1 милијарди долара (Vu, 2008).

наставници не деле мишљење да су резултати тестова показатељ квалитета рада наставника. Насупрот таквом становишту, наставници оцењују да постоји раскорак између онога што тестови испитују и садржаја који би требало да буде наглашен у наставном раду (Herman & Golan, 1991).

Хати, Јејг и Бонд (Hattie, Jaeger & Bond, 1999) посебно разматрају позитивне, а посебно негативне ефекте тестирања. Када је реч о позитивним искуствима, она се односе на то да су наставници више фокусирани на садржај курикулума него раније, и теже да остваре све циљеве дефинисане курикулумом, како би се ученици што боље припремили за тест. Негативни ефекти су, чини се, бројнији: мање се обраћа пажња на стратегије учења код ученика, повећава се број ученика који напуштају школу, наставницима је теже да мотивишу ученике на перманентно образовање, опада поверење јавности у ефикасност школског система (пошто се највише пажње посвећује школама које не остварују добре резултате).

Важност резултата тестирања за будућност младих и школе довела је до непопуларне праксе „варања“ на тесту, не само ученика већ и наставника. Све је више истраживача који показују интересовање за ову врсту проблема, бавећи се разлозима за такво поступање, питањима заступљености варања у ситуацији тестирања, питањима начина на који се то остварује, као и предлагањем начина да се таква пракса детектује на што објективнији начин (Porham, 2001; Jacob & Levitt, 2004; Van der Linden & Jeon, 2012).

Пофам (2001) указује на неке од типичних начина на које наставници директно утичу на резултате тестова: а) продужавање времена за решавање теста; б) давање наговештаја тачних одговора на поједине ајтеме у тесту; в) прегледање предатих тестова и враћање ученицима да још једном „размисле“ или „провере да ли су одговорили тачно“. Нетипичне, али забележене праксе су, директно преправљање одговора на тесту пошто ученик преда свој рад (Jacob & Levitt, 2004) или упућивање ученика да одговоре најпре напишу на помоћном листу папира, како би наставник могао да их провери (Porham, 2001). Више софистициран начин деловања на резултате тестова јесте „подучавање за тест“. У Америци наставници имају увид у садржај и форму

задатака који се могу наћи на тесту, тако да се одлучују да са ученицима вежбају управо идентичне задатке или „дупликате“ задатака, који су незнатно другачији од задатака који ће се појавити у тесту (Porham, 2001: 23). Озбиљан недостатак такве праксе је поистовећивање предстојеће успешности у решавању увежбаних задатака са бољом савладаношћу градива, што не мора бити случај (Porham, 2001).

Колико год да је варање у ситуацији тестирања озбиљан проблем, реч је о пракси која није толико распрострањена да би угрожавала интегритет наставничке професије. Процењује се да је у Америци од 3 до 6% наставника склоно таквој пракси. Врло често се као разлог наводи њихов страх да ће ученици подбацити на тесту, барем делимично зато што одрастају у неповољним околностима (сиромашне породице, припадност мањинској групи), чији лош утицај школа не успева да компензује (Jacob & Levitt, 2004). Вишегодишње праћење резултата тестова минималних компетенција у Америци, показало је да ученици једноставно не уче више него пре увођења датих тестова, чак је и број ученика који напуштају школу у порасту (Catterall, 1989).

Пофам (2003) разматра услове под којима стандардизовани тестови знања могу допринети унапређивању наставног процеса. Први услов је њихово утемељење на стандарима знања. За наставнике је најкорисније уколико резултати теста знања могу да укажу на успешност сваког појединачног ученика на сваком од испитаних стандарда знања. Све док наставник не сазна које стандарде ученици нису усвојили, неће моћи да унапреди свој рад у настави. Извештавање о успешности ученика на сваком од предвиђених стандарда знања и даље омогућава поређење ученика и школа, тако да једна од полазних интенција примене стандардизованих тестова и даље може бити остварена.

Други услов је да приликом конструкције стандардизованих тестова, задаци у тесту обухвате само неколико кључних стандарда знања предвиђених курикулумом. Пофам (2003) подсећа да је корисније тестом темељно испитати неколико кључних стандарда знања, него површно испитати сваки од

предвиђених стандарда. Како би се избегла редукција курикулума, Пофам (2003) предлаже да се школама омогући приступ серијама инструмената путем којих се може испитивати постигнуће ученика у вези са стандардима који неће бити заступљени на завршним тестовима, али је за ученике корисно да овладају референтним знањима и вештинама.

Трећи услов је да знања и вештине у основи кључних стандарда буду описани на начин који је наставницима разумљив. На тај начин, наставни процес неће бити оптерећен подучавањем ученика тачним решењима задатака у тесту, већ усмерен ка остваривању јасних и конкретних образовних исхода (Popham, 2003).

Стратегија развоја образовања до 2020. године у нашој земљи предвиђа да се резултати националних и међународних испитивања користе за: а) унапређивање наставног процеса и учења и б) креирање образовне политике (*Стратегија развоја образовања у Србији до 2020. године*, 2012). Да би се дефинисани циљеви могли остварити, неопходно је довести у везу постигнуће ученика са контекстом у коме се постигнуће остварује. Применом стандардизованих тестова стиче се увид у знања и вештине које су ученици стекли на одређеном стадијуму образовања. Применом других истраживачких техника (анкетирањем, интервјуисањем, анализом садржаја) прикупљају се информације о околностима у којима се постигнуће остварује: социоекономском статусу породица ученика, мотивацији ученика за учење, мерама подршке у учењу, школској клими, наставним методама, стручности наставника и њиховим професионалним активностима, материјално-техничким условима у којима раде школе. Укрштањем одговарајућих података тежи се објашњавању резултата које ученици остварују, уочавају се слабости једног образовног система у поређењу са системом који се показао успешнијим. На темељу уочених разлика могу се планирати промене у домену образовне политике једне земље. На пример, учешће у PISA тестирању 2000. године довело је до корених промена у образовном систему Немачке: курикулум је централизован, успостављена је већа контрола образовања учитеља и

наставника и финансирања школа, дефинисани су национални стандарди постигнућа ученика (Максимовић, 2013).

Иако би се могло рећи да је пут од тестирања до предложених решења логичан, његова реализација није једноставна. Потребне науке и потребе праксе у области педагошких истраживања није лако усагласити. Научна сазнања посебно је тешко пренети у област просветне политике и праксе (Keiner, 2002; према: Вујисић-Живковић, 2013). Дискусије о раскораку између теорије и праксе надовезују се на расправу о томе како конципирати истраживање које је „корисно“, шта је оно што „даје резултате“ („*what works*“) у педагогији (Howe, 2004), колико су за педагогију значајна истраживања за која не можемо да кажемо да ће нужно унапредити образовну политику и праксу (Whitty, 2006; Bridges, Smeyers & Smith, 2008), који су домети „на доказима засноване“ (*evidence-based*) концепције реформисања образовања (Вујисић-Живковић, 2013).

Када је реч о националним тестирањима, полази се од уверења да је њихова реализација корисна, јер обезбеђује заједничку основу за поређење појединаца и школа (Davey, 1992; према: Rudner & Schafer, 2002). Међутим, тешко је доказати да такви тестови сами по себи воде позитивним променама у образовању. Такође, стандарди знања на којима су тестови у националним испитивањима утемељени ретко имају подршку целе стручне јавности, а поједина испитивања указују на изражену пристрасност тестова који се користе у националним тестирањима (Davey & Neill, 1992; према: Rudner & Schafer, 2002).

Када је реч о резултатима међународних испитивања, полази се од становишта да су поређења образовних система могућа. Противници таквог становишта подсећају да су разлике које се идентификују испитивањем само делимично одраз разлика у образовном систему. Међународна испитивања полазе од резултата које ученици остваре на тесту. Међутим, јединица анализе је школа, а не ученик. Без контроле над полазним разликама међу ученицима не може се бранити основаност процене квалитета рада школе на основу ученичких постигнућа (Chatterji, 2013). Голдштајн (Goldstein, 2004) и Рошекс (Rochex, 2006) истичу да је исправније друштвене и културне специфичности

узети у обзир, него конструистати тестове који су неосетљиви на ту врсту разлика. Постоје и друга органичења међународних студија која доводе у питање могућност генерализације њихових закључака. Ограничења се доводе у везу са неусаглашеношћу наставних садржаја који се изучавају у школама различитих земаља учесница и немогућношћу формирања узорка испитаника према унапред дефинисаном моделу у свим земљама учесницама.

Питања на која, дакле, немамо коначне одговоре су: можемо ли говорити о томе да су образовни системи заиста упоредиви; да ли резултат који ученик оствари на тесту знања довољно говори о квалитету наставног рада у школи коју похађа; да лоше постигнуће ученика на тесту знања несумњиво говори о томе да његов наставник не ради добро свој посао?

Пофам (1999) указује на неколико аргумената због којих наведена становишта нису одржива, а која се тичу методолошких карактеристика самих стандардизованих тестова. Први аргумент од кога Пофам полази односи се на садржај самих стандардизованих тестова. Наиме, интенција која стоји иза конструкције стандардизованих тестова знања била је да се њима захвати врло широк опсег знања које ученици стичу у школи. Међутим, да би се тестом обухватило „све што је битно“, тест би морао да буде доста дуг. У циљу конструкције теста који би се решавао у оптималном времену (на пример, за сат времена), тест мора да буде знатно краћи, и самим тим мора да захвати само „узорак“ школског градива. Додатна потешкоћа за компаније које се баве конструкцијом стандардизованих тестова у САД је што се образовни циљеви могу разликовати од државе до државе. Како би ипак успели у намери да своје тестове пласирају на што ширем подручју, компаније се одлучују да понуде садржај који ће „углавном одговарати свима“ (Popham, 1999: 4). Последице, дешава се значајна дискрепанца између онога што тест мери и онога што је садржај наставног процеса (Banks, 1980).

Следећи аргумент који заокупља Пофамову пажњу, односи се на психометријску тенденцију да се након пробног испитивања из теста искључе задаци који нису довољно дискриминативни. То су ајтеми које решава велики број ученика, или ајтеми које решава мали број ученика. Пофам задржава

пажњу на ајтемима које решава велики број ученика и доводи их у везу са наставним процесом. Ако је наставник одређеном садржају посветио посебну пажњу, очекује се да ће већина ученика ајтеме који се на тај садржај односе решавати успешно. То нас води ка закључку, да што су наставници успешнији у подучавању одређених садржаја, то је мања шанса да ће се такви садржаји наћи на тесту знања (Popham, 1999; Popham, 2001). Пофам не упућује посебан коментар који би се односио на категорију задатака које решава мали број ученика. Могло би се очекивати да је реч о задацима које већина ученика не решава успешно због сложености захтева (или необрађености у настави) или неадекватне формулације (двосмислености, лоших дистрактора, употребе непознатих речи или стручних термина).

Следеће питање које уобичајену намену стандардизованих тестова чини спорном је: шта све поред знања стеченог у школи мере ајтеми у тесту? Пофам низом различитих примера показује да одређени ајтеми заиста мере знање које је стечено у школи. Истовремено, примерима доказује и да постоје ајтеми који мере интелектуалну сналажљивост ученика, пре него било какво школско знање.¹² Образложење зашто се такви ајтеми појављују у тестовима, Пофам проналази у захтевима психометрије – потребно је обезбедити што правилнију дистрибуцију резулата добијених тестом. Највеће шансе да се то оствари јесте да се ослонимо на мерење својства које се нормално распоређује у популацији, као што је интелигенција (Popham, 1999; Popham, 2001; Popham, 2003). На узорку готово свих доступних стандардизованих тестова знања за основну школу на почетку новог миленијума, Пофам (2001) је утврдио да се заступљеност таквих ајтема креће од 20% у области математике, преко 35% у области језичке културе, до 55% у области природних наука. Несумњиво, такви ајтеми ни на који начин не упућују на квалитет образовног процеса.

¹² Један од примера који Пофам приказује је и следећи:

Који од следећих изума је омогућио фармерима да производе храну?

1. усисивач
2. машина за веш
3. камера
4. плуг

(Popham, 2001: 70)

Такође спорни су и ајтеми који се односе на свакодневна искуства ученика¹³. Аутор закључује да такви ајтеми праве разлику међу ученицима сходно социоекономском статусу њихових породица. Разумљиво, они не говоре о квалитету наставног процеса, али постају део тестова знања из истог разлога као и ајтеми који мере интелектуалну сналажљивост, како би дистрибуција добијених одговора одговарала нормалној расподели (Popham, 1999; Popham, 2001; Popham, 2003). На узорку готово свих тада доступних стандардизованих тестова знања за основну школу, Пофам (2001) је утврдио да се заступљеност таквих ајтема креће од 5% у области математике, преко 45% у области природних наука, до чак 65% у области језичке културе.

Пофам даје још један интересантан пример. Замислите да сте директор школе коју уписују махом деца из сиромашних породица. Шта ће се догодити ако у тесту преовладају задаци који се односе на свакодневно искуство ученика? Највероватније, ученици Ваше школе ће подбацити на тесту. Да ли то значи да наставници у Вашој школи не раде добро свој посао? „Наравно да не значи“ (Popham, 1999: 8). Напоменимо да је на неправедност оцењивања квалитета рада школе на основу постигнућа ученика који потичу из породица различитог социоекономског статуса указано и у неким ранијим радовима (на пример: Banks, 1980).

Када све наведене чиниоце имају у виду, шта запослени у образовању могу да ураде? Пофам (1999) саветује да свако ко је заинтересован посвети време анализи садржаја стандардизованих тестова како би се уверио шта тестови заиста мере. Такође, потребно је повећати свест свих актера васпитно-образовног процеса да такви тестови не могу да мере квалитет наставе. Наставницима се, такође, саветује да пронађу алтернативне начине како би показали резултате свог рада – на пример, јачањем когнитивних вештина као

¹³ Задатак који Пофам коментарише гласио је:

Плодови воћа увек имају семе. Који од наведених примера није воће?

1. поморанца
2. бундева
3. јабука
4. целер

(Popham, 1999: 8; Popham, 2001: 62)

што је решавање сложених математичких проблема или примена знања из историје како би се анализирали актуелни политички проблеми. Уколико се за то одлуче, наставници могу самостално да креирају иницијалне и завршне тестове којима би показали, родитељима или школским надзорницима, колико су ученици напредовали у току школске године (Porham, 1999).

Пофам (2001) не пориче да је перцепција успешности образовног система у јавности изједначена са успешношћу на тестовима знања. Квалитет наставе доводи се у директну везу са резултатима тестова. У начелу, Пофам није противник идеје о стандардизованим тестовима знања. Његова критика је упућена постојећим тестовима који нису осетљиви на потребе наставне праксе, односно не помажу наставницима да унапреде свој рад са ученицима. Дакле, није погрешна идеја о постојању стандардизованих тестова знања, већ су постојећи тестови „лоши“ (Porham, 2001: 18).

Упркос свим критикама упућеним стандардизованим тестовима знања, без њихове примене готово да се не може добити јасна слика о томе како се спроводи школски курикулум (Brennan, 2001; Phelps, 2006). Испитивање Фелпса (Phelps, 2006) показало је да однос присталица стандардизованих тестова знања и њихових опонената у Америци износи 12:1. Како Фелпс истиче, готово сваки Американац сматра да је корисно „да се такви тестови користе барем једанпут годишње“ (2006: 19). Фелпс се у свом раду осврће на најчешће критике упућивање стандардизованим тестовима. Коментаришући појаву редуковања наставног садржаја услед увођења стандарда знања, Фелпс истиче да када не би било дефинисаних стандарда, наставни садржаји не би били нужно обухватнији. Заправо, десило би се да буду редуковани на оне садржаје за које наставници лично процене да су значајнији од других. Када је реч о склоности „варању“ од стране ученика у ситуацији решавања стандардизованих тестова, Фелпс подсећа да је такво понашање карактеристично за готово сваку ситуацију испитивања писаним путем. Разлика је само у томе што је, у стандардизованим околностима тестирања, такву појаву много лакше детектовати него у другим околностима. Склоност наставника ка „варању“ у ситуацији тестирања, такође не би требало некритички везивати само за

решавање стандардизованих тестова. Фелпс упозорава да је инфлација оцена тренд који је преплавио школе независно од примене стандардизованих тестова. Наставници сами признају да су генерално под притиском да дају високе оцене ученицима који их не заслужују (Phelps, 2006). Аутор се осврће и на критике да је поређење ученика са „нормом“ неправедно. Фелпс је један од присталица становишта да нормативни тестови пружају информације које су вишеструко корисне за све запослене у образовању и које се не би могле добити на неки други начин (Brennan, 2001; Phelps, 2006). Алтернатива поређењу постигнућа ученика са нормом за дати узраст је поређење оцене коју ученик има са оценама других ученика. Према становишту Фелпса, оцене се такође могу посматрати као „норме“, али норме које важе на нивоу конкретне школске установе. На крају, аутор се осврће и на наглашавање разлике између тестова које наставници сами праве и стандардизованих тестова. У критикама, обично се наглашава да су тестови које наставници самостално израђују боље усклађени са потребама школске праксе. Међутим, Фелпс подсећа да стандардизоване тестове конструишу особе које су најстручније у домену одређене науке или наставне области и чије познавање теорије и принципа мерења знања далеко превазилази знање које у тим областима поседују појединачни наставници (Phelps, 2006).

Поставља се питање да ли наставници могу да користе резултате стандардизованих тестова како би унапредили сопствени рад у настави? Таква могућност постоји уколико су наставницима доступни резултати које остварују појединачни ученици, што није увек случај. Резултате добијене испитивањем, наставници могу користити како би идентификовали ученике којима је потребна помоћ у учењу, односно ученике који остварују високе резултате и како би известили родитеље о постигнућу, односно напредовању ученика у учењу (*Assessment in the primary school curriculum: guidelines for schools*, 2007).

Када је реч о ученицима са тешкоћама у учењу, резултати стандардизованог теста могу да помогну да се идентификује њихова заступљеност. Међутим, резултати мало говоре о природи проблема коју ученик има. У циљу утврђивања узрока потешкоћа, потребно је применити

додатне дијагностичке тестове. Такође, када је реч о ученицима који остварују високо постигнуће, резултати испитивања помажу наставнику само да такве ученике идентификује. Даљи кораци у вези са радом на развијању талента ученика зависиће од могућности које школа пружа у том домену, пре свега од могућности самог наставника да подржи даљи развој ученика кроз редован наставни рад, рад у оквиру додатне наставе и одговарајућих секција.

Када се има у виду разноврсност података коју екстерна испитивања обезбеђују, чини се да би сваком наставнику било корисно да се информише о резултатима тестирања. Могућности да се такав потенцијал искористи ограничава више чинилаца. Најпре, екстерна испитивања у нашој земљи реализују се повремено у последњих десет година. Пошто нису дуго заступљена на нашем подручју и нису реализована у свим школама нашег образовног система, још увек се не доживљавају као интегрални део школске праксе. Истовремено, уколико је наставник провео само кратко време радећи у једној школској установи, он није у могућности да прати процес истраживања, стекне увид у то како истраживање напредује и када се могу очекивати резултати. Можда најважније од свега је то да је улога наставника у процесу екстерних испитивања често сведена на улогу техничког лица које нема увид у полазне оквире испитивања и нема могућност да учествује у интерпретацији резултата. Резултати националних и међународних испитивања у нашој земљи немају последице у смислу материјалног подстицаја или санкције за рад наставника. Када се имају у виду наведене околности, разумљиво је да ће тек мали број наставника бити заиста мотивисан да учешће у испитивању схвати као део своје професионалне активности, да се детаљније информише о резултатима тестирања и евентуално се одлучи да начини промену у свом раду.

МЕТОДОЛОГИЈА ИСТРАЖИВАЊА

ПРЕДМЕТ ИСТРАЖИВАЊА И ДЕФИНИСАЊЕ ОСНОВНИХ ПОЈМОВА

Предмет истраживања је примена тестова знања у пракси основних школа.

Под тестовима знања у овом раду подразумеваће се инструменти израђени у сврху проверавања знања ученика. Конкретније, реч је о тестовима знања које наставници примењују у наставном раду и чијој конструкцији могу приступити самостално, ослонити се на помоћ колега и стручних сарадника, или користити готове инструменте. Истраживањем се обухватају и тестови знања намењени проверавању знања ученика који су део екстерне евалуације рада школе и то: Мале матуре, националних и међународних испитивања ученичких постигнућа.

Специфичност тестова као инструмената је да се састоје од међусобно повезаних задатака који су истоврсни за све испитанике, прецизирани су начини њиховог решавања и планирани су начини вредновања постигнутих резултата. У ситуацији тестирања, испитаник користи један од следећих начина решавања задатака у тесту: заокружује или подвлачи тачан одговор, уписује одговарајући симбол, уписује одговор речима, примењује рачунске операције над бројевима, конструише решење задатка. Решавање теста може бити ручно или посредством рачунара.

Примена тестова знања у школи разматра се на три плана. Први план подразумева могућности примене тестова знања у наставној пракси које би биле сагледане на основу разматрања: тестирања као начина проверавања знања ученика; процене наставника о властитој припремљености за израду тестова знања, тешкоћама у процесу израде тестова знања и начинима њиховог превазилажења; учесталости и динамике коришћења тестова знања у настави.

Други план се односи на карактеристике тестова које наставници самостално израђују и које би биле сагледане на основу разматрања процеса конструкције и методолошких карактеристика тестова знања са становишта садржаја, структуре и начина утврђивања постигнућа на тесту.

Трећи план се односи на начине на које се могу користити резултати добијени применом тестова знања са становишта унапређивања васпитно-образовног рада школе.

ЦИЉ И ЗАДАЦИ ИСТРАЖИВАЊА

Циљ истраживања је да се испитају могућности примене тестова знања у наставној пракси, карактеристике примењених тестова и начине на које се могу користити резултати добијени применом тестова знања.

У складу са циљем истраживања, постављени су следећи задаци:

1. испитати у чему наставници виде предности, а у чему ограничења тестова знања као начина провере знања ученика
2. испитати како наставници процењују властиту припремљеност за израду тестова знања
3. утврдити врсте тешкоћа са којима се наставници суочавају у процесу израде тестова знања и начине њиховог превазилажења
4. утврдити учесталост и диманику коришћења тестова знања током школске године
5. регистровати фазе конструкције теста и утврдити методолошке карактеристике тестова знања које наставници самостално израђују: садржај, структуру теста знања, као и начин утврђивања постигнућа на тесту
 - испитати да ли постоји разлика у фазама и методолошким карактеристикама тестова знања које наставници српског језика и математике самостално израђују
6. испитати на које начине се могу користити резултати добијени испитивањем образовних постигнућа ученика са становишта унапређивања васпитно-образовног рада школе

7. испитати да ли постоје статистички значајне разлике међу наставницима друштвено-језичке и природно-математичке групе предмета у:
- сагледавању предности и ограничења тестова знања као начина провере знања ученика
 - процени властите припремљености за израду тестова знања
 - врсти тешкоћа са којима се суочавају у процесу израде тестова знања и начинима њиховог превазилажења
 - учесталости и динамици коришћења тестова знања у настави
 - сагледавању начина коришћења резултата добијених испитивањем образовних постигнућа ученика са становишта унапређивања васпитно-образовног рада школе
8. испитати да ли постоје статистички значајне разлике, међу наставницима различитих година радног стажа у:
- сагледавању предности и ограничења тестова знања као начина провере знања ученика
 - процени властите припремљености за израду тестова знања
 - врсти тешкоћа са којима се суочавају у процесу израде тестова знања и начинима њиховог превазилажења
 - учесталости и динамици коришћења тестова знања у настави
 - сагледавању начина коришћења резултата добијених испитивањем образовних постигнућа ученика са становишта унапређивања васпитно-образовног рада школе.

ХИПОТЕЗЕ У ИСТРАЖИВАЊУ

У раду се полази од следећих хипотеза:

1. могућности за примену тестова знања у наставној пракси су повољне, с обзиром на то како наставници сагледавају предности и ограничења тестова знања у односу на друге поступке испитивања знања ученика,

властиту процену припремљености за израду тестова знања, и тешкоће са којима се суочавају у изради тестова

2. фазе конструкције теста и методолошке карактеристике тестова знања које наставници самостално израђују одступају од фаза и карактеристика тестова знања дефинисаних у методологији педагошких истраживања
 - 2.1. постоји разлика у фазама конструкције и методолошким карактеристикама тестова знања који су примењени у настави српског језика и математике
3. резултати добијени применом тестова знања, доминантно се користе у сврху пружања повратне информације о савладаности градива од стране ученика, али се у извесној мери користе и за унапређивање васпитно-образовног рада школе
4. постоје статистички значајне разлике међу наставницима друштвено-језичке и природно-математичке групе предмета у:
 - сагледавању предности и ограничења тестова знања као начина провере знања ученика
 - процени властите припремљености за израду тестова знања
 - врсти тешкоћа са којима се суочавају у процесу израде тестова знања и начинима њиховог превазилажења
 - учесталости и динамици коришћења тестова знања у настави
 - сагледавању начина коришћења резултата добијених испитивањем образовних постигнућа ученика са становишта унапређивања васпитно-образовног рада школе
5. постоје статистички значајне разлике међу наставницима различитих година радног стажа у:
 - сагледавању предности и ограничења тестова знања као начина провере знања ученика
 - процени властите припремљености за израду тестова знања
 - врсти тешкоћа са којима се суочавају у процесу израде тестова знања и начинима њиховог превазилажења

- учесталости и динамици коришћења тестова знања у настави
- сагледавању начина коришћења резултата добијених испитивањем образовних постигнућа ученика са становишта унапређивања васпитно-образовног рада школе.

ВАРИЈАБЛЕ У ИСТРАЖИВАЊУ

Независне варијабле у истраживању су: школски предмет који наставник предаје и године радног стажа наставника.

За потребе истраживања, школски предмети су груписани на следећи начин:

а) друштвено-језичка група предмета: српски језик, страни језици, историја, веронаука и грађанско васпитање

б) математика и предмети природних наука: математика, биологија, физика, хемија, географија, информатика и техничко образовање.

Варијабла „године радног стажа наставника“, обухватила следеће категорије: мање од 10 година; од 10 до 20 година; преко 20 година.

Зависне варијабле су:

1. процена наставника о предностима и ограничењима тестова знања као начина проверавања знања ученика
2. наставничка процена власите припремљености за израду тестова знања
3. врсте тешкоћа са којима се наставници суочавају у процесу израде тестова знања и начини њиховог превазилажења
4. учесталост и динамика коришћења тестова знања у настави
5. начини на које се користе резултати добијени применом тестова знања.

Предности и ограничења тестова знања сагледане су у односу на друге начине испитивања ученика и то са становишта: обухвата градива, објективности, економичности и аутентичности ученичких одговора.

Наставничка процена власите припремљености обухватала је: процену укупне припремљености за израду тестова знања; процену припремљености за израду тестова знања у току студија и процену припремљености за израду

тестова знања посредством различитих облика стручног усавршавања, са посебним освртом на похађање семинара стручног усавршавања.

Врсте тешкоћа са којима се наставници суочавају у процесу израде тестова знања сагледане су са становишта конструкције, примене и тумачења резултата теста знања. Начини превазилажења тешкоћа доведени су у везу са различитим облицима сарадње са колегама и стручним сарадницима школе, самосталним истраживањем стручне литературе, применом методе „покушаја и погрешака“.

У раду су разматрани начини коришћења резултата Мале матуре, националних и међународних тестирања. Коришћење резултата тестирања доведено је у везу са унапређивањем квалитета наставе, сарадњом са другим учесницима васпитно-образовног процеса у школи и планирањем и организацијом наставних и ваннаставних активности у школи.

Посебан сегмент истраживања односи се на анализу конкретних тестова знања коју наставници самостално израђују за потребе наставне праксе. У том домену, статус независне варијабле има школски предмет, док сама варијабла има два модалитета: српски језик и математика. Статус зависне варијабле имају фазе конструкције теста и методолошке карактеристике тестова знања који су примењени у пракси.

Када је реч о фазама кроз које конструкција тестова знања пролази, опис процеса израде конкретних тестова знања које наставници примењују у настави упоређен је са фазама конструкције теста дефинисаним у методологији педагошких истраживања. То су фазе израде нацрта инструмента, састављања конкретних задатака и финалног обликовања теста знања. Прецизније, реч је о следећим корацима: утврђивање циља који се жели постићи применом теста знања, одређивање садржаја теста, одређивање садржаја задатака, нивоа знања који се задацима мере, тежине задатака, одређивање дужине теста, утврђивање редоследа задатака, избор форме задатака, израда кључа за оцењивање теста знања, утврђивање начина оцењивања теста, састављање упутства за решавање теста, пробно испитивање, формирање, чување и развијање базе задатака.

Методолошке карактеристике примењених тестова знања односиле су се на садржај, структуру теста знања и начин утврђивања постигнућа на тесту. Садржај теста знања процењен је са становишта: циља примене теста, полазне основе за разраду садржаја теста знања, заступљености задатака којима се региструју различити нивои знања (ниво познавања чињеница, ниво разумевања градива, ниво примене), заступљености задатака различите тежине, упутства за решавање теста.

У вези са структуром теста знања регистровани су: 1) дужина теста знања, 2) редослед задатака, 3) заступљеност задатака различите форме (задаци типа бирања одговора, навођења, спаривања и сређивања) и њихове карактеристике (колико алтернатива је понуђено у задатку бирања одговора; број понуђених алтернатива у задатку спаривања; број тачних одговора на нивоу задатка); 4) повезаност примене одређене форме задатка и: а) нивоа знања који задатак утврђује и б) тежине задатка.

Утврђивање постигнућа на тесту обухватило је израду кључа за оцењивање теста и начин оцењивања теста. Као релевантна, издвојена су следећа питања: 1) да ли постоји кључ за оцењивање теста знања; 2) да ли се бодују само потпуно тачни или и делимично тачни одговори; 3) да ли задаци имају исти или различит број поена и какав је распон поена; 4) како се бодују одговори на нивоу различитих типова задатака; 5) да ли има негативних поена; 6) начин утврђивања укупног скорa (уколико има негативних поена); 7) начин на који се поени преводе у оцену.

УЗОРАК ИСТРАЖИВАЊА

У вези са могућностима примене тестова знања и начинима коришћења резултата добијених применом тестова, подаци су сакупљени испитивањем наставника предметне наставе, запослених у основним школама.

Приликом формирања узорка, узета је у обзир чињеница да је Србија подељена на 5 статистичких региона: Београд, Војводину, Јужну и источну Србију, Шумадију и западну Србију и Косово. Истраживањем су обухваћени

сви региони изузев Косова. Када је реч о региону Београда, као највећем статистичком региону у нашој земљи, истраживањем су обухваћене одређене централно-градске школе и школе из појединих приградских насеља. Реч је о следећим школама: ОШ „Народни херој Синиша Николајевић“ (Врачар), ОШ „Вељко Дугошевић“ (Звездара), ОШ „Драгојло Дудић“ (Звездара), ОШ „Марија Бурсаћ“ (Звездара), ОШ „Лаза Костић“ (Нови Београд), ОШ „Радојка Лакић“ (Стари град), и ОШ „Карађорђе“ (Чукарица); ОШ „Јован Ристић“ (Борча), ОШ „Никола Тесла (Винча), ОШ „Момчило Живојиновић“ (Младеновац), ОШ „Вук Караџић“ (Лазаревац), ОШ „Рудовци“ (Лазаревац), ОШ „Јелица Милосављевић“ (Сопот), и ОШ „22.октобар“ (Сурчин). Поред београдских школа, истраживањем је обухваћено још 6 школа које припадају градовима и општинама из статистичких региона Војводина (Кикинда и Панчево), Јужна и источна Србија (Смедерево и Пирот) и Шумадија и западна Србија (Чачак и Лозница). Реч је о следећим школама: ОШ „Фејеш Клара“ (Кикинда), ОШ „Јован Јовановић Змај“ (Панчево), ОШ „Бранко Радичевић“ (Смедерево), ОШ „8. септембар“ (Пирот), ОШ „Танаско Рајић“ (Чачак) и ОШ „Петар Тасић“ (Лозница).

Истраживањем је обухваћено 20 основних школа. Процентуално најзаступљенији су испитаници из Београда и београдских приградских насеља, који заједно чине преко 65% од укупног броја испитаника. Према заступљености следе испитаници из региона Војводине (12,2%) и Шумадије са западном Србијом (12,2%). Процентуално најмање заступљени су испитаници из региона Јужне и источне Србије (8,6%) (табела 8).

Табела 8. Структура узорка према статистичким регионима у којима се налазе школе

Статистички региони	f	%
Београд	243	67,1
Војводина	44	12,2
Шумадија и западна Србија	44	12,2
Јужна и источна Србија	31	8,6
укупно	362	100,0

Узорак је обухватио 362 наставника предметне наставе. Структура узорка према полу показује да наставнице чине скоро 80% од укупног броја испитаника (табела 9). Такав податак не изненађује, јер су припаднице женског пола бројније међу запосленима у школама и на нивоу читаве популације.

Табела 9. Структура узорка према полу испитаника

Пол	f	%
женски	289	79,8
мушки	73	20,2
укупно	362	100,0

Већина испитаника завршила је факултет (95%), а мањи број вишу школу (5%) (табела 10). Када је о факултетима реч, поред факултета београдског универзитета (Филолошки факултет, Математички факултет, Филозофски факултет, Географски факултет, Биолошки факултет, Физички факултет, Хемијски факултет и тако даље), заступљени су и факултети осталих универзитетских центара широм региона (на пример: Филозофски факултет у Новом Саду; Природно-математички факултет у Крагујевцу; Филозофски факултет у Нишу; Филолошки факултет у Приштини; Филозофски факултет у Сарајеву; Филозофски факултет у Никшићу; Филолошки факултет у Скопљу; Филозофски факултет у Задру).

Табела 10. Структура узорка према завршеној школи испитаника

Завршена школа	f	%
факултет	344	95,0
виша школа	18	5,0
укупно	362	100,0

За потребе истраживања, према предмету који предају, наставници су груписани у две категорије: друштвено-језичку и природно-математичку (табела 11).

Табела 11. Структура узорка према групи предмета којој припада предмет који наставник предаје

Група предмета	f	%
друштвено-језичка	187	51,7
природно-математичка	175	48,3
укупно	362	100,0

Када су у питању године радног стажа наставника у просвети, преко 40% испитаника има мање од 10 година радног стажа. Такав податак указује на релативно „младу“ структуру узорка у истраживању (табела 12).

Табела 12. Структура узорка према годинама радног стажа наставника у просвети

Године радног стажа	f	%
мање од 10	154	42,5
10-20	142	39,2
више од 20	66	18,2
укупно	362	100,0

За потребе истраживања, из постојећег узорка испитаника издвојен је подузорак који су чинили наставници српског језика и математике. И овог пута, процентуално најзаступљенији су испитаници из Београда и београдских приградских насеља, који заједно чине 70% подузорка (табела 13).

Табела 13. Структура подузорка према статистичким регионима у којима се налазе школе

Статистички региони	f	%
Београд	56	70,0
Војводина	8	10,0
Шумадија и западна Србија	8	10,0
Јужна и источна Србија	8	10,0
укупно	80	100,0

Структура подузорка према полу показује да наставнице чине 90% од укупног броја испитаника у подузорку (табела 14).

Табела 14. Структура подузорка према полу испитаника

Пол	f	%
женски	72	90,0
мушки	8	10,0
укупно	80	100,0

Већина испитаника у подузорку завршила је факултет (88,7%), а мањи број је завршио вишу школу (11,3%) (табела 15).

Табела 15. Структура подузорка према завршеној школи испитаника

Завршена школа	f	%
факултет	71	88,7
виша школа	9	11,3
укупно	80	100,0

Од укупно 80 испитаника у подузорку, половину су чинили наставници српског језика, а половину наставници математике (табела 16).

Табела 16. Структура подузорка према групи предмета којој припада предмет који наставник предаје

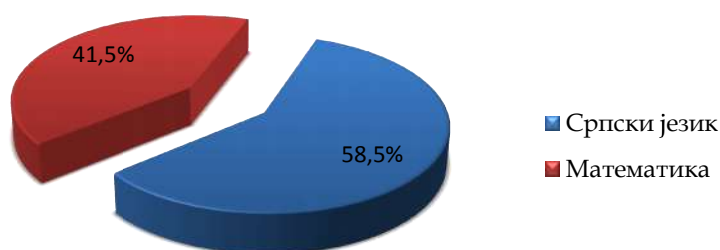
Група предмета	f	%
српски језик	40	50,0
математика	40	50,0
укупно	80	100,0

Када су у питању године радног стажа у просвети, наставници у подузорку испитаника у просеку имају 12,6 година радног стажа. Може се рећи да у погледу година радног стажа испитаника, подузорок испитаника добро одсликава дистрибуцију година радног стажа забележену на нивоу узорка у целини (табела 17).

Табела 17. Структура подзорка према годинама радног стажа наставника у просвети

Године радног стажа	f	%
мање од 10	44	55,0
10-20	19	23,8
више од 20	17	21,3
укупно	80	100,0

Истраживањем је обухваћено 40 тестова из српског језика и 40 тестова из математике. Тестови из српског језика садржали су укупно 382 задатка (58,5%), а тестови из математике 271 задатак (41,5%) (слика 5).



Слика 5. Заступљеност задатака из српског језика и математике у тестовима обухваћеним истраживањем

Нацртом истраживања планирано је да се испитивањем обухвати 400 наставника основне школе (200 наставника друштвено-језичке групе предмета и 200 наставника природно-математичке групе предмета), као и да се реализује минимум 30 интервјуа са наставницима српског језика и минимум 30 интервјуа са наставницима математике. Одзив наставника у првом делу испитивања (попуњавање инструмента за наставнике) није био задовољавајући, а објашњење се барем делимично може потражити у чињеници да се испитивање временски преклопило са штрајком просветних радника. У датим околностима, било је неопходно повећати број школа које су укључене у истраживање. Када се приближио крај школске године, донета је одлука да се

испитивање заустави при броју од 20 посећених школа. На датом узорку школа, добијено је 377 попуњених инструмената. Из даље анализе искључено је 15 инструмената који су се показали непотпуним, чиме је добијен број од 362 комплетна инструмента за наставнике. Када је реч о реализацији интервјуа, сарадња са наставницима је била задовољавајућа. У свакој од 20 посећених школа било је могуће реализовати интервјуе, тако да се број планираних интервјуа повећао са 60 на 80, обухватајући 40 наставника српског језика и 40 наставника математике.

МЕТОДЕ, ТЕХНИКЕ И ИНСТРУМЕНТИ ИСТРАЖИВАЊА

У истраживању је коришћена дескриптивно-аналитичка метода истраживања, технике анкетирања, скалирања, анализе садржаја и интервјуисања. Инструмент намењен наставницима садржи комбинацију питања анкетног типа, петостепене дескриптивне скале и скале ранга. Поред тога, за потребе истраживања припремљен је структурирани протокол интервјуа са наставницима и структурирани протокол анализе садржаја за анализирање примењених тестова знања.

Садржај *инструмента за наставнике* концентрисан је око проблема:

1. могућности примене тестова знања и
2. начина на које наставници користе резултате добијене применом тестова знања.

Наставници су одговарали на питања која се односе на учесталост и динамику коришћења тестова знања у редовној настави предмета који предају. Такође, наставници су попуњавањем одговарајућих дескриптивних скала упоређивали: а) усмено испитивање са тестовима „папир и оловка“; б) писане вежбе са тестовима „папир и оловка“ и в) електронске тестове знања са тестовима „папир и оловка“.

Наставници су имали могућност да путем одговарајућих питања у инструменту извисте о властитој методолошкој припремљености за израду

тестова знања. Наставници су одговарали на питања у вези са: личном проценом укупне припремљености за израду тестова знања, припремом за израду тестова знања током иницијалног професионалног образовања и припремом у домену стручног усавршавања. Одговарајућа питања у инструменту односила су се на испитивање тешкоћа са којима се наставници суочавају у процесу израде тестова знања и начинима њиховог превазилажења. У вези са начинима коришћења резултата наставници су извештавали о: а) циљевима са којима су лично користили резултате тестова знања и б) начинима коришћења резултата Мале матуре, националних и међународних тестирања на нивоу школе. Такође, наведени аспект испитивања обухватио је и испитивање значаја Мале матуре, националних и међународних тестирања за становишта унапређивања васпитно-образовног рада школе.

Структурираним протоколом интервјуа добили смо одговоре на питања у вези са: одређењем циља теста; полазном основом за разраду садржаја теста; задацима којима се региструју различити нивои знања; задацима различите тежине; припремом ученика за тестирање; садржајем упутства за решавање теста; кључем за оцењивање теста; бројем поена које задаци носе; делимичним признавањем одговора, предвиђањем негативних поена за погрешне одговоре; начинима утврђивања укупног скорa и скалом превођења поена у оцене; пробним испитивањем; формирањем, чувањем и развијањем базе задатака; коришћењем готових тестова или задатака.

Структурираним протоколом анализе садржаја прикупљени су подаци који се односе на: дужину теста; редослед задатака у тесту; заступљеност одређеног типа задатака према форми; глаголе и глаголске облике који се употребљавају у формулацији задатка; повезаност примене одређене форме задатка и: а) нивоа знања који задатак утврђује и б) тежине задатка; садржај писаног упутства за решавање теста¹.

¹ У случају да конкретан тест садржи упутство у писаној форми.

НИВО И НАЧИН ОБРАДЕ ПОДАТАКА

Обрада података добијених истраживањем, превасходно је квантитативна, и подразумевала је примену дескриптивне статистике и статистике закључивања. Поред фреквенција и процената, коришћени су и: аритметичка средина, стандардна девијација, хи-квадрат тест, t-тест, анализа варијансе и коефицијент контингенције. Обрада података извршена је у програму SPSS Statistics 17.

ОРГАНИЗАЦИЈА И ТОК ИСТРАЖИВАЊА

Истраживање је спроведено у периоду од марта до јуна 2015. године. Школе су посећене у претходно договореним терминима са директорима школе. Приликом прве посете школи, директорима или стручним сарадницима су остављени инструменти за наставнике и утврђен је временски период у току кога би требало сакупити попуњене инструменте. По истеку дефинисаног временског периода, истраживач би поново посетио школу и преузео попуњене инструменте.

Приликом прве посете, истраживач је уз помоћ директора школе, договарао термине интервјуа са наставницима српског језика и математике. Наставници су замољени да за разговор припреме примерак теста знања који су применили у раду, заједно са кључем за оцењивање теста. Интервјуи су заказивани, а потом и реализовани у терминима који су наставницима били најповољнији. У просеку, интервју је трајао 30 минута. Наставници су били спремни да уступе истраживачу оригиналан примерак теста или да обезбеде увид у тест. Део анализе методолошких елемената теста остварен је уз помоћ наставника. Реч је о утврђивању нивоа знања који се задатком испитује и процени тежине задатка. По завршетку интервјуа, истраживач је самостално реализовао други део планиране методолошке анализе, који је трајао од 15 до 90 минута у зависности од дужине теста и сложености захтева у тесту.

**АНАЛИЗА И ИНТЕРПРЕТАЦИЈА РЕЗУЛТАТА
ИСТРАЖИВАЊА**

МОГУЋНОСТИ ПРИМЕНЕ ТЕСТА ЗНАЊА У НАСТАВИ

Предности и ограничења теста знања у односу на друге облике праћења и вредновања рада ученика

Тестови знања су један од облика праћења и вредновања рада ученика које наставници примењују у свом раду. На плану теоријских разматрања, тестови знања су најчешће упоређивани са усменим испитивањима, при чему су истицане вишеструке предности тестова (Ebel, 1979). Последњих деценија актуелне су компарације класичних и електронских тестова у којима је предност обично на страни електронских тестова. Са методолошког становишта, тестови се не могу поистоветити са писаним вежбама које се такође користе у циљу испитивања знања ученика, иако су им одређене карактеристике заједничке (истоветност захтева који се постављају пред ученике и економичност). У односу на писане вежбе, конструкција тестова знања је прецизније дефинисана. Такође, тестови најчешће садрже већи број задатака, разноврснијих према форми и тежини.

У овом раду, предности и ограничења тестова знања сагледани су поређењем тестова знања са усменим испитивањем, писаним вежбама и електронским тестовима знања. Претпоставили смо да је сагледавање добрих и лоших страна примене тестова одређено различитим чиниоцима, што је водило томе да критеријуми компарације буду унапред дефинисани. Поређење је остварено с обзиром на 9 критеријума: 1) обухват градива испитивањем, 2) сложеност процеса припреме испитивања, 3) сложеност реализације испитивања, 4) сложеност оцењивања ученичких одговора, 5) истоветност захтева који се постављају пред ученике, 6) истоветност критеријума оцењивања, 7) ефикасност поступка испитивања са становишта броја ученика који се поступком може обухватити, 8) могућност понављања поступка испитивања, 9) могућност „варања“ приликом одговарања.

У односу на 9 задатих критеријума, више од две трећине наставника као предности тестова у односу на усмено испитивање наводи већу економичност поступка тестирања са становишта броја ученика који се тестирањем може обухватити (75,1%), истоветност критеријума оцењивања (67,7%) и истоветност захтева који се постављају пред ученике (61,3%). Мање од половине наставника (40,1%) обим градива који се тестом може обухватити сматра предношћу тестова знања у односу на усмено испитивање. Око трећине испитаника (32,0%) предношћу тестова означава специфичност реализације на часу, односно специфичност поступка оцењивања (26,8%). Мањи број наставника (13,8%) специфичност припреме теста знања сматра њиховом предношћу у односу на усмено испитивање. Тек 12,4% наставника могућност да се поступак провере знања понови више пута са истим питањима сматра предношћу тестова знања (табела 18)

Табела 18. Предности тестова знања у односу на усмено испитивање

Критеријум поређења	да		не	
	f	%	f	%
истоветност захтева	222	61,3	140	38,7
истоветност критеријума оцењивања	245	67,7	117	32,3
припрема поступка испитивања	50	13,8	312	86,2
реализација поступка испитивања	116	32,0	246	68,0
оцењивање одговора	97	26,8	265	73,2
обухват градива испитивањем	145	40,1	217	59,9
обухват ученика испитивањем	272	75,1	90	24,9
поновљивост поступка испитивања	45	12,4	317	87,6
могућност ученика да „варају“	10	2,8	352	97,2

N = 362

Само мали број наставника (2,8%) сматра да је умањена могућност „варања“ предност тестова знања у односу на усмено испитивање.

Резултати хи-квадрат теста показују да међу одговорима наставника у односу на сваки наведени критеријум постоји статистички значајна разлика на нивоу 0,01 (табела 19).

Табела 19. Поређење тестова знања и усменог испитивања

Критеријум поређења	истоветност захтева који се постављају ученицима	истоветност критеријума оцењивања	припрема поступка провере	реализација поступка провере	оцењивање одговора	обим градива који се поступком може обухватити	број ученика који се поступком може обухватити	поновљивост поступка провере	могућност ученика да „варају“
χ^2	18.575	45.260	189.624	46.685	77.967	14.320	91.503	204.376	323.105
df	1	1	1	1	1	1	1	1	1
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Када је реч о поређењу тестова и писаних вежби, подаци добијени испитивањем представљени су табелом 20.

Табела 20. Предности тестова знања у односу на писане вежбе

Критеријум поређења	да		не		укупно	
	f	%	f	%	f	%
истоветност захтева	158	43,6	204	56,4	362	100,0
истоветност критеријума оцењивања	199	55,0	163	45,0	362	100,0
припрема поступка испитивања	54	14,9	308	85,1	362	100,0
реализација поступка испитивања	101	27,9	261	72,1	362	100,0
оцењивање одговора	177	48,9	185	51,1	362	100,0
обухват градива испитивањем	140	38,7	222	61,3	362	100,0
обухват ученика испитивањем	131	36,2	231	63,8	362	100,0
поновљивост поступка испитивања	46	12,7	316	87,3	362	100,0
могућност ученика да „варају“	15	4,1	347	95,9	362	100,0

Укупно 55,0% наставника сматра да је истоветан критеријум оцењивања предност тестова знања у односу на писане вежбе. Специфичност оцењивања ученичких одговора, 48,9% наставника види као предност тестова знања у односу на писане вежбе. Резултати хи-квадрат теста показују да међу одговорима испитаника у односу на наведене критеријуме не постоји статистички значајна разлика.

Нешто мање од половине испитаника (43,6%) истоветност захтева који се постављају пред ученике означава предношћу тестова у односу на писане вежбе. Обухват градива предношћу тестова означава 38,7% испитаника, а обухват ученика 36,2%. Предност тестова у домену реализације поступка испитивања сагледава 27,9% испитаника, у домену припреме поступка испитивања 14,9%, а у домену поновљивости поступка провере знања 12,7%. Разматрајући критеријум „ученици не могу да варају“, тек 4,1 % сматра да тестови у том домену имају предност у односу на писане вежбе. Резултати хи-квадрат теста показују да међу одговорима наставника у односу на наведене критеријум постоји статистички значајна разлика (табела 21).

Табела 21. Поређење писаних вежби и тестова знања

Критеријум поређења	истоветност захтева који се постављају ученицима	истоветност критеријума оцењивања	припрема поступка провере	реализација поступка провере	оцењивање одговора	обим градива који се поступком може обухватити	број ученика који се поступком може обухватити	поновљивост поступка провере	могућност ученика да “варају”
χ^2	5.845	3.580	178.221	70.718	0.177	18.575	27.624	201.381	304.486
df	1	1	1	1	1	1	1	1	1
p	0.016	0.058	0.000	0.000	0.674	0.000	0.000	0.000	0.000

Када је реч о електронским тестовима знања, приближно половина испитаника (48,9%) не користи ову врсту тестова знања, тако да је поређење тестова „папир и оловка“ са електронским тестовима остварено на нивоу половине узорка (51,1%). Значајно је напоменути да су поједини наставници у оквиру коментара на крају инструмента назначили да, заправо, нису сигурни шта су електронски тестови знања, па су се определили за одговор да их не користе. Овакви коментари упућују на извесно ограничење самог инструмента који су наставници попуњавали. Наиме, прецизнији подаци би се могли добити да је било назначено шта су електронски тестови знања или да је

постојао још један понуђени одговор, на пример: „Нису ми познати електронски тестови знања“.

Анализом одговора наставника који су поредили две наведене врсте тестова знања, установљено је да мали број наставника указује на предности класичних тестова над електронским тестовима (табела 22).

Табела 22. Предности класичних тестова у односу на електронске тестове

Критеријум поређења	да		не	
	f	%	f	%
истоветност захтева	62	17,1	300	82,9
истоветност критеријума оцењивања	70	19,3	292	80,7
припрема поступка испитивања	42	11,6	320	88,4
реализација поступка испитивања	81	22,4	281	77,6
оцењивање одговора	54	14,9	308	85,1
обухват градива испитивањем	48	13,3	314	86,7
обухват ученика испитивањем	64	17,7	298	82,3
поновљивост поступка испитивања	24	6,6	338	93,4
могућност ученика да „варају“	24	6,6	338	93,4

N = 362

Резултати хи-квадрат теста показују да међу одговорима наставника у односу на наведене критеријуме постоји статистички значајна разлика на нивоу 0,01 (табела 23).

Табела 23. Поређење класичних и електронских тестова знања

Критеријум поређења	истоветност захтева који се постављају ученицима	истоветност критеријума оцењивања	припрема поступка провере	реализација поступка провере	оцењивање одговора	обим градива који се поступком може обухватити	број ученика који се поступком може обухватити	поновљивост поступка провере	могућност ученика да „варају“
χ^2	156.475	136.144	213.492	110.497	178.221	195.459	151.260	272.365	272.365
df	1	1	1	1	1	1	1	1	1
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Наставници су имали могућност да укажу на недостатке тестова знања у односу на усмено испитивање, писане вежбе и електронске тестове знања. Питање је било отвореног типа, и захтевало је да наставници остваре поређење у све три околности, што се показало као сложен захтев. Значајан број наставника није пружио одговор на наведено питање. Прецизније, недостатке тестова знања у односу на усмено испитивање сагледало је 70,4% наставника, недостатке тестова знања у односу на писане вежбе 51,4% наставника, а недостатке тестова знања у односу на електронске тестове сагледало је тек 20,2% наставника.

Недостаци тестова знања у односу на усмено испитивање на које су наставници указали груписани су у 6 категорија (табела 24).

Табела 24. Поређење тестова знања и усменог испитивања

Недостаци тестова знања	f	%
не може се помоћи ученику	70	31,8
одговор је мање квалитетан	64	29,1
неки ученици су мање успешни на тесту	43	19,6
ученици могу да „варају“	27	12,3
нема индивидуализације захтева	12	5,5
припрема и оцењивање су сложени	4	1,8
укупно	220	100,0

Приказани подаци показују да се као три најчешћа недостатка тестова знања у односу на усмено испитивање издвајају немогућност да се ученику помогне приликом одговарања (31,8%), мањи квалитет добијеног одговора (29,1%) и мања успешност појединих ученика у решавању теста (19,6%).

Категорију „мањи квалитет добијеног одговора“ најчешће истичу наставници друштвено-језичке групе предмета. Категорију можемо илустровати следећим примерима: „У усменом одговору можемо проценити (...) контекстуалну употребу вокабулара, способност изражавања на страном језику, владање граматичким конструкцијама, док тестови знања обично садрже јако профилисана питања“; „На тесту не видимо да ли ученик разуме градиво“; „Ученици учећи за тест не повезују градиво, већ само меморишу чињенице. Путем теста не

могу да дају критичко мишљење и своју процену о неком догађају“; „Нема аналитичког мишљења“; „Усменим излагањем добијамо шире одговоре и образложења одговора“.

Категорију „мања успешност појединих ученика у решавању теста“ можемо илустровати следећим примерима: „Постоје ученици који мање знања показују на тесту него приликом усменог одговора“; „На тесту, нека деца имају мањак концентрације“; „Једноставно, неки ученици не умеју да се изразе кратко и прецизно као што се тражи у тесту“; „Слабији ученици на тесту остварују још слабије резултате“.

Мали број испитаника (9,4%) који су одговорили на дато потпитање понудило је одговор који комбинује неке од претходно размотрених одговора. Два одговора указују на то да тестови знања немају недостатака у односу на усмено испитивање.

Одговори наставника на други део питања који се односио на недостатке тестова знања у односу на писане вежбе, груписани су у 7 категорија (табела 25).

Табела 25. Поређење тестова знања и писаних вежби

Недостаци тестова знања	f	%
одговор је мање квалитетан	94	65,3
ученици могу да „варају“	19	13,2
неки ученици су мање успешни на тесту	11	7,6
припрема је сложена	11	7,6
нема индивидуализације захтева	5	3,5
тест је дужи	3	2,1
не може се више пута поновити	1	0,7
укупно	144	100,0

Убедљиво најчешће истакнут недостатак тестова знања у односу на писане вежбе је мањи квалитет ученичког одговора (65,3%). Остале категорије одговора су значајно мање заступљене.

Категорију „мањи квалитет добијеног одговора“ у одговору на дато потпитање, можемо илустровати следећим примерима: „Тестови знања не мере богатство језика“; „На тесту се не уочава писмена зрелост“; „На тесту ученик не

изражава мисли у целим реченицама“; „Ученик на писменој вежби може да повеже узрок и последицу“; „Писани састави показују комплексно познавање језика и креативност“; „Писмене вежбе омогућавају лакше изражавање и ширу примену наученог градива“.

Мали број испитаника (3,8%) који су одговорили на дато потпитање понудило је одговор који комбинује неке од претходно размотрених одговора. Овога пута, нешто више наставника (7,5%) указало је на то да тестови знања немају недостатака у односу на писане вежбе.

Одговори наставника на трећи део питања који се односио на недостатке тестова знања у односу на електронске тестове, груписани су у 7 категорија (табела 26)

Табела 26. Поређење класичних и електронских тестова знања

Недостаци класичних тестова знања	f	%
оцењивање одговора је сложеније	15	30,0
мање су економични	10	20,0
мање мотивишу ученике на даље учење	8	16,0
више времена је потребно да се да повратна информација	7	14,0
поступак припреме је сложенији	4	8,0
мање је флексибилно време решавања	3	6,0
мањи је обухват градива	3	6,0
укупно	50	100,0

Као три најчешћа недостатка тестова знања у односу на електронске тестове издвајају се: теже оцењивање ученичких одговора (30,0%), мања економичност тестова „папир и оловка“ (20,0%) и већа мотивисаност ученика за учење након решавања електронских тестова знања (16,0). Међу испитаницима који су одговорили на дато потпитање, изврстан број (13,7%) ипак наглашава да „тест знања нема недостатака у односу на електронски тест“.

Захтев за указивањем на недостатке тестова знања у односу на друге начине испитивања ученика, показао се тешким за један број испитаника. На

такав закључак упућује чињеница да су наставници наводили недостатке усменог испитивања, писаних вежби, односно електронских тестова у односу на класичне тестове уместо да учине обратно. Приликом поређења усменог испитивања и тестова знања, погрешно тумачење захтева регистровано је код 3,5% испитаника. Приликом поређења писаних вежби и тестова знања, погрешно тумачење захтева регистровано је код 11,3% испитаника. Приликом поређења електронских и класичних тестова знања, погрешно тумачење захтева регистровано је код 17,8% испитаника.

* * *

Највећи број наставника предности тестова знања у односу на усмено испитивање сагледава у доменима веће економичности поступка тестирања са становишта броја ученика који се тестирањем може обухватити, истоветности критеријума оцењивања и истоветности захтева који се постављају пред ученике. Добијени налаз обухвата неке од најчешће сагледаваних предности тестова знања у односу на усмено испитивање (Хавелка, 2000; Бјекић и Голубовић, 2003). Мање од половине наставника обим градива који се тестом може обухватити сматра предношћу тестова знања у односу на усмено испитивање. Могло би се очекивати да је сагледавање наведене предности тестова више заступљено, сходно резултатима ранијих испитивања (Хавелка, 2000; Бјекић и Голубовић, 2003). Могуће образложење је да поједини наставници под тестом подразумевају инструмент који је сличан контролној вежби и који обично садржи мањи број задатака (на пример, 5 задатака). Тест који би на такав начин био конципиран заиста не би обухватао значајно више градива него усмено испитивање.

Налази да мање од трећине наставника предношћу тестова означава сложеност припреме теста, реализације и оцењивања теста, не представљају изненађење. Наиме, припрема теста подразумева детаљну разраду нацрта и одабирање задатака који према садржају и форми одговарају дефинисаном

нацрту. За разлику од припреме теста, усменом испитивању обично не претходи израда нацрта. Реализација тестирања подразумева одговарајућу стандардизацију услова (од просторно-временских до понашања свих учесника у ситуацији тестирања) која није наглашена у ситуацији усменог испитивања. Оцењивање теста, такође се остварује према унапред дефинисаним правилима: претпоставља постојање кључа који се мора изградити пре реализације тестирања и чија примена је обавезујућа за наставника. Такав приступ оцењивању ученичких одговора није карактеристичан за усмено испитивање.

Мањи број наставника могућност да се поступак провере знања понови више пута са истим питањима сматра предношћу тестова знања у односу на усмено испитивање. Добијени налаз је у супротности са одређеним теоријским становиштима (Хавелка, 2000; Бјекић и Голубовић, 2003). Налаз се и овог пута може објаснити уколико се има у виду често поистовећивање теста са контролном вежбом. У ситуацији када се тест састоји само до неколико задатака, ученици могу лако да запамте задатке и да о томе информишу ученике који ће касније бити тестирани. Из тог разлога, разумљиво је да се наставници неће лако одредити да поново примене истоветан тест. То је посебно важно уколико се задацима испитује ниво познавања чињеница.

Најмањи број наставника могућност ученика да „варају“ на тесту означава предношћу тестова над усменим испитивањем. Добијени налаз се може сматрати очекиваним. Иако се „варање“ ученика може појавити и у ситуацији усменог испитивања, оно се успешније може контролисати у тој ситуацији.

Иако се објективност оцењивања означава предношћу тестова знања над осталим поступцима испитивања у школској пракси (Хавелка, 2000; Бјекић и Голубовић, 2003; *Guide to Assessment*, 2015), резултати показују да је наставницима подједнако тешко оцењивање писаних вежби и тестова знања. Такав налаз се може објаснити ако се има у виду чињеница да уобичајена форма писаних вежби обухвата неколико питања есејског типа. Бројна испитивања су показала да је вредновање одговора на таква питања посебно тешко за наставнике (Buser, 1996; Porham, 2003). Питањима есејског типа у писаним вежбама, најсличнији су задаци испуњавања налога у тесту, мада

задаци испуњавања налога тек повремено захтевају да одговор испитаника добије структуру есеја (одговор најчешће обухвата неколико реченица). Адекватан дизајн теста знања подразумева да су задаци испуњавања налога једна од 7 различитих форми задатака које могу бити заступљене у тесту. Самим тим, тестови знања у односу на писане вежбе садрже већи број задатака које је потребно оценити, али писане вежбе садрже више питања чије одговоре је „теже“ оценити. Могуће је да отуда потиче и приближна уједначеност у одговорима наставника на претходно питање.

С обзиром на сложену израду нацрта теста знања, очекивано, припрему поступка тестирања тек мали број наставника сагледава као предност тестова над припремом писаних вежби. Очекивало би се да обухват градива и поновљивости поступка провере знања буду сагледани као убедљивија предност тестова знања, међутим резултати то не потврђују.

Када је реч о компарацији класичних и електронских тестова знања, установљено је да тек мали број наставника истиче предности класичних тестова над електронским тестовима. Добијени налаз има подршку и у савременој литератури о тестовима знања (He & Tumm, 2005). Ипак, о наведеној компарацији известила је половина од укупног броја испитаника, што указује на недовољно познавање карактеристика електронских тестова на нивоу узорка.

Резултати показују да највише наставника као недостатке тестова знања у односу на усмено испитивање издваја немогућност да се ученику помогне приликом одговарања, мањи квалитет добијеног одговора и мању успешност појединих ученика у решавању теста. Добијени налази су у сагласности са раним критикама тестова знања (Бјекић и Голубовић, 2003). Ипак, критика да се ученику не може помоћи током решавања теста, више се не сматра актуелном, јер превиђа могућност васпитног деловања наставника након примене теста (Гргин, 2001). Такође, квалитет одговора који се добија на тесту не сматра је иманентним недостатком теста, већ се доводи у везу са његовом конструкцијом. Уколико ученици на тесту показују искључиво владање чињеницама, али не и сложеније облике мишљења, извор проблема се може

потражити у захтевима који се пред ученике постављају, а не у ограничењу теста као инструмента. Мања успешност појединих ученика у решавању теста, такође, могла би се довести у везу са искуством ученика у решавању тестова, односно неразвијеношћу мисаоних вештина које су неопходне за успешно решавање задатака (Дубљанин, 2010).

Највише наставника као недостатак тестова знања у односу на писане вежбе истиче мањи квалитет ученичког одговора. Решавање задатака у тесту наставници доводе у везу са давањем кратких одговора који не остављају простора ученику да аргументује свој одговор, исказе другачији став или креативност. Као што је претходно наглашено, дати приступ не сагледава сав потенцијал тестова знања као инструмената којима се могу испитивати различити нивои мишљења и који могу садржати задатке различите форме.

Највише наставника као недостатке тестова знања у односу на електронске тестове издвајају теже оцењивање ученичких одговора, мању економичност класичних тестова и већу мотивисаност ученика за учење након решавања електронских тестова знања. Иако би се предности електронских тестова могле потражити и осталим доменима компарације, може се закључити да су уочени недостаци класичних тестова у складу са резултатима других испитивања (Фајгел, 2003; He & Tumm, 2005; Маравић-Чисар, 2011).

Припремљеност наставника за израду теста знања

Уколико се наставници одлуче да примене тестова знања у свом раду, поставља се питање да ли су и у којој мери припремљени за њихову израду. Могло би се очекивати да су наставници током иницијалног професионалног образовања стекли знања и вештине потребне за израду тестова. Уколико је таква припрема била организована, разлике се могу очекивати у погледу садржаја и организације саме припреме, и последично, у нивоу припреме коју су наставници стекли. Међутим, ако се има у виду чињеница да поједини наставници пре запослења у школи, нису похађали факултете који обухватају

методичку припрему, поставља се питање да ли су недостатак припреме могли да надокнаде похађањем одговарајућих облика стручног усавршавања? О којим облицима је реч? Каква је позиција семинара стручног усавршавања у односу на друге облике? С обзиром на различите садржаје и организацију облика стручног усавршавања и овог пута се могу очекивати разлике у нивоу припреме коју су наставници стекли.

Процена припремљености наставника за израду тестова знања сагледана је на три нивоа: а) глобално, б) са становишта потребних знања и в) са становишта потребних вештина. Глобална процена припремљености за израду тестова знања на нивоу узорка показала се изузетно повољном. Просечна процена, на скали теоријског распона од $M = 1,00$ до $M = 5,00$, износила је $M = 4,11$. Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика на нивоу $0,01$ ($\chi^2 = 342,97$; $df = 4$; $p = 0,00$; Прилог 3, табела 1).

Процена припремљености за израду тестова знања са становишта потребних знања и вештина, такође се показала повољном. Просечна процена припремљености са становишта потребних знања, на скали теоријског распона од $M = 1,00$ до $M = 5,00$, износила је $M = 4,19$ ($SD = 0,82$). Просечна процена припремљености са становишта потребних вештина, на скали теоријског распона од $M = 1,00$ до $M = 5,00$, износила је $M = 4,02$ ($SD = 0,82$). Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика на нивоу $0,01$ и када је реч о потребним знањима ($\chi^2 = 333,80$; $df = 4$; $p = 0,00$) и када је реч о потребним вештинама ($\chi^2 = 316,94$; $df = 4$; $p = 0,00$) (табела 27).

Табела 27. Припремљеност наставника за израду тестова знања са становишта потребних знања и вештина

Процена припремљености са становишта:	потребних знања		потребних вештина	
	f	%	f	%
потпуно сам припремљен/а	138	38,4	98	27,5
доста сам припремљен/а	171	47,6	188	52,7
осредње сам припремљен/а	37	10,3	55	15,4
мало сам припремљен/а	8	2,2	11	3,1
нисам припремљен/а	5	1,4	5	1,4
укупно	359	100,0	357	100,0

Дистрибуције одговора потврђују да високу процену укупне припремљености за израду тестова знања прати и висока процена припремљености са становишта потребних знања и вештина. Оба пута, најчешће бирана категорија одговора је „доста сам припремљен/а“, а већ следећа по учесталости је категорија „потпуно сам припремљен/а“.

Допринос иницијалног професионалног образовања припреми наставника за израду теста знања

С обзиром на релативно „млад“ узорак испитаника, очекивало би се да високу процену припремљености за израду тестова знања прати исто тако високо вредновање припреме на студијама. У наставку, приказаћемо резултате којима се стиче јаснији увид у припрему наставника током студијама.

Око половине испитаника (51,4%) изјавило је да су се током студија припремали за израду тестова знања. Уколико је припрема за израду тестова знања била организована током иницијалног професионалног образовања, најчешће се спроводила у оквиру обавезног студијског предмета (62,9%). Значајно ређе, организована је у оквиру изборног студијског предмета (5,9%) или у оквиру и обавезног и изборног предмета (2,2%). Немали број наставника (28,0%) није могао да се сети какав је био статус предмета у оквиру кога је

припрема била организована. Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика и када је реч о похађању припреме током студија ($\chi^2 = 5,08$; $df = 1$; $p = 0.02$; Прилог 3, табела 2) и када је реч о статусу студијског предмета у оквиру кога је припрема организована ($\chi^2 = 182,23$; $df = 2$; $p = 0.00$; Прилог 3, табела 3).

Највећи број наставника (53,2%) који су се припремали за израду тестова знања током студија, извештава да су у оквиру припреме разматрана питања појма и врста тестова знања. Нешто мањи број наставника (45,7%) потврђује да је припрема обухватала питања могућности примене тестова знања у настави. Следећи по учесталости су садржаји који се односе на врсте и карактеристике задатака у тесту знања (36,0%), а тек на четвртом месту налазе се садржаји који се тичу процеса израде теста знања (30,6%) (табела 28).

Табела 28. Садржаји обухваћени припремом за израду тестова знања током студија

Садржаји	f	%
појам и врсте тестова знања	99	53,2
могућности примене тестова знања у настави	85	45,7
врсте и карактеристике задатака у тесту знања	67	36,0
процес израде теста знања	57	30,6
нешто друго	1	0,5
		N = 186

Око половине наставника (48,4%) процењује да је припрема током студија више допринела усвајању теоријских знања из области тестирања него усклађивању теоријских и практичних знања (34,9%), односно развијању потребних вештина (34,4%) (табела 29).

Табела 29. Допринос припреме током студија актуелној припремљености за израду тестова знања

Аспекти доприноса	f	%
усвајање теоријских знања из области тестирања	90	48,4
усклађивање теоријских и практичних знања	65	34,9
развијање вештина потребних за израду тестова знања	64	34,4
нешто друго	1	0,5
		N = 186

И коначно, наставници су имали могућност да процене укупан допринос студија њиховој актуелној припремљености за израду тестова знања. Просечна процена доприноса припреме током студија, на скали теоријског распона од $M = 1,00$ до $M = 5,00$, износила је $M = 2,37$ ($SD = 1,31$). Одговори наставника указују на релативно лошу оцену доприноса студија припремљености за израду тестова знања – више од половине наставника (55,3%) допринос припреме током студија оцењује као мали или непостојећи. Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика на нивоу 0,01 ($\chi^2 = 83,85$; $df = 4$; $p = 0.00$). Подаци говоре у прилог томе да студије ипак нису разлог високог самопоуздања наставника у домену израде тестова (табела 30).

Табела 30. Допринос припреме током студија актуелној припремљености за израду тестова знања

Процена доприноса	f	%
веома много	24	6,9
много	54	15,6
осредње	77	22,2
мало	64	18,4
нимало	128	36,9
укупно	347	100,0

Нешто више од половине од укупног броја испитаника (56,1%) дало је препоруке у вези са унапређивањем припреме за израду тестова знања током студија. Највише препорука односило се на измену начина рада у оквиру неког од постојећих студијских курсева који предвиђа разматрање тематике тестова знања и на увођење одговарајућег студијског предмета или садржаја који би се односио на тестове знања (табела 31).

Табела 31. Унапређивање припреме за израду тестова знања током студија

Предлози	f	%
измена начина рада у оквиру постојећег предмета	66	41,3
увођење обавезног студијског предмета	48	30,0
измена садржаја наставе постојећег предмета	35	21,9
увођење изборног предмета	6	3,8
укључивање студената у наставни процес у школи	5	3,1
укупно	160	100,0

Наставници наглашавају да је потребно преиспитати наставне методе на којима је припрема утемељена, мада се изванредан број препорука тиче и измене садржаја у оквиру припреме. Наставници желе да сугеришу да је потребно акценат са теоријске припреме померити у правцу практичне примене стеченог знања. На пример: *„Методика наставе историје би морала да садржи више практичних, а мање теоријских знања“*; *„Методика наставе француског језика није имала практични део, само теорија, теорија, теорија...“*; *„Да се избаки теорија, а убаци пракса“*; *„Нисмо учили о процесу израде теста, него само о могућностима примене“*; *„Увели бих више часова посвећених изради тестова знања са тачно разрађеном методологијом, као и часове вежби на којима би се коментарисали типови тестова“*; *„Потребно је да нас науче како се прави тест и како се оцењује“*; *„Увели бих праксу израде, примене и анализе резултата тестова знања“*; *„Требало би увести процес израде тестова и у вези са тим начине градирања тежине појединачних задатака“*; *„Важно је научити студенте како да саставе валидан тест“*.

Препоруке једног броја наставника крећу се у правцу увођења одговарајућег студијског предмета, односно увођења релевантног садржаја у програм већ постојећег студијског предмета (на пример, методике наставе одговарајућег предмета). Мањи број испитаника је оценио да: у вези са припремом за израду тестова знања током студија није потребно ништа мењати (12,3%); да се већ доста тога променило у односу на време њихових студија (4,4%); да им није познато да ли је у међувремену нешто промењено

(2,5%); односно да нема потребе за организовањем припреме за израду тестова знања током студија (2,0%).

Допринос стручног усавршавања припреми наставника за израду теста знања

С обзиром на претходно размотрене резултате, остаје нам да утврдимо да ли је висока процена властите припремљености за израду тестова знања утемељена на похађању обука у домену стручног усавршавања наставника. Уз консултације Правилника о сталном стручном усавршавању и стицању знања наставника, васпитача и стручних сарадника (2012) утврђена је листа облика сталног стручног усавршавања који су обухваћени даљом анализом¹. Облици који су преузети из Правилника били су: семинари стручног усавршавања, учешће на стручном скупу, летњој/зимској школи, боравак на стручном/студијском путовању. Листа је допуњена категоријама за чију релевантност налазимо потврду у домаћој литератури (Боговац и сар, 1972; Банђур и Поткоњак, 1999). Реч је о следећим категоријама: „консултовање са искуснијим колегама“, „сарадња са стручним сарадником у школи“ и „читање релевантне методолошке литературе“. Наставници су имали задатак да међу понуђеним облицима означе облике стручног усавршавања које су похађали и да, према потреби, допишу облик/облике који нису наведени (табела 32).

Табела 32. Стручно усавршавање наставника у домену израде тестова знања

Облици стручног усавршавања	f	%
консултовање са искуснијим колегама	294	81,2
читање релевантне методолошке литературе	270	74,6
сарадња са стручним сарадником у школи	198	54,7
семинари стручног усавршавања	194	53,6
учешће на стручном скупу	103	28,5
учешће у летњој/зимској школи	21	5,8
боравак на стручном/студијском путовању	12	3,3
неки други облик	12	3,3

N = 362

¹ У наставку текста користићемо скраћени назив „Правилник о стручном усавршавању“.

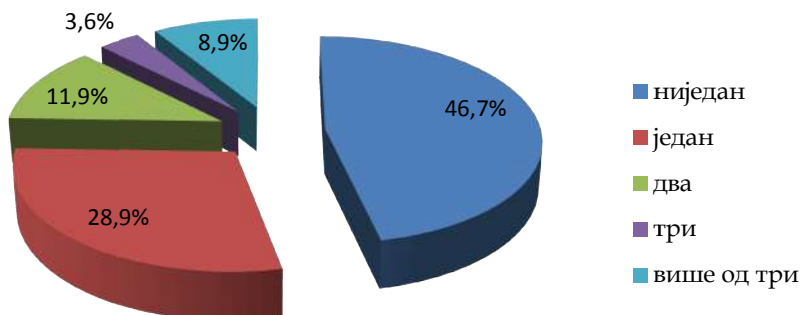
Како резултати показују, наставници обухваћени испитивањем највише сазнања о процесу израде тестова знања стичу кроз консултовање са искуснијим колега (81,2%). Према учесталости следи читање релевантне методолошке литературе (74,6%), сарадња са стручним сарадником у школи (54,7%) и семинари стручног усавршавања (53,6%). Учешћа на стручним скуповима су ређе заступљена, а посебно су ретки облици стручног усавршавања који захтевају издвајање већих финансијских средстава (летње/зимске школе и стручна/студијска путовања). Категорији облика који нису били понуђени у инструменту, наставници су прикључили следеће: „учешће у изради стандарда знања“, „анализа готових тестова и тестова за такмичење“, „анализа примера тестова колега из земље и иностранства који се могу пронаћи на интернету (форуми, блогови)“, „организација трибине: каква питања воде до знања“.

Мањи број наставника (6,9%) изјавио је да нема искуство стручног усавршавања у домену израде тестова знања.

Уколико из табеле 32. издвојимо само облике усавршавања који су препознати Правилником о стручном усавршавању (2012), приметимо да су међу њима најзаступљенији семинари. Добијени налаз је у складу са очекивањима. Пре свега, похађање семинара је обавезно. Такође, семинари су облик усавршавања који се може организовати у школској установи и захтева мања материјална улагања од учешћа на стручним скуповима, летњим школама, студијским путовањима. Управо због веће заступљености семинара од осталих облика стручних усавршавања, даљом анализом утврђене су карактеристике семинара које су наставници похађали у домену израде тестова знања: број семинара које су наставници похађали, садржај обуке, процена ефеката обуке, иницијатива за похађање семинара и финансирање обуке.

Највећи број наставника (46,7%) до тренутка испитивања није похађао ниједан семинар посвећен изради тестова знања. Један семинар у референтној области похађало је нешто мање од трећине испитаника (28,9%). Свега 43 наставника (11,9%) похађала су два семинара, а нешто мањи број испитаника (8,9%) похађао је три или више семинара (слика 6). Резултати хи-квадрат теста

показују да међу одговорима наставника постоји статистички значајна разлика на нивоу 0,01 ($\chi^2 = 224,47$; $df = 4$; $p = 0.00$).



Слика 6. Број семинара које је наставник похађао у домену израде тестова знања

У поређењу са припремом током студија, садржаји семинара битно су другачији. Најкрупнија промена коју доносе семинари стручног усавршавања је акцентовање процеса израде теста знања (развија садржаја теста, састављање задатака, оцењивање теста и слично). Уколико је припрема током студија постојала, наставници су извештавали о доминантно теоријским садржајима припреме и слабо заступљеној практичној примени знања. Иако се и на семинарима разматрају питања као што су појам и врсте тестова, ипак се процесу израде теста посвећује већа пажња (табела 33).

Табела 33. Садржаји обухваћени семинарима у домену израде тестова знања

Садржаји	f	%
процес израде теста знања	118	61,5
могућности примене тестова знања у настави	112	58,3
појам и врсте тестова знања	91	47,4
врсте и карактеристике задатака у тесту знања	82	42,7
нешто друго	10	5,2
		N = 192

Размотримо како су наставници проценили ефекте семинара које су похађали (табела 34).

Табела 34. Допринос семинара припремљености за израду тестова знања

Аспекти доприноса	f	%
усвајање теоријских знања из области тестирања	80	41,9
развијање вештина потребних за израду тестова знања	127	66,5
усклађивање теоријских и практичних знања	105	55,0
нешто друго	6	3,1
		N = 191

И овога пута потврђено је да припрема током студија и семинари стручног усавршавања остварују различите исходе – семинари доминантно развијају вештине потребне за израду тестова знања, док припрема током студија доминантно доприноси усвајању теоријских знања из области тестирања.

Како резултати показују, лична иницијатива наставника за похађањем семинара доминира над иницијативом других (75,9%). Значајно ређе иницијатива долази од стране стручног актива, Завода за унапређивање образовања и васпитања/Министарства просвете, директора школе или стручних сарадника школе (табела 35).

Табела 35. Иницијатива за похађање семинара у домену израде тестова знања

Иницијатор	f	%
лично	145	75,9
стручни актив наставника	27	14,1
Завод/Министарство просвете	25	13,1
директор школе	22	11,5
стручни сарадници школе	18	9,4
неко други	2	1,1
		N = 191

Више од половине наставника (57,6%) потврђује да је школа искључиви финансијер семинара стручног усавршавања у домену израде тестова знања (табела 36). Следећи одговор по учесталости је „неко други“. Анализа

добијених одговора показује да су то издавачи уџбеника (Клет, Бигз, *The English Book*), стране организације (*British Council*, Гете Институт), поједини факултети (нису наведени називи), али се међу одговорима срећу и следећи: „Семинар је био бесплатан“; „Половину трошкова је платила школа, половину сам платио сам“; „Само један је финансирао школа, остале сам платио сам (...)“ и „Не знам ко је платио семинар“. Немали број наставника (17,8%) самостално је финансирао своју обуку, а скоро подједнако је заступљено (17,3%) учешће Министарства просвете у финансирању семинара. Општина и локална заједница тек повремено су имали учешћа у финансирању обука.

Табела 36. Финансирање семинара у домену израде тестова знања

Финансијер	f	%
школа	110	57,6
заинтересовани наставници	34	17,8
Министарство просвете	33	17,3
општина	13	6,8
локална заједница	8	4,2
неко други	37	19,4

N = 191

Нешто више од половине испитаника који су похађали семинаре у домену израде тестова знања, оцењују допринос семинара у распону од осредњег ка значајном (табела 37).

Табела 37. Допринос похађања семинара усавршавању у процесу израде тестова знања

Процена доприноса	f	%
веома много	37	19,3
много	65	33,9
осредње	66	34,4
мало	21	10,9
нимало	3	1,6
укупно	192	100,00

Врло слична дистрибуција одговора добијена је и у вези са проценом доприноса учешћа на стручним скуповима усавршавању у процесу израде тестова знања (табела 38).

Табела 38. Допринос учешћа на стручном скупу усавршавању у процесу израде тестова знања

Процена доприноса	f	%
веома много	13	10,4
много	36	28,8
осредње	61	48,8
мало	12	9,6
нимало	3	2,4
укупно	125	100,0

Допринос сарадње са колегама оцењен је повољније. Око две трећине наставника који имају искуство сарадње са колегама оцењују допринос таквог облика стручног усавршавања у распону од значајног ка веома значајном (табела 39).

Табела 39. Допринос сарадње са колегама усавршавању у процесу израде тестова знања

Процена доприноса	f	%
веома много	104	32,7
много	135	42,5
осредње	56	17,6
мало	21	6,6
нимало	2	0,6
укупно	318	100,0

Нешто више од половине наставника који имају искуство сарадње са стручним сарадником школе оцењују допринос таквог облика стручног усавршавања у распону од осредњег ка значајном (табела 40).

Табела 40. Допринос сарадње са стручним сарадником у савршавању у процесу израде тестова знања

Процена доприноса	f	%
веома много	41	17,4
много	94	39,8
осредње	58	24,6
мало	32	13,6
нимало	11	4,7
укупно	236	100,0

Приближно две трећине наставника који се опредељују за проучавање стручне литературе у домену стручног усавршавања, оценило је њен допринос усавршавању у процесу израде тестова знања у распону од значајног ка веома значајном (табела 41).

Табела 41. Допринос консултовања стручне литературе усавршавању у процесу израде тестова знања

Процена доприноса	f	%
веома много	93	30,8
много	116	38,4
осредње	66	21,9
мало	25	8,3
нимало	2	0,7
укупно	302	100,0

Када се упореде пондерисане средње вредности, може се закључити да се највиша оцена приписује доприносу сарадње са колегама ($M = 4,00$), потом праћењу стручне литературе ($M = 3,90$), семинара ($M = 3,58$), и на крају сарадњи са стручним сарадником ($M = 3,51$) и стручним скуповима ($M = 3,35$).

Укупно 66,3% испитаника одговорило је на питање како би се могло унапредити стручно усавршавање у домену израде тестова знања. Одговори су груписани у сродне категорије и добијена је дистрибуција коју приказујемо у табели 42.

Табела 42. Унапређивање стручног усавршавања у области израде тестова знања

Начини унапређивања	f	%
промена организације и начина рада на семинарима	92	39,1
размена искустава са колегама	21	8,9
више примењивати тестове у раду	10	4,3
већа доступност савремене стручне литературе	6	2,6
облици усавршавања који се ослањају на употребу рачунара и интернета	5	2,2
организовање трибина на нивоу стручних актива	3	1,3
обука у току приправничког стажа	1	0,4
комбинација више наведених облика	97	41,3
укупно	235	100,0

Предлози наставника за унапређивање облика стручног усавршавања најчешће су се односили на комбиновање различитих облика усавршавања (41,3% одговора). Неки од карактеристичних одговора на дато питање су: „Похађањем бесплатних семинара и праћењем стручне литературе“; „Консултације у оквиру актива и праћење нових тестова на интернету“; „Личним ангажовањем у изради тестова и похађањем семинара који су боље осмишљени“; „Семинари и сарадња са стручним сарадником“; „Да се окупе практичари који примењују квалитетне тестове које су сами израдили и да Министарство подржи акредитацију семинара који би они креирали“; „Међусобним договором, консултацијама са колегама уз помоћ стручне литературе“.

Следећи по учесталости је предлог да се промени организација или начин рада на постојећим семинарима (39,1%). На пример: „Похађањем семинара који доприносе да се ускладе теоријска и практична знања“; „На семинару, уз довољно простора за практични део“; „Организовањем семинара у нашем граду који би био бесплатан“; „Похађање семинара који ће водити запослени у школама и који тестове успешно користе“; „Плаћеним семинарима“; „Организовањем обавезног семинара на коме би стручњаци објаснили сваки корак у састављању теста, како теоријски, тако и практично“; „Похађањем семинара са провером стеченог знања“. Сви остали предлози за унапређивање стручног усавршавања у домену израде тестова, добили су мање од по 10% сагласности.

Напоменимо и то да се међу испитаницима који су понудили одговоре на наведено питање, 1,3% наставника изјаснило да не зна на који начин се стручно усавршавање у области израде тестова знања може унапредити, а 0,8% испитаника је оценило да нема потребе за таквом врстом усавршавања.

* * *

Резултати показују да се око половине испитаника није припремало за израду тестова током студија, док је друга половина испитаника незадовољна припремом која је била организована. Недостатке припреме током иницијалног професионалног образовања могуће је потражити у неадекватно осмишљеном садржају или неодговарајућем избору наставних метода. По завршетку студија, наставницима превасходно недостају вештине потребне за израду тестова знања.

До тренутка реализације испитивања, мали број наставника је похађао семинаре посвећене изради тестова знања. Добијени налаз се може објаснити када се има у виду релативно скромна понуда семинара посвећених изради тестова знања у Каталогу програма стручног усавршавања (*Каталог програма сталног стручног усавршавања наставника, васпитача и стручних сарадника за школску 2014/2015 и 2015/2016*). Скромна понуда је у извесном раскораку са потребама наставника за обукама у домену вредновања постигнућа ученика, утврђеним ранијим испитивањима (Кундачина и Стаматовић, 2012).

Док припрема током студија доминантно доприноси усвајању теоријских знања из области тестирања, мишљење наставника је да семинари доминантно развијају вештине потребне за израду тестова знања. Наведени налаз можемо оценити као повољан, како са становишта компензаторног деловања семинара стручног усавршавања, тако и са становишта квалитета семинара. Наиме, испитивања показују да овладавање вештинама примењивим у пракси за наставнике представља једну од кључних карактеристика „добрих“ семинара (Кундачина и Стаматовић, 2012).

Како резултати показују, лична иницијатива наставника за похађањем семинара доминира над иницијативом других: стручног актива, Завода за

унапређивање образовања и васпитања/Министарства просвете, директора школе или стручних сарадника школе. Такође, немали број наставника самостално је финансирао своју обуку. Констатовано стање можемо оценити као повољно, јер показује да наставници препознају своје потребе у домену професионалног усавршавања. То је управо једна од полазних претпоставки успешног планирања стручног усавршавања запослених у образовању (Стаматовић, 2006). Семинари стручног усавршавања јесу препознати као облик стручног усавршавања који може да допринесе унапређивању праксе у домену израде тестова знања, уколико је акценат на стицању практичних знања и вештина и уколико се учине приступачнијим.

Упркос ограниченом доприносу иницијалног професионалног образовања и семинара стручног усавршавања, већина наставника је своју укупну припремљеност за израду тестова знања оценила високом оценом. Објашњење се може потражити у домену других облика стручног усавршавања за које се наставници опредељују. Истраживањем је установљена спремност наставника да се у вези са питањима конструкције тестова знања посаветују са искуснијим колега или да самостално истражују у процесу израде тестова.

Тешкоће наставника у процесу израде теста знања у настави и начини њиховог превазилажења

С обзиром на сложеност и разноврсност фаза конструкције може се очекивати да је процес израде теста праћен одређеним тешкоћама. Поставља се питање, са којим врстама тешкоћа се наставници суочавају у процесу израде теста? Које тешкоће су учесталије од других? На који начин наставници постојеће тешкоће превазилазе?

Око трећине испитаника (33,1%) истиче да нема тешкоће у процесу израде теста. Извештавање преосталих 66,9% наставника о различитим врстама

тешкоћа са којима се суочавају, било је могуће упоредити са фазама израде теста знања које се разматрају у методолошкој литератури (табела 43).

Табела 43. Тешкоће са којима се наставници суочавају приликом састављања тестова знања

Врста тешкоћа	f	%
одређивање тежине задатака	83	33,3
одређивање броја поена које задатак носи	67	26,9
одређивање нивоа знања који задатак мери	61	24,5
одређивање дужине теста	59	23,7
избор адекватне врсте задатака за одређени садржај	53	21,3
превођење поена на тесту у оцене	42	16,9
развијање и чување базе задатака	20	8,0
одређивање редоследа задатака у тесту	16	6,4
развијање садржаја теста	15	6,0
реализовање пробног испитивања	11	4,4
састављање упутства за решавање теста	8	3,2
нешто друго	16	6,4

N = 249

Наставници који извештавају о тешкоћама у процесу израде тестова, најчешће наводе тешкоће у одређивању тежине задатака (33,3%), број поена које задатак носи (26,9%), одређивању нивоа знања који задатак мери (24,5%), одређивању дужине теста (23,7%), избору врсте задатака (21,3%) и превођењу поена на тесту у оцене (16,9%). У мањој мери, заступљене су тешкоће у развијању и чувању базе задатака, одређивању редоследа задатака у тесту, разради садржаја теста, реализовању пробног испитивања и састављању упутства за решавање теста. Такође, мањи број наставника извештава о тешкоћама које не припадају фазама израде теста и које су сврстане у категорију „нешто друго“. На пример: „У свим доменима бих волела да се усавршим“; „Тешко ми је да прилагодим тестове знања ученицима који слабо напредују (у петом разреду не знају таблицу множења, а треба да множе разломке)“; „Конкретно, за напредни ниво немам довољно материјала и литературе“.

У циљу превазилажења тешкоћа у процесу конструкције теста знања наставници ће се обратити најпре својим колегама (67,5%). Према учесталости следи самостално решавање насталих потешкоћа (43,4%) и консултовање

стручне литературе (38,2%). Консултације са педагогом школе налазе се на четвртом месту према заступљености (20,5%), док су консултације са психологом још ређе (8,4%) (табела 44).

Табела 44. Начин решавања тешкоћа у процесу израде тестова знања

Начин решавања тешкоћа	f	%
консултовање са искуснијим колегама	168	67,5
самостално, путем „покушаја и погрешака“	108	43,4
читање релевантне методолошке литературе	95	38,2
сарадња са педагогом школе	51	20,5
сарадња са психологом школе	21	8,4
неки други начин	19	7,6

N = 249

Међу одговорима наставника у вези са тешкоћама у изради тестова, предњаче потешкоће које се непосредније односе на састављање задатака. Добијени налази су очекивани, јер се најсложенији кораци у изради теста знања тичу управо задатака: одређивања нивоа знања који испитују, форме, тежине, броја поена који носе. Такође, налази су у сагласности са резултатима испитивања других аутора (Cooper & Leiter, 1980; Stiggins & Bridgeford, 1985; Rudner & Schafer, 2002; Frey *et al.*, 2005).

Занимљиво је да трећина наставника (33,1%) изјављује да нема тешкоћа у процесу израде тестова знања. У односу на испитивање Богавца и сарадника (1972), где је већина наставника указивала да има потешкоћа, резултат добијен нашим испитивањем је повољнији. Можда би се налаз могао довести у везу са бољом информисаношћу наших испитаника о томе шта су све карактеристике доброг теста знања и бољим овладавањем знањима и вештинама потребним за израду теста знања. У извесној мери, то је и разумљиво, јер у време испитивања Богавца и сарадника нису реализовани завршни испити у форми теста знања и није било различитих евалуативних испитивања који су део професионалног искуства савремених наставника.

Учесталост и диманика коришћења теста знања у настави

Испитивање је омогућило да установимо колико су тестови знања заступљени у школској пракси као средство праћења и вредновања постигнућа ученика и у којим околностима се наставници опредељују за њихову примену. Напоменимо на почетку да највећи број наставника (96,1%) обухваћених испитивањем користи тестове знања, док тек мали број (3,9%) то не чини. Од укупног броја наставника који користе тестове, укупно 94,5% испитаника тестове узрађује самостално, док то не чини 5,5% испитаника. Уколико користе тестове, али се не одлучују да их самостално припреме, наставници извештавају да су им доступни тестови које је неко други припремио. Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика на нивоу 0,01 када је реч о самосталности у изради теста ($\chi^2 = 285.43$; $df = 1$; $p = 0.00$; Прилог 3, табела 4)

Међу наставницима који примењују тестове, већина (38,4%) се одлучује за примену тестова четири пута годишње. Према учесталости, следе категорије „два пута у току школске године“ (24,6%) и „пет и више пута у току школске године“ (22,8%) (табела 45). Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика на нивоу 0,01 када је реч о учесталости израде тестова знања ($\chi^2 = 182.53$; $df = 5$; $p = 0.00$).

Табела 45. Учесталост коришћења тестова знања у току школске године

Учесталост коришћења у току школске године	f	%
једанпут	25	7,2
два пута	85	24,6
три пута	24	6,9
четири пута	133	38,4
пет и више пута	79	22,8
укупно	346	100,00

Близу трећине наставника тестове израђује четири пута годишње. Ипак, неопходно је заузети извештајан опрез у тумачењу добијених података. Наиме, узорак обухваћен истраживањем чине наставници предмета у којима су могућности примене тестова знања релативно велике (на пример: српски језик, математика, биологија, хемија, страни језици). Процењујемо да би подаци били битно другачији да су узорком били обухваћени и наставници неких других предмета као што су предмети вештина или уметности (на пример: ликовно, музичко, физичко).

Даљом анализом, установљено је да највећи број наставника (63%) тестове задаје након обраде веће наставне целине. Према учесталости, следе одговори „на полугодишту“ (33,1%), „на тромесечју“ (23,5%) и „у неком другом тренутку“ („Скоро сваке недеље“; „Након обраде мање наставне целине“; „Често задајем петоминутне тестове да би се проверило знање из претходне лекције и да би се ученици подстакли да редовно уче“; „На почетку школске године“; „На крају школске године“; „Када проценим да је прави тренутак“; „При поправљању оцена“) (15,2%).

Правовремено тестирање често се доводи у везу са остваривањем бољих резултата у учењу (McDaniel, Roediger III & McDermott, 2007). За разлику од резултата нашег истраживања, друга испитивања показују да најповољнији тренутак за задавање теста наступа већ недељу дана након првог сусрета ученика са датом наставном материјом (Ali Salmani Nodoushan, 2010).

Иако би се сходно одређеним истраживачким налазима препоручивало чешће тестирање ученика (Bangert-Drowns, Kulik & Kulik, 1991; Gipps, 2003; Бјекић и Голубовић, 2003; McDaniel, Roediger III & McDermott, 2007), у актуелној пракси, наставници усклађују своје планове са захтевима Правилника о оцењивању ученика у основном образовању и васпитању (у даљем тексту: Правилника о оцењивању), односно са радом осталих наставника у школи. Наиме, Правилником о оцењивању (2013: члан 12) наглашено је да наставник сваку усмену или писану проверу ученичких знања дужу од 15 минута мора најавити најкасније до краја треће недеље полугодишта. У одређивању динамике испитивања, важно је да се поштује одредба према којој је могуће

реализовати највише две провере дуже од 15 минута, у истом одељењу, у једној недељи. Пошто се може догодити да је више наставника заинтересовано да зада тест у истом временском периоду, неопходно је да се индивидуални планови усагласе и дефинише распоред провера знања на нивоу школе. Распоред испитивања утврђује директор школе на предлог одељенског већа (*Правилник о оцењивању ученика у основном образовању и васпитању*, 2013: члан 12). Усмене или писмене провере постигнућа у трајању до 15 минута не морају бити најављене и спроводе се са циљем утврђивања остварености циља часа и савладаности дела градива. Оцена стечена на провери која је краћа од 15 минута не уписује се у дневник (*Правилник о оцењивању ученика у основном образовању и васпитању*, 2013: члан 13). Дакле, уколико није реч о краћим тестовима знања, учесталост тестирања ученика одређена је распоредом провера знања на нивоу школе. У том смислу, наставник нема потпуну слободу да тест планира када процени да је за то прави тренутак, што се може одразити на учесталост тестирања уопште.

Једно од ретких испитивања динамике тестирања код нас, истраживање Богавца и сарадника (1972), показало је да наставници основних школа тестове знања најчешће примењују на тромесечју (36,6%), након обраде веће наставне целине (33,8%) и на полугодишту (33,4%). Мањи број наставника (8,2%) тестове користи само на крају школске године (Богавцац и сар, 1972). Динамика коришћења тестова током школске године регистрована нашим испитивањем, делимично је подударна са резултатима испитивања Богавца и сарадника (1972). Као и динамика тестирања регистрована у испитивању Богавца и сарадника, и динамика тестирања регистрована нашим испитивањем суштински одговара динамици сумативног оцењивања у тумачењу Правилника о оцењивању (2013: члан 3).

Предмет који наставници предају и могућности примене теста знања у настави

Да би се потпуније сагледало мишљење наставника о предностима и ограничења тестова знања у односу на друге облике праћења и вредновања

рада ученика, доприносу иницијалног професионалног образовања и стручног усавршавања припреми наставника за израду тестова знања, тешкоћама са којима се наставници суочавају у процесу израде тестова знања, начинима превазилажења тешкоћа и учесталости и динамици коришћења тестова знања, приступило се испитивању разлика у одговорима наставника с обзиром на предмет који предају.

Резултати показују да су статистички значајне разлике у одговорима наставника друштвено-језичке и природно-математичке групе предмета регистроване само у домену сагледавања појединих предности тестова знања у односу на друге начине праћења и вредновања постигнућа ученика, и домену припремљености за израду тестова знања који се односи на иницијално образовање.

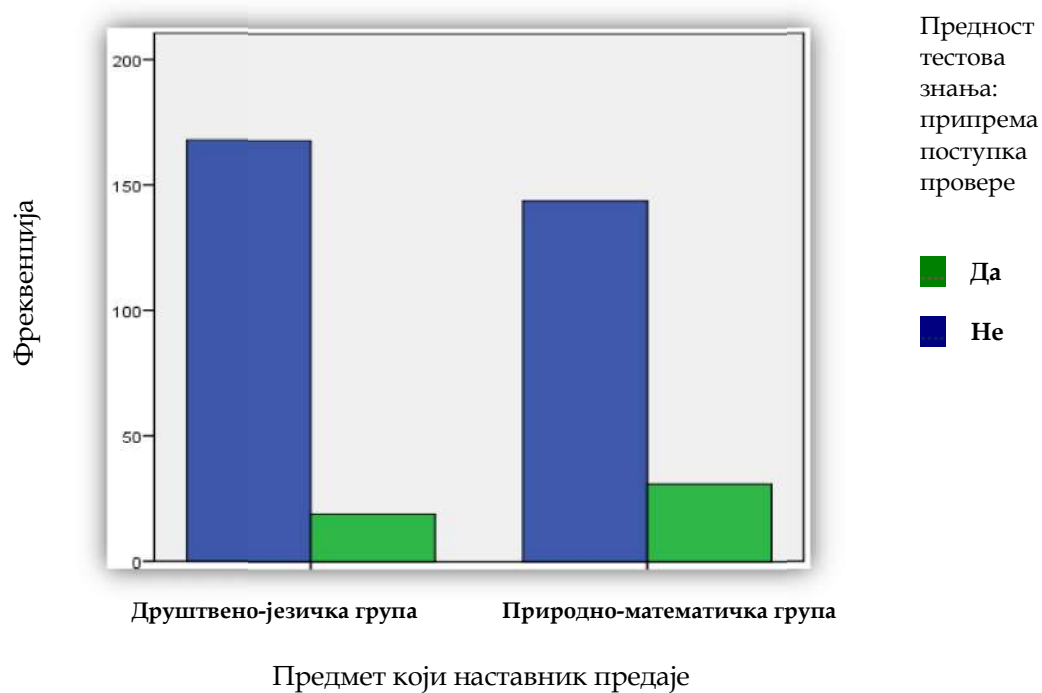
Када је реч о поређењу тестова знања и усменог испитивања, истраживањем је регистрована тенденција да наставници који предају предмете природно-математичке групе јединствени критеријум оцењивања чешће посматрају као предност тестова знања у односу на наставнике који предају предмете друштвено-језичке групе. Разлика у одговорима наставника је статистички значајна на нивоу 0,05 (табела 46).

Табела 46. Предмет који наставник предаје и предности тестова знања у односу на усмено испитивање с обзиром на истоветност критеријума оцењивања

Предмет који наставник предаје	Предност тестова знања: истоветност критеријума оцењивања					
	не		да		укупно	
	f	%	f	%	f	%
друштвено-језичка група	71	38,0	116	62,0	187	100,0
природно-математичка група	46	26,3	129	73,7	175	100,0

$$\chi^2 = 5,64; df=1; p = 0.02; C = 0,12$$

Такође, резултати показују да наставници који предају предмете природно-математичке групе у припреми поступка провере чешће виде предност тестова знања у односу усмено испитивање. Разлика у одговорима наставника је статистички значајна на нивоу 0,05 ($\chi^2=4,33$; $df=1$; $p=0.04$; $C= 0,11$) (слика 7).



Слика 7. Предмет који наставник предаје и предности тестова знања у односу на усмено испитивање с обзиром на сложеност припреме поступка испитивања

Када је реч о осталим предностима тестова знања у односу на усмено испитивање, као и предностима тестова у односу на писане вежбе, одговори наставника који предају различите предмете нису се статистички значајно разликовали (Прилог 3, табеле 66-67).

Када је реч о поређењу класичних и електронских тестова, статистички значајна разлика у одговорима наставника различитих предмета установљена је у односу на истоветност критеријума оцењивања. Међу наставницима који у односу на наведени критеријум предност дају класичним тестовима, постоји разлика у зависности од предмета који предају. Наведену предност класичних

тестова чешће сагледавају наставници природно-математичке групе предмета. Разлика у одговорима наставника је статистички значајна на нивоу 0,05 (табела 47). У односу на остале критеријуме поређења није установљено постојање статистички значајне разлике у одговорима наставника различитих предмета (Прилог 3, табела 68).

Табела 47. Предмет који наставник предаје и предности тестова знања у односу на електронске тестове с обзиром на истоветност критеријума оцењивања

Предмет који наставник предаје	Предност тестова знања: истоветност критеријума оцењивања					
	не		да		укупно	
	f	%	f	%	f	%
друштвено-језичка	160	85,6	27	14,4	187	100,0
природно-математичка	132	75,4	43	25,6	175	100,0

$$\chi^2 = 5,95; df=1; p = 0.02; C = 0,13$$

У разматрању препорука за унапређивање припреме на студијама у домену израде тестова знања, груписањем одговора у две категорије – „унапређивање припреме током студија је потребно“ и „унапређивање припреме током студија није потребно“ – установљено је да су наставници друштвено-језичке групе предмета статистички значајно чешће исказивали потребу за унапређивање припреме током студија у односу на наставнике природно-математичке групе (табела 48). Статистичка значајност разлике у одговорима наставника постоји на нивоу 0,05.

Табела 48. Предмет који наставник предаје и потреба за унапређивањем припреме током студија у домену израде тестова знања

Предмет који наставник предаје	Потреба за унапређивањем припреме током студија					
	да		не		укупно	
	f	%	f	%	f	%
друштвено-језичка група	97	89,8	11	10,2	108	100,0
природно-математичка група	58	76,3	18	23,7	76	100,0
укупно	155	84,2	29	15,7	184	100

$$\chi^2 = 6,12; df = 1; p = 0.01; C = 0,18$$

У домену процене укупне припремљености за израду тестова знања (Прилог 3, табела 69), припреме за израду тестова знања током иницијалног професионалног образовања (Прилог 3, табела 70) и стручног усавршавања у домену израде тестова знања (Прилог 3, табеле 71-72), нису регистроване статистички значајне разлике у одговорима наставника који предају различите предмете.

Када је реч о врсти тешкоћа са којима се суочавају у процесу израде тестова знања и начинима њиховог превазилажења, заступљености и учесталости самосталне израде тестова знања у настави, није установљено постојање статистички значајних разлика у одговорима наставника друштвено-језичке и природно-математичке групе предмета (Прилог 3, табеле 73-74).

У целини посматрано, у разматрању издвојених аспеката примене тестова знања међу наставницима друштвено-језичке и природно-математичке групе предмета постоји више сличности него разлика. Иако обе групе испитаника неповољно оцењују припрему за израду тестова током иницијалног професионалног образовања, регистрована је тенденција да наставници друштвено-језичке групе предмета више исказују потребу за променама у домену припреме за израду тестова знања током студија. Добијени налаз се може протумачити ако се има у виду чињеница да се наставници друштвено-језичке групе предмета традиционално више ослањају на усмена испитивања и писане вежбе (Ebel, 1979). У том смислу, може се очекивати и да је методичка припрема на студијама мање оријентисана на израду тестова знања.

У настави природно-математичке групе предмета, тестови знања су традиционално више заступљени (у форми низова задатака објективног типа и у форми „правих“ тестова знања). Можда се управо у традицији може тражити образложење зашто су наставници природно-математичке групе предмета ређе изражавали потребу за променом припреме на студијама, односно предњачили у указивању на одређене предности тестова знања у односу на усмено испитивање и електронске тестове. Чешће указивање на одређене предности тестова знања у односу на усмено испитивање од стране наставника природно-математичке групе предмета може се објаснити и са становишта природе наставног предмета. У наставни предмета као што су математика, физика, хемија, провера знања се ефикасније остварује применом теста знања него што је то могуће остварити усменим испитивањем (на пример, решавање одговарајућих једначина, примена образаца и формула). Ипак, истицање предности класичних над електронским тестовима с обзиром на јединственост критеријума оцењивања није очекивани налаз. У основи, електронски тест знања обезбеђује аутоматско скоровање, што је објективнији начин вредновања ученичких одговора у односу на ситуацију када то чини наставник.

Образложење добијеног налаза, барем делимично, може се потражити у недовољном познавању концепције електронских тестова знања.

Сходно добијеним резултатима, можемо закључити да хипотеза која се односи на испитивање разлика међу наставницима друштвено-језичке и природно-математичке групе предмета у: сагледавању предности и ограничења тестова знања као начина провере знања ученика; процени властите припремљености за израду тестова знања; врсти тешкоћа са којима се суочавају у процесу израде тестова знања и начинима њиховог превазилажења; учесталости и динамици коришћења тестова знања у настави и сагледавању начина коришћења резултата добијених испитивањем образовних постигнућа ученика – у највећој мери није потврђена.

Године радног стажа наставника у просвети и могућности примене теста знања у настави

Разматрање питања предности и ограничења тестова знања у односу на друге облике праћења и вредновања рада ученика, доприноса иницијалног професионалног образовања и стручног усавршавања припреми наставника за израду тестова знања, тешкоћа са којима се наставници суочавају у процесу израде тестова знања, начина превазилажења тешкоћа и учесталости и динамици коришћења тестова знања, допуњено је испитивањем разлика у одговорима испитаника различитих година радног стажа у просвети.

Истраживањем су регистроване статистички значајне разлике у одговорима наставника различитих година радног стажа у односу на поједина питања предности тестова знања у односу на друге начина праћења и вредновања постигнућа ученика, поједина питања припремљености за израду тестова знања, извесне аспекте тешкоћа са којима се наставници суочавају у изради тестова знања и учесталости самосталне израде тестова током школске године.

Резултати показују да су у поређењу тестова знања и других начина праћења и вредновања постигнућа ученика, статистички значајне разлике утврђене у домену поређења класичних и електронских тестова знања. Наиме, статистички значајна разлика у одговорима наставника различитих година радног стажа установљена је у односу на следеће критеријуме: „захтеви су исти за све ученике“, „критеријум оцењивања је исти за све ученике“ и „ученици не могу да 'варају'“. Истоветност захтева и критеријума оцењивања чешће су предношћу класичних тестова означавали наставници са више година радног стажа. Разлика у одговорима наставника је статистички значајна на нивоу 0,05 у случају истоветности захтева ($\chi^2 = 6,65$; $df = 2$; $p = 0,04$; $C = 0,13$; Прилог 3, табела 5) и у случају истоветности критеријума оцењивања ($\chi^2 = 8,08$; $df = 2$; $p = 0,02$; $C = 0,15$; Прилог 3, табела 6). Умањену могућност „варања“ на тесту, чешће су предношћу класичних тестова означавали наставници са мање од 10 година радног стажа, него њихове искусније колеге. Наиме, могуће је да је млађим наставницима ближе сазнање да електронски тестови у условима лошије стандардизације не гарантују аутентичност одговора ученика. Разлика у одговорима наставника је статистички значајна на нивоу 0,05 (табела 49).

Табела 49. Године радног стажа наставника у просвети и предности тестова знања у односу на електронске тестове с обзиром на могућност да ученици „варају“

Године радног стажа	Предност тестова знања: могућност да ученици „варају“					
	не		да		укупно	
	f	%	f	%	f	%
0-10	137	89,0	17	11,0	154	100,0
10-20	139	98,0	3	2,1	142	100,0
20 и више	62	94,9	4	6,1	66	100,0
укупно	338	93,4	24	6,6	362	100,0

$$\chi^2 = 9,55; df = 2; p = 0,01; C = 0,16$$

У поређењу класичних и електронских тестова с обзиром на преостале задате критеријуме, тестова знања и усменог испитивања, тестова и писаних

вежби, нису регистроване статистички значајне разлике у одговорима наставника различитих година радног стажа (Прилог 3, табеле 78-80).

Процена припремљености за израду тестова знања није се разликовала с обзиром на године радног стажа наставника (Прилог 3, табела 81). Када је реч о припреми за израду тестова знања током студија, установљено је да разлике постоје у односу на поједине садржаје припреме и процену доприноса припреме на студијама развијању вештина потребних за израду тестова знања (резултати преосталих статистичких анализа приказани су у Прилогу 3, табела 82). Статистичка значајна разлика у одговорима наставника различитих година радног стажа регистрована је у односу на следеће садржаје припреме: а) појам и врсте тестова знања ($\chi^2 = 14,80$; $df = 2$; $p = 0.01$; $C = 0,20$; Прилог 3, табела 7) и б) врсте и карактеристике задатака у тесту знања ($\chi^2 = 13,26$; $df = 2$; $p = 0.01$; $C = 0,19$; Прилог 3, табела 8). Наставници са мање година радног стажа чешће су извештавали о заступљености наведених садржаја у току припреме на студијама.

У поређењу са колегама са више педагошког искуства, наставници са мање година радног стажа чешће су извештавали о доприносу припреме на студијама развијању вештина потребних за израду тестова знања. Статистичка значајност разлике у одговорима наставника постоји на нивоу 0,05 (табела 50).

Табела 50. Године радног стажа наставника у просвети и допринос припреме на студијама развијању вештина потребних за израду тестова знања

Године радног стажа	Допринос припреме на студијама развијању вештина потребних за израду тестова знања					
	не		да		укупно	
	f	%	f	%	f	%
0-10	118	76,6	36	23,4	154	100,0
10-20	122	85,9	20	14,1	142	100,0
20 и више	58	87,9	8	12,1	66	100,0
укупно	298	82,3	64	17,7	362	100,0

$$\chi^2 = 6,10; df = 2; p < 0.05; C = 0,13$$

У домену стручног усавршавања, разлике у одговорима наставника различитих година радног стажа регистроване су с обзиром на похађање и укупан број семинара стручног усавршавања у домену израде тестова знања, поједине садржаје разматране на семинарима, допринос семинара развијању вештина потребних за израду тестова знања, консултовање методолошке литературе и похађање летњих/зимских школа (резултати преосталих статистичких анализа приказани су у Прилогу 3, табеле 83-84).

У поређењу са колегама са мање педагошког искуства, наставници са више година радног стажа чешће су извештавали о похађању семинара стручног усавршавања у домену израде тестова знања. Статистичка значајност разлике у одговорима наставника постоји на нивоу 0,01 (табела 51).

Табела 51. Године радног стажа наставника у просвети и похађање семинара стручног усавршавања у домену израде тестова знања

Године радног стажа	Похађање семинара стручног усавршавања у домену израде тестова знања					
	не		да		укупно	
	f	%	f	%	f	%
0-10	82	53,3	72	46,8	154	100,0
10-20	70	49,3	72	50,7	142	100,0
20 и више	20	30,3	46	69,7	66	100,0
укупно	172	47,5	190	52,5	362	100,0

$$\chi^2 = 10,10; df = 2; p < 0,01; C = 0,16$$

У погледу броја семинара које су наставници различитих година радног стажа похађали, утврђена је разлика у корист наставника са више година радног стажа ($\chi^2 = 21,17; df = 8; p < 0,01; C = 0,24$; Прилог 3, табела 9). Статистичка значајност разлике у одговорима наставника на дато питање постоји на нивоу 0,01. Међу наставницима који нису похађали ниједан семинар

посвећен тематици тестова знања значајно је више наставника са мање од 10 година радног стажа, него наставника са више година радног стажа. У категорији наставника који су похађали 5 и више семинара посвећених тематици тестова знања, значајно је више наставника са 20 и више година радног стажа у односу на наставнике са мање година радног стажа у просвети. Добијени налаз могао би се протумачити као очекиван уколико имамо у виду чињеницу да дужи радни стаж значи и више шанси да наставници буду укључени у семинаре стручног усавршавања. Истовремено, налаз се не би могао сматрати добрим уколико се има у виду чињеница да је наставницима на почетку професионалне каријере потребно више помоћи око израде тестова знања него искуснијим наставницима.

Када је реч о садржајима који су разматрани у оквиру семинара стручног усавршавања, статистички значајна разлика у одговорима наставника различитих година радног стажа регистрована је у односу на:

- појам и врсте тестова знања ($\chi^2 = 12,00$; $df = 2$; $p = 0.00$; $C = 0,18$; Прилог 3, табела 10) и
- процес израде тестова знања ($\chi^2 = 8,01$; $df = 2$; $p = 0.02$; $C = 0,15$; Прилог 3, табела 11).

Наставници са више година радног стажа чешће су извештавали о заступљености наведених садржаја на семинарима које су похађали, што је у складу са претходно размотреним налазом о њиховом чешћем учешћу у програмима стручног усавршавања посвећеним изради тестова знања.

Такође, наставници са више година радног стажа чешће су извештавали о доприносу семинара развијању вештина потребних за израду тестова знања (табела 52). Статистичка значајност разлике у одговорима наставника на дато питање постоји на нивоу 0,01.

Табела 52. Године радног стажа наставника у просвети и допринос семинара развијању вештина потребних за израду тестова знања

Године радног стажа	Допринос семинара развијању вештина потребних за израду тестова знања					
	не		да		укупно	
	f	%	f	%	f	%
0-10	110	71,4	44	28,6	154	100,0
10-20	93	65,5	49	34,5	142	100,0
20 и више	32	48,5	34	51,5	66	100,0
укупно	235	64,9	127	35,1	362	100,0

$$\chi^2 = 10,71; df = 2; p < 0.01; C = 0,17$$

Испитивањем је установљено да су наставници са више година радног стажа у просвети чешће учествовали у летњој или зимској школи посвећеној изради тестова знања у односу на наставнике са мање година радног стажа ($\chi^2 = 8,00; df = 2; p = 0.02; C = 0,15$; Прилог 3, табела 12). Статистички значајна разлика на нивоу 0,05 утврђена је и када је реч о консултовању методолошке литературе у домену усавршавања у процесу израде тестова знања (табела 53).

Табела 53. Године радног стажа наставника у просвети и консултовање методолошке литературе у домену усавршавања у процесу израде тестова знања

Године радног стажа	Консултовање методолошке литературе у домену усавршавања у процесу израде тестова знања					
	не		да		укупно	
	f	%	f	%	f	%
0-10	21	13,6	126	81,8	154	100,0
10-20	41	28,9	101	71,1	142	100,0
20 и више	23	34,8	43	65,2	66	100,0
укупно	92	25,4	270	74,6	362	100,0

$$\chi^2 = 8,25; df = 2; p = 0.02; C = 0,15$$

Овог пута, испитивањем је утврђено да су наставници са мање од 10 година радног стажа чешће извештавали о опредељењу за дати облик стручног усавршавања у односу на колеге са више година радног стажа.

Даљом анализом, установљено је да се одговори наставника различитих година радног стажа углавном нису значајно разликовали када су у питању тешкоће са којима се суочавају у конструкцији тестова знања (Прилог 3, табела 85). Разлика је утврђена једино у извештавању о постојању тешкоћа у одређивању дужине теста. Наиме, наставници који имају мање од 10, као и наставници са 20 и више година радног стажа, чешће су извештавали о тешкоћама у одређивању дужине теста знања у односу на наставнике који имају између 10 и 20 година радног стажа. Разлика у одговорима наставника статистички је значајна на нивоу 0,05 (табела 54).

Табела 54. Године радног стажа наставника у просвети и тешкоће у одређивању дужине теста знања

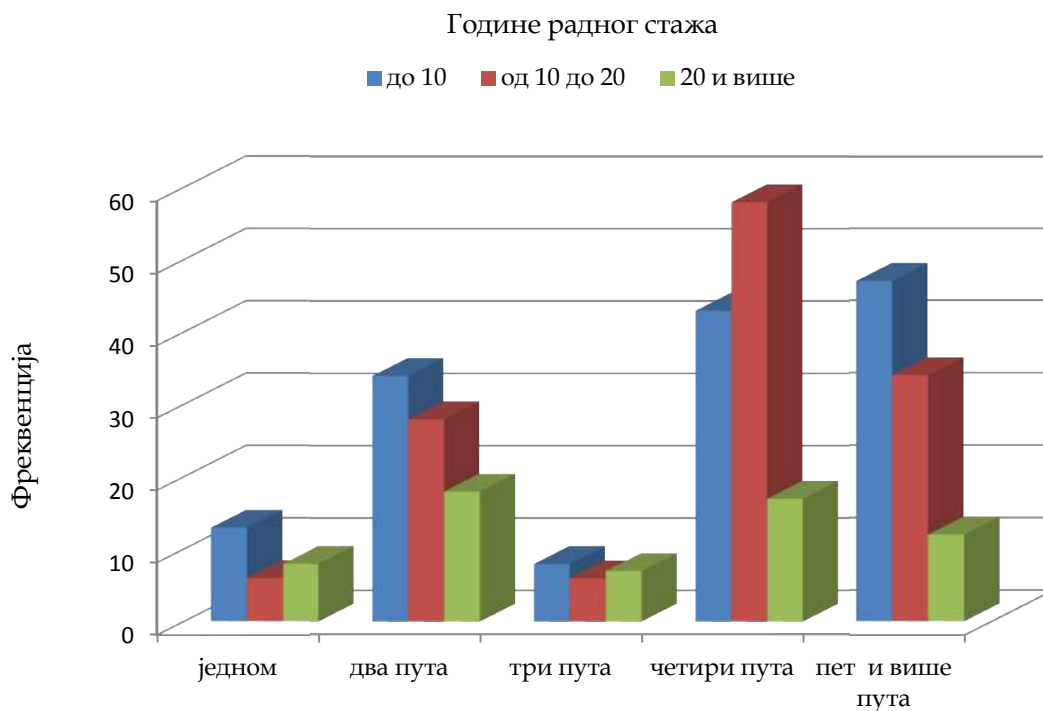
Године радног стажа	Постојање тешкоћа у одређивању дужине теста знања					
	не		да		укупно	
	f	%	f	%	f	%
0-10	121	78,6	33	21,4	154	100,0
10-20	128	90,1	14	9,9	142	100,0
20 и више	54	81,8	12	18,2	66	100,0
укупно	303	83,7	59	16,3	362	100,0

$$\chi^2 = 7,46; df = 2; p = 0.02; C = 0,14$$

У односу на различите начине решавања тешкоћа у изради тестова знања, статистички значајна разлика утврђена је само у односу на консултовање са искуснијим колегама (Прилог 3, табела 85). Наиме, наставници са мање година радног стажа чешће су извештавали о консултовању са колегама у циљу решавања потешкоћа у односу на наставнике са више година радног стажа.

Разлика у одговорима наставника статистички је значајна на нивоу 0,05 ($\chi^2 = 7,58$; $df = 2$; $p = 0,02$; $C = 0,14$; Прилог 3, табела 13).

Иако у погледу учесталости примене тестова знања током школске године није регистрована статистички значајна разлика у одговорима наставника различитих година радног стажа (Прилог 3, табела 85), разлика је утврђена с обзиром на учесталост самосталне израде теста (слика 8).



Учесталост самосталне израде тестова знања током школске године

Слика 8. Године радног стажа наставника у просвети и учесталост самосталне израде тестова знања током школске године

Како подаци показују, у категорији високе учесталости (примена тестова 4, 5 и више пута током године) бројнији су наставници са мање од 10 година радног стажа, односно наставници који имају између 10 и 20 година радног стажа, у односу на наставике са преко 20 година радног стажа. Статистичка значајност разлике у одговорима наставника на дато питање постоји на нивоу 0,05 ($\chi^2 = 16,56$; $df = 8$; $p = 0,04$; $C = 0,22$).

У поређењу са предметом који наставници предају, године радног стажа могу се посматрати као чинилац који више доприноси разлици у одговорима у односу на разматране аспекте примене тестова знања. Иако се и овог пута одговори испитаника нису статистички значајно разликовали у вези са појединим питањима израде и примене тестова знања, у односу на одређена питања разлике су биле значајне. Најупечатљивије разлике регитроване су у домену припремљености наставника за израду тестова знања. Резултати показују да је нешто повољнија процена доприноса иницијалног професионалног образовања карактеристична за наставнике са мање година радног стажа, док је повољнија процена доприноса појединих облика стручног усавршавања карактеристична за наставнике са више година радног стажа.

Значајно чешће од млађих наставника, искуснији наставници су извештавали да не могу да се сете садржаја који су разматрани у оквиру припреме на студијама ($\chi^2 = 13,08$; $df = 2$; $p = 0.01$; $C = 0,19$; Прилог 3, табела 14). Такав налаз је разумљив, с обзиром на временску дистанцу са које наставници извештавају о својим студијама. То може бити једно од објашњења добијених налаза у вези са припремом током студија. Друго објашњење у вези са слабијим извештавањем старијих наставника о похађању припреме током студија могу бити промене студијских програма које су се догодиле у периоду по окончању њиховог иницијалног професионалног образовања.

С обзиром на чињеницу да са дужим стажом, наставник има више шанси да буде укључен у различите облике стручног усавршавања, није изненађујућ податак о већем броју семинара које су похађали искуснији наставници, односно о њиховом чешћем похађању летњих/зимских школа. За разлику од искуснијих колега, млађи наставници у процесу израде тестова више се ослањају на помоћ колега. Такав налаз је разумљив, јер је у стицању професионалних знања и вештина млађим колегама од посебног значаја помоћ колега који имају више искуства (*Професионални развој на нивоу школе/вртића*, 2013). Истовремено, искуснији наставници ће ређе имати потребу да помоћ

затраже од колега, посебно у околностима када они сами имају највише професионалног искуства у колективу.

У нашем испитивању, млађи наставници чешће су указивали на предности електронских тестова знања у односу на класичне тестове које су у складу са савременим теоријских становиштима. Једно од могућих објашњења добијеног налаза је да су млађи наставници нешто боље информисани о електронским тестовима него њихове искусније колеге. Испитивање указује и на тенденцију да се млађи наставници чешће опредељују да самостално израде тест знања када говоримо о високо фреквентној изради (4, 5 и више пута током школске године). С обзиром на то да су млађи наставници скорије завршили студије, могло би се очекивати да су спремнији да примене стечена сазнања (Стојановић, 2011). Када је реч о учесталости самосталне израде тестова знања, ипак се не може рећи да искуснији наставници драстично заостају за млађим колегама када је реч о мање честим тестирањима, те би наредна испитивања омогућила да се добијени налаз провери. Напоменимо да је једно раније испитивање показало да наставници са преко 20 година радног стажа имају израженију потребу за коришћењем нових начина испитивања и оцењивања ученика у односу на наставнике који имају мање од 10 година радног стажа (Максимовић, 2012).

Сходно добијеним резултатима, можемо закључити да је хипотеза која се односи на испитивање разлика међу наставницима различитих година радног стажа у: сагледавању предности и ограничења тестова знања као начина провере знања ученика; процени властите припремљености за израду тестова знања; врсти тешкоћа са којима се суочавају у процесу израде тестова знања и начинима њиховог превазилажења; учесталости и динамици коришћења тестова знања у настави и сагледавању начина коришћења резултата добијених испитивањем образовних постигнућа ученика са становишта унапређивања васпитно-образовног рада школе – делимично потврђена.

МЕТОДОЛОШКЕ КАРАКТЕРИСТИКЕ ТЕСТОВА ЗНАЊА У НАСТАВИ СРПСКОГ ЈЕЗИКА И МАТЕМАТИКЕ

Фазе конструкције теста знања

Испитивањем су добијени подаци о томе кроз које фазе пролази процес конструкције тестова примењених у настави српског језика и математике. Анализа добијених одговора омогућила је да се фазе конструкције теста о којима наставници извештавају упореде са фазама конструкције дефинисаним у домену методологије педагошких истраживања. Компарација је остварена у односу на теоријски модел који дефинише следеће фазе: утврђивање циља који се жели постићи применом теста знања, одређивање садржаја теста, одређивање нивоа знања који се задатком мери, одређивање тежине тестовских задатака, одређивање дужине теста, утврђивање редоследа задатака, избор форме задатка, израда кључа за оцењивање теста знања, утврђивање начина оцењивања теста, састављање упутства за решавање теста, пробно испитивање и формирање, чување и развијање базе задатака (табела 55).

Резултати показују да највећи број наставника (77,5%²) конструкцијом теста обухвата одређивање нивоа знања који се испитује сваким појединачним задатком. Приближно исти број (72,5%) одређује садржај теста. Нешто ређе, али и даље више од половине наставника, утврђује начин оцењивања теста (56,3%) и израђује кључ за оцењивање теста (53,8%). Препознавање преосталих 8 фаза дефинисаних у методолошкој литератури знатно је ређе ($\leq 20\%$). Чак је и одређивање циља теста, тежине задатака, дужине теста, избор форме задатака, пало у сенку коришћења примера готових задатака из различитих извора (30%) и „осталих фаза“ (41,3%). Само три наставника обухваћена испитивањем извештавају о реализацији пробног испитивања теста, два

² Иако је узорак наставника у овом делу истраживања мањи од 100 (обухвата 40 наставника српског језика и 40 наставника математике), у приказу добијених резултата, поред фреквенција користе се и проценти. С обзиром на то да на поједина питања у инструменту није уједначен број одговора наставника српског језика и математике, а у намери да се одговори те две групе пореде, користе се и проценти у приказу резултата.

наставника утврђују редослед задатака, један наставник извештава да формира, развија и чува базу својих задатака, док ниједан наставник не саставља упутство за решавање теста.

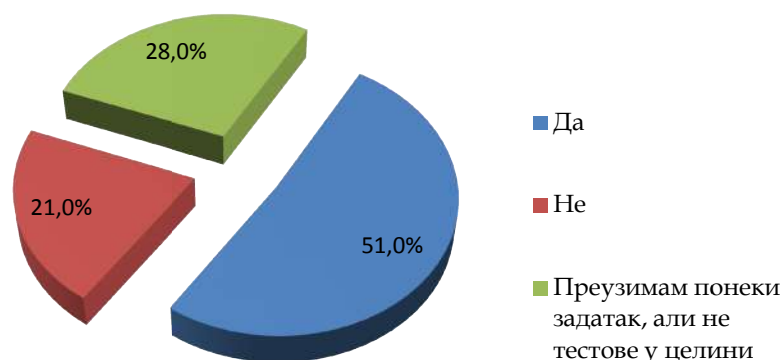
Табела 55. Фазе којима наставници описују процес конструкције теста знања

Фазе у конструкцији теста знања	f	%
одређивање нивоа знања који се задатком мери	62	77,5
одређивање садржаја теста	58	72,5
утврђивање начина оцењивања теста	45	56,3
израда кључа за оцењивање теста	43	53,8
одређивање тежине тестовских задатака	16	20,0
одређивање дужине теста	15	18,8
избор форме задатка	11	13,8
утврђивање циља теста	9	11,3
пробно испитивање	3	3,8
утврђивање редоследа задатака	2	2,5
формирање, чување и развијање базе задатака	1	1,3
састављање упутства за решавање теста	0	0,0
остале фазе	33	41,3
коришћење примера задатака из различитих извора	24	30,0
		N = 80

Поједини наставници су, поред поменутих фаза, извештавали и о постојању фаза које нису обухваћене претходним низом. Међутим, због њихове специфичне природе и учесталости јављања, постале су део даље анализе. Реч је о категоријама „остале фазе“ и „коришћење примера задатака из различитих извора“. Категорија „остале фазе“ обухватила је одговоре као што су: „Узимам у обзир какав је састав одељења“; „Консултујем се са колегиницом“; „Узнем у обзир да ли у одељењу има ученика који раде по ИОП-у и припремим тест за њих ако је потребно“; „Откуцам и одштапам тестове“.

Истраживања показују да се наставници радо ослањају на тестове које је израдио неко други (Porham, 2001). У методолошкој литератури таква пракса се повремено и препоручује (Фурлан, 1964). У случају наших испитаника који су потврдили да користе задатке/тестове других аутора, занимало нас је да

утврдимо ко је аутор наведених тестова и колико често се наставници опредељују да употребе задатке/тестове других аутора.



Слика 9. Коришћење тестова других аутора

Како приказује слика 9, око половине интервјуисаних наставника (51,0%) користи тестове које је саставио неко други. Такође, заступљена је и пракса да се из готових тестова преузимају поједини задаци (28,0%). Ипак, 21,0% испитаника не практикује да користи тестове које нису самостално израдили.

Табела 56. приказује ко су аутори тестова, односно на које изворе се наставници ослањају када трагају за тестовима који се могу преузети у целини или делимино.

Табела 56. Аутори/извори тестова које наставници користе

Аутори/извори тестова	f	%
друге колеге	43	53,8
тимови стручњака које окупља издавачка кућа	33	41,3
тестове који се налазе на Интернету	17	21,3
неко други	3	3,8

N = 80

Дакле, најчесталија пракса је да се користе тестови других колега (53,8%), мада је значајан број наставника који користе тестове различитих

издавачких кућа (41,3%), односно проналазе на Интернету (21,3%). Око трећине наставника консултује тестове других аутора сваки пут када приступа изради теста (табела 57).

Табела 57. Учесталост коришћења тестова других аутора

Учесталост	f	%
једном годишње	5	7,9
два пута годишње	17	27,0
четири пута годишње	11	17,5
скоро сваког месеца	4	6,3
једном месечно	2	3,2
сваки пут када се израђује тест	19	30,2
неки други одговор	5	7,9
укупно	63	100,0

Занимало нас је да сазнамо ко су издавачи које наставници бирају када је реч о преузимању тестова знања. У области српског језика, најчешће су бирани *Нови Логос* (28%), *Клет* (25%), а знатно ређе *Завод за уџбенике*, *Едука* и други. Занимљиви су и следећи коментари наставника: „Користим *Клет*, али не одговара ми све. На пример, ми не дајемо тако дуге тестове као што они имају у понуди“; „*Радна свеска* је довољна и задаци су разврстани по нивоима. Не купујем друге издаваче, усталом нису њихови тестови прилагођени нашој пракси.“

У области математике, најчешће су бирани: *Клет* (15%), *Математички лист* (13%), *Математископ* (10%), а знатно ређе *Круг*, *Креативни центар*, *Венцловић* и *Бигз*.

Када је реч о садржајима са Интернета, наставници српског језика су издвојили сајтове: „Знање није баук“, „Анина учионица“, „Наки 21“, „Учитељи Тошлице“ и Фејсбук групу „Наставник српског језика“. Наставници математике су издвојили сајтове: „Математика на длану“, „Учи слободно“, „WordPress“ и „е-Збирка“.

Резултати показују да се конструкција теста знања у школској пракси најчешће изједначава са одређивањем наставног садржаја који ће се тестом обухватити, нивоа знања који се задатком испитује и утврђивањем начина

оцењивања теста у складу са израђеним кључем. Судећи према добијеним подацима, јасно је да је наставницима у процесу конструкције теста најважније да саставе задатке који адекватно репрезентују наставно градиво и да адекватно варирају ниво знања који се задатком мери. Иако се може рећи да то јесу кључне фазе, конструкција теста знања обухвата и знатан број других фаза чије превиђање може значајно да угрози квалитет теста. Све фазе у конструкцији теста које су дефинисане у методолошкој литератури, емпиријски су провераване, односно њихово разматрање показало се релевантним за квалитет теста у целини. Укажимо само на чињеницу да ако се не разуме исправно циљ теста, тешко се разрађује садржај и нивои знања које задаци мере. Као што смо видели, преко 70% наставника у конструкцији теста разматра садржај теста, а тек 11,3% прецизно одређује његов циљ.

Претходно приказани резултати о тешкоћама наставника у изради теста могу се посматрати као софистицирана анализа корака кроз које израда теста знања пролази, а који нису тако видљиви када наставници самостално описују процес конструкције теста знања. На пример, у описивању фаза кроз које пролази конструкција теста, укупно 20% наставника је одређивање тежине задатака сагледало као фазу у процесу израде теста знања, а 33,3% извештава да се суочава са тешкоћама у одређивању тежине задатака. Такође, избор адекватне врсте задатака за одређени наставни садржај, издвојен је као посебна фаза конструкције у 13,8% случајева, а као потешкоћа у 21,3% случајева. Сличне разлике примећене су и у вези са развијањем и чувањем базе задатака, утврђивањем редоследа задатка и израдом упутства за испитанике.

Садржај теста знања

Разматрање садржаја теста знања омогућава да се сагледају полазне фазе у конструкцији теста знања. Реч је о фазама од којих зависи разрада осталих фаза у конструкцији и којима се пружају одговори на питања: шта је циљ израде теста; шта представља основу за разраду садржаја теста знања; који се нивои

знања испитују одговарајућим задацима; како су задаци формулисани; која је тежина задатака у тесту; шта чини садржај упутства за решавање теста? Испитивање је омогућило да утврдимо одговоре на наведена питања када је реч о тестовима које су примењени у настави српског језика и математике.

Најчесталија врста тестова који су анализирани (38,8%) јесу тестови који се задају са циљем провере ученичких знања у току школске године, решавају се током целог часа и могу се оценити. Већ следећи по учесталости су контролни задаци (30,0%), а потом следе иницијални тестови знања (10,0%) и писмени задаци (8,8%). Укупно 6 наставника (7,5%) приложило је полугодишњи тест знања, три наставника (3,8%) су се одлучила за петнаестоминутни тест, а један наставник (1,3%) за годишњи тест знања.

Примена иницијалних тестова знања налази своје утемељење у Правилнику о оцењивању ученика у основном образовању и васпитању (2013; члан 10) који предвиђа да се на почетку школске године процењују претходна постигнућа ученика која су од значаја за наставни предмет. Резултати иницијалног оцењивања се не оцењују и служе за планирање рада наставника и даље праћење напредовања ученика. Дакле, иницијални тест је једна од форми којом се може реализовати иницијално испитивање.

Петнаестоминутни тестови које су наставници такође прилагали за анализу, у основи су замишљени као *микротестови* (Кундачина и Банђур, 1997). Микротестови или петоминутне провере знања примењују се на часовима обраде новог градива, са циљем испитивања продуктивности одржаног часа. Иако би према дефиницији, такве провере требало да трају максимално 5 минута, у пракси оне углавном трају дуже (Кундачина и Банђур, 1997). Њихова сврха одговара испитивањима која не морају бити најављена, на основу којих се не изводи оцена, и могу трајати до 15 минута (*Правилник о оцењивању ученика у основном образовању и васпитању*, 2013; члан 13).

Полугодишњи и годишњи тестови знања такође налазе своје место у пракси школског испитивања ученичких знања. Иако би разлоге за њихову примену требало детаљније испитати, постоје назнаке да су поједини издавачи уџбеника подстакли наставнике да размишљају о овој врсти тестова

објављивањем полугодишњих и годишњих тестова намењених примени у редовном наставном раду. Такође, Завод за вредновање квалитета образовања и васпитања у оквиру својих редовних делатности периодично спроводи иницијална и годишња тестирања ученика у циљу процена остварености циљева, општих и посебних стандарда постигнућа ученика, што се такође може сматрати подстицајем да се израде дате форме тестова.

Када је реч о циљу задавања анализираног теста, табела 58. приказује дистрибуцију одговора добијених на дато питање.

Табела 58. Циљеви задавања теста знања

Циљ	f	%
оцењивање ученика	66	82,5
провера савладаности градива без извођења оцене	58	72,5
утврђивање претходног знања без извођења оцене	9	11,4
провера остварености стандарда на основном нивоу	4	5,0
планирање даљег рада	3	3,4
нешто друго	3	3,4

N = 80

Дакле, најчешћи одговор у вези са циљем задавања теста јесте оцењивање ученика, мада је доста фреквентан одговор и провера савладаности градива која нема за циљ извођење оцене. Остале категорије одговора су значајно мање заступљене. Категорији одговора „нешто друго“ припадају одговори: *„вредновање рада наставника“*; *„припрема за полугодишњи тест“* и *„припрема за завршни тест“*.

Судећи према подацима добијеним испитивањем, наши наставници још увек чешће приступају тестирању са циљем испитивања ученичких знања него са циљем унапређивања процеса учења. Када се каже „тест“ вероватније је да се мисли на „резултат учења“, него на „процес учења“ (Cooper & Leiter, 1980; Buser, 1996). Рефокусирање процеса тестирања које се на плану теорије догодило крајем XX века (Porham, 2001), још увек није упечатљив део наше школске стварности.

Највећи број наставника у разради садржаја теста полази од наставног програма (85,0%) и од стандарда знања дефинисаних за одговарајући предмет (72,5%) (табела 59).

Табела 59. Основа за разраду садржаја теста знања

Основа	f	%
наставни програм	68	85,0
стандарди знања дефинисани за наставни предмет	58	72,5
уџбеник	28	35,0
садржај завршног испита	5	6,3
градиво које наставник сматра битним за наставак учења	5	6,3
садржај који је обрађиван на часу	4	5,0
интересовања ученика	2	2,5
нешто друго	4	5,0

N = 80

Око трећине испитаника (35,0%) у разради садржаја теста узима у обзир садржај уџбеника. Остале категорије одговора су значајно мање заступљене, указујући на тенденцију да се садржај теста у настави обликује према захтевима завршног испита (Мале матуре), односно према личној процени наставника шта је ученицима битно за наставак учења одговарајућег предмета или наставак школовања. Такође, четири наставника су нагласила да је основа за разраду теста садржај који је обрађиван на часу, а који се не мора нужно поклопити са садржајем уџбеника (на пример: *„Оно што смо вежбали на часу је кључно, не садржај уџбеника“*). У категорији „нешто друго“ заступљени су одговори који указују на ослањање наставника на садржај приручне литературе и на садржај са интернета.

Може се рећи да су добијени подаци очекивани. У разрадама садржаја теста знања, неопходно је поћи од наставног програма. Исто тако, важно је да се као основа користе и стандарди дефинисани за наставни предмет. Стандарди помажу наставницима да препознају да садржај теста не чини само градиво, већ и мисаоне операције које се примењују над тим градивом (Crocker & Algina, 1986). У том смислу, стандарди су конкретнија основа за разраду теста

у односу на програм или уџбеник. Њихова ређа заступљеност у односу на наставни програм у домену разраде садржаја теста може се објаснити чињеницом да стандарди постоје тек од 2009. године. У извесном смислу, подаци су и охрабрујући, јер показују да је ослањање наставника на властити суд о томе шта би требало да сачињава тест знања ипак ређе заступљено него ослањање на програм, стандарде или уџбеник.

У наставку, приказаћемо заступљеност задатака који испитују различите нивое знања. О нивоима знања који се одређеним задатком у тесту испитују, извештавали су сами наставници. Резултати показују да се у тестовима које наставници самостално конструишу, задаци могу разврстати у три категорије према нивоу знања који мере. Реч је о нивоима знања који одговарају нивоима познавања чињеница, разумевања и примене знања, према Блумовој класификацији нивоа знања. Иако је постојала могућност да се укаже и на друге нивое знања (дефинисане Блумовом класификацијом или класификацијом других аутора), испитивањем није утврђена заступљеност задатака који такве нивое знања испитују.

Највећи број (86,3%) анализираних тестова из српског језика и математике садржи задатке који испитују три нивоа знања. Значајно је мањи број тестова (11,3%) који испитују два нивоа знања. Прецизније, шест тестова испитује нивое познавања чињеница и разумевања, два теста испитују нивое разумевања и примене, а један тест испитује ниво познавања чињеница и примене. Само два теста садрже задатке који испитују један ниво знања: један тест испитује ниво познавања чињеница, а други ниво примене знања. У анализираним тестовима није било задатака којима се испитују нивои анализе, синтезе и евалуације (табела 60).

Табела 60. Укупан број нивоа знања који се испитују тестом

Број нивоа знања	Број тестова	
	f	%
1	2	2,5
2	9	11,3
3	69	86,3
укупно	80	100,0

Анализирани тестови из српског језика и математике садрже највише задатака који испитују ниво познавања чињеница (40,1%), нешто је мање задатака који испитују ниво разумевања градива (34,3%), а најмање задатака примене знања (25,6%) (табела 61).

Табела 61. Укупан број задатака обухваћених анализом према нивоима знања које испитују

Нивои знања који се испитују задатком	f	%
познавање чињеница	262	40,1
разумевање	224	34,3
примена	167	25,6
укупно	653	100,0

Дистрибуција података у табели 62. је доста разуђена, али показује да највећи број анализираних тестова (75,1%) садржи од једног до четири задатка који испитују ниво познавања чињеница. Нешто заступљенији од других су тестови који садрже укупно два задатка који испитују ниво познавања чињеница.

Табела 62. Укупан број задатака који испитују ниво познавања чињеница на нивоу појединачног теста знања

Број задатака	Број тестова	
	f	%
0-1*	2	2,5
1	12	15,0
1-2	5	6,3
2	16	20,0
2-3	2	2,5
3	11	13,8
4	14	17,5
4-5	1	1,3
5	6	7,5
6	5	6,3
7	4	5,0
8	1	1,3
15	1	1,3
укупно	80	100,0

Када је реч о задацима који испитују ниво разумевања градива, табела 63. показује да око половине (55,1%) анализираних тестова садржи два или три задатка који испитују ниво разумевања градива. Нешто заступљенији од других су тестови који садрже укупно два задатка који испитују ниво разумевања градива.

* Приближно 90% задатака обухваћених истраживањем не дели се на ајтеме различите према нивоу знања који мере, тежини или форми. У случају задатака код којих постоји таква подела, емпиријски распон поделе задатака на ајтеме кретао се од 0,25 до 0,5. Тако, вредност 0,25 означава један од четири ајтема на које се задатак дели. Вредност 0,3 подразумева један од три ајтема на које се задатак дели. Вредност 0,5 означава један од два ајтема на које се задатак дели. То значи да распон, на пример, од 0 до 1 може да обухвати вредности 0,25; 0,3 и 0,5; распон од 1 до 2 вредности 1,25; 1,30 и 1,50. С обзиром на изразиту доминацију задатака који се не деле на ајтеме према нивоу знања који мере, тежини или форми, одлучили смо се за приказ броја задатака у целим бројевима. Навођење распона између два цела броја користи се и у наставку текста.

Табела 63. Укупан број задатака који испитују ниво разумевања градива на нивоу појединачног теста знања

Број задатака	Број тестова	
	f	%
0	3	3,8
1	10	12,5
1-2	3	3,8
2	27	33,8
2-3	1	1,3
3	16	20,0
4	9	11,3
4-5	2	2,5
5	4	5,0
6	2	2,5
9	2	2,5
12	1	1,3
укупно	80	100,0

Када је реч о задацима који испитују ниво примене знања, већина (73,8%) анализираних тестова садржи један, два или три задатка који испитују ниво примене знања (табела 64).

Табела 64. Укупан број задатака који испитују ниво примене знања на нивоу појединачног теста знања

Број задатака	Број тестова	
	f	%
0	10	12,5
1	22	27,5
1-2	2	2,5
2	15	18,8
3	20	25,0
4	5	6,3
5	4	5,0
6	2	2,5
укупно	80	100,0

Већина анализираних тестова садржи задатке који испитују три различита нивоа знања – ниво познавања чињеница, разумевања градива и примене знања. Ипак, најзаступљенији су задаци који утврђују ниво познавања чињеница. Добијени резултат је у сагласности са резултатима ранијих испитивања спроведених у нашој земљи (Гргин, 2001; Баковљев, 1997; Хавелка, 2000; Бјекић и Голубовић, 2003; Girps, 2003; Дубљанин, 2010; Илић-Голубовић, 2011). Када је реч о задацима за испитивање нивоа познавања чињеница, број задатака на нивоу теста варира у распону од једног ајтема до 15 задатака. Најчешће, анализирани тестови садрже до четири задатка на нивоу познавања чињеница. Када је реч о задацима за испитивање нивоа разумевања градива, број задатака на нивоу теста варира у распону од једног до 12 задатака. Најчешће, анализирани тестови садрже до три задатка на нивоу разумевања. Када је реч о задацима за испитивање нивоа примене знања, број задатака на нивоу теста варира у најмањем распону, крећући се од једног до 6 задатака. Најчешће, анализирани тестови садрже до три задатка на нивоу примене знања. У целини посматрано, међу анализираним тестовима, типичан тест садржи два задатка на нивоу познавања чињеница, два задатка на нивоу разумевања градива и један задатак примене знања.

Анализом узорка задатака који нису подељени на ајтеме различите тежине, нивоа знања које мере или различите форме, регистрована је заступљеност глагола који се користе у формулацији задатака. Резултати показују да глаголи који се најчешће користе у формулацијама задатака јесу: глагол „одредити“ (16,9%), глагол „написати“ (16,4%) и глагол „израчунати“ (14,4%). У значајној мери (15,6%) заступљени су и задаци који не садрже директиван глагол, односно задаци у чијој формулацији се не експлицира коју мисаону радњу је потребно ученик да примени. Реч је о задацима који ученика не упућују на то да одређене појмове објасни, да реши постављену једначину, исправно напише наведене именице и томе слично, већ је то имплицирано у формулацији задатка (на пример: *„Милица је имала 30 јабука. Ако је продала петину, колико јој је остало?“*). Наведене задатке, наставници

обично означавају текстуалним задацима, тако да ћемо се у приказу резултата у наставку користити управо датим називом.

На нивоу познавања чињеница, у формулацијама задатака преовладава глагол „написати“ (18,9%), а потом следе текстуални задаци (14,3%) и формулације задатака које садрже глагол „израчунати“ (13,1%). На нивоу разумевања, преовладава глагол „одредити“ (21,7%), а потом следе глагол „израчунати“ (14,0%) и глагол „написати“ (13,6%). На нивоу примене знања, најзаступљенији је глагол „одредити“ (18,5%), друго место деле глаголи „написати“ и „израчунати“ и текстуални задаци (16,7%), а на трећем месту је глагол „подвући“ (10,5%).

Може се приметити да су најфреквентнији глаголи на различитим нивоима знања готово идентични. Правилан избор глагола у формулацији задатака важан је са становишта упућивања ученика на примену одговарајућих мисаоних операција у циљу решавања задатка. Имајући то у виду, добијени налаз се не може оценити као повољан, јер упозорава да код наставника није довољно развијена свест о важности формулације задатке. Податак који такође привлачи пажњу јесте висока заступљеност текстуалних задатака на нивоу познавања чињеница. Реч је о задацима за које има основа тврдити да су ученицима тешки (Дубљанина, 2010), а што не би требало да буде карактеристика задатака на нивоу познавања чињеница. Добијени налази указују на потребу да се темељније испита интенција наставника да одговарајућом формулацијом задатака усмере мисаоне процесе код ученика.

Као што задаци могу варирати према нивоу знања који испитују, задаци могу имати и различиту тежину. У испитивању су разликована три нивоа тежине задатака: лаки задаци, средње тешки и тешки задаци. Као и у случају одређивања нивоа знања који се задатком мери, и овог пута, о тежини задатака извештавали су сами наставници.

Како подаци показују, највећи број (82,5%) анализираних тестова садржи задатке са три нивоа тежине. Значајно је мањи број тестова који садрже задатке са два нивоа тежине (10,0%). Прецизније, четири теста садрже лаке и средње тешке задатке, три теста садрже средње тешке и тешке задатке, а један тест

обухвата лаке и тешке задатке. Укупно 6 тестова садржи задатке једног нивоа тежине: четири теста садрже само средње тешке задатке, а један тест садржи само лаке задатке (табела 65).

Табела 65. Укупан број нивоатежине који се тестом испитује

Број нивоа тежине	Број тестова	
	f	%
1	6	7,5
2	8	10,0
3	66	82,5
укупно	80	100,0

Анализирани тестови из српског језика и математике садрже највише задатака средње тежине (40,0%). На нивоу узорка анализираних тестова, уочава се и тенденција да лаких задатака (35,8%) буде више него тешких задатака (24,3%) (табела 66).

Табела 66. Укупан број задатака обухваћених анализом према нивоима тежине

Нивои тежине задатака	f	%
лаки	233,5	35,8
средње тешки	261	40,0
тешки	158,5	24,3
укупно	653	100,0

Око три четвртине (71,5%) анализираних тестова садрже један, два, три или четири лака задатка. У погледу заступљености, благу предност имају тестови који садрже два лака задатка (табела 67).

Табела 67. Укупан број лаких задатака на нивоу појединачног теста знања

Број задатака	Број тестова	
	f	%
0	7	8,8
1	13	16,3
1-2	1	1,3
2	16	20,0
2-3	1	1,3
3	13	16,3
3-4	1	1,3
4	12	15,0
5	7	8,8
6	4	5,0
7	5	6,3
Укупно	80	100,0

Више од половине (63,9%) анализираних тестова садржи два, три или четири средње тешка задатка. Нешто учесталији од других су тестови који садрже два средње тешка задатка (табела 68).

Табела 68. Укупан број средње тешких задатака на нивоу појединачног теста знања

Број задатака	Број тестова	
	f	%
0	2	2,5
0-1	1	1,3
1	6	7,5
1-2	4	5,0
2	20	25,0
2-3	1	1,3
3	16	20,0
3-4	1	1,3
4	13	16,3
4-5	1	1,3
5	6	7,5
6	3	3,8
7	3	3,8
9	1	1,3
10	2	2,5
укупно	80	100,0

Већина анализираних тестова (73,9%) садржи један, два или три тешка задатка. Приближно једнако су учестали тестови са једним и са два тешка задатка, а потом следи категорија тестова са три тешка задатка. Укупно 12 анализираних тестова не садржи тешке задатке (табела 69)

Табела 69. Укупан број тешких задатака на нивоу појединачног теста знања

Број задатака	Број тестова	
	f	%
0	12	15,0
1	20	25,0
1-2	3	3,8
2	19	23,8
2-3	2	2,5
3	15	18,8
4	4	5,0
5	2	2,5
6	1	1,3
7	2	2,5
Укупно	80	100,0

Добијени подаци показују да анализирани тестови садрже највише средње тешких задатака. Такав налаз се може сматрати оправданим са становишта класичне теорије тестирања (Буквић, 1996). Када је реч о лаким задацима, број задатака на нивоу теста варира у распону од једног до 7 задатака. Најчешће, анализирани тестови садрже до четири лака задатка. Када је реч о средње тешким задацима, број задатака на нивоу теста варира у нешто већем распону, крећући се од једног до 10 задатака. Најчешће, анализирани тестови садрже два, три или четири средње тешка задатка. Када је реч о тешким задацима, број задатака на нивоу теста варира у распону од једног до 7 задатака. Најчешће, анализирани тестови садрже до три тешка задатка. У целини посматрано, међу анализираним тестовима, типичан тест садржи два лака задатка, два средње тешка задатка и један или два тешка задатка.

Однос тежине задатка и нивоа знања који се њиме мери није детаљније истраживан у испитивању које смо спровели, али је несумњиво важно методолошко питање. Стиче се утисак да се нивои знања о којима смо разговарали са наставницима повремено поистовећују са нивоима постигнућа према образовним стандардима. Истовремено, дефинисање стандарда на три нивоа лако се може помешати са распонима тежине задатака, што за последицу има поистовећивање нивоа знања који задатак мери са тежином задатка. По свему судећи, разграничење појмова „ниво знања“, „тежина задатка“ и „ниво постигнућа према стандардима“ за поједине наставнике представља проблем.

Када говоримо о садржају теста знања, потребно је да се осврнемо и на питање садржаја упутства за испитанике. У методолошкој литератури састављање упутства за испитанике једна је од фаза конструкције теста (Taylor, Greer & Mussio, 1978; Буквић, 1996; Банђур и Поткоњак, 1999; Мужих, 2004; Фајгел, 2004). Упутство је најчешће интегрални део теста знања, мада може бити саопштено и усменим путем. Значајан број наставника обухваћених нашим испитивањем (36,3%) ученицима не саопштава било каква упутства пре почетка рада. Међу наставницима који припремају упутство за решавање теста, убедљива већина (96,1%) опредељује се за саопштавање упутства искључиво усменим путем.

Структура теста знања

Питање структуре теста надовезује се на разматрање садржаја теста знања. У вези са структуром теста, испитивањем су прикупљени подаци који се односе на: дужину теста, редослед задатака у тесту, заступљеност задатака различите форме, карактеристике задатака одређене форме (колико одговора је понуђено на нивоу задатка, број тачних одговора на нивоу задатка), повезаност форме задатка и нивоа знања који се задатком испитује, повезаност форме и тежине задатка.

Дужина анализираних тестова утврђивана је на основу броја задатака које тест садржи. Број задатака у тесту креће се од 2 до 18 (табела 70). Најучесталији су тестови са 5 задатака (22,5%), а затим следе тестови са 6 задатака (15,0%), односно са 10 задатака (15,0%).

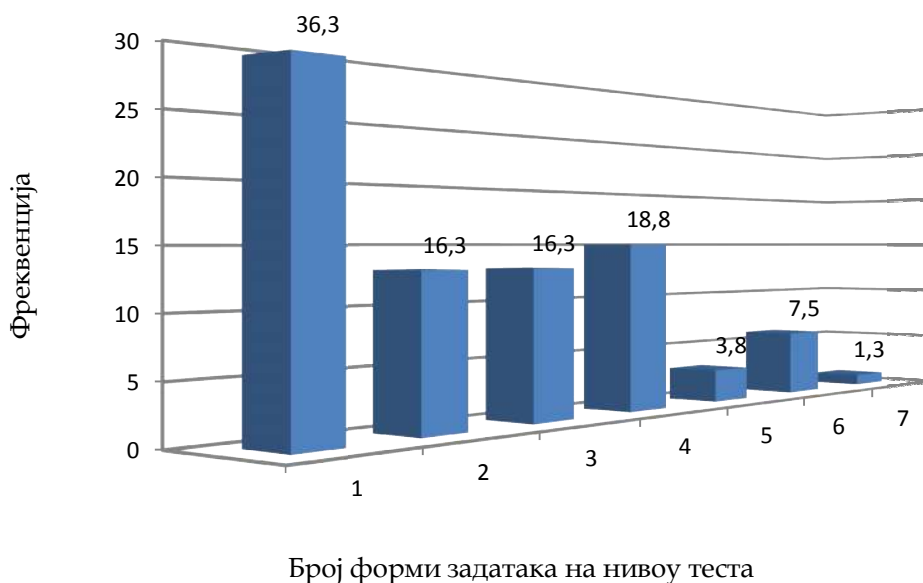
Табела 70. Укупан број задатака у тесту

Број задатака	Број тестова	
	f	%
2	1	1,3
3	1	1,3
4	7	8,8
5	18	22,5
6	12	15,0
7	2	2,5
8	2	2,5
9	6	7,5
10	12	15,0
11	3	3,8
12	6	7,5
13	3	3,8
14	1	1,3
15	4	5,0
18	2	2,5
укупно	80	100,0

Прегледом тестова, утврђено је да редослед задатака у значајној мери варира од теста до теста. Укупно 20,0% тестова у потпуности уважава принцип уређивања задатака од лакших, преко средње тешких ка тешким задацима, који се препоручује у литератури (Taylor, Greer & Mussio, 1978; Савовић и сар, 2006). Близу половине анализираних тестова (48,8%) само започиње лаким задатком, чиме се делимично уважава наведени принцип (Taylor, Greer & Mussio, 1978). Преосталих 31,2% анализираних тестова већ са првим задатком одступа од препорученог принципа.

Даља анализа тестова омогућила је да се утврди заступљеност одговарајућих форми задатака у тестовима. Форме које су обухваћене анализом биле су: задаци двочланог избора, вишечланог избора, допуњавања, испуњавања налога, задаци са кратким одговорима, задаци спаривања, задаци сређивања (Баковљев, 1997). Известан број задатака издвојен анализом није се могао сврстати ни у једну од наведених форми, тако да су обухваћени посебном категоријом означеном као „задаци неке друге форме“.

Највећи број анализираних тестова (36,3%) садржи само једну форму задатака. Према учесталости следе тестови са четири форме задатака (18,8%), потом са две (16,3%) и са три (16,3%) (слика 10). Приближно трећина задатака испуњава препоручени стандард да се на нивоу теста варирају три или четири форме задатака (Богавац и сар, 1972).



Слика 10. Укупан број форми задатака на ниво појединачног теста

Задаци испуњавања налога су најзаступљенија форма задатака у анализираним тестовима (57,2%), а према учесталости следе задаци са кратким

одговорима (18,4%). Остале форме су заступљене са уделом мањим од 10% (табела 71).

Табела 71. Укупан број задатака одређене форме у анализираним тестовима

Задаци	Број задатака	
	f	%
испуњавања налога	373,5	57,2
са кратким одговорима	120	18,4
вишечланог избора	56	8,6
допуњавања	42,5	6,5
друге форме	34,5	5,3
спаривања	20	3,1
двочланог избора	4,5	0,7
сређивања	2	0,3
укупно	653	100,0

Задатке испуњавања налога садрже готово сви анализирани тестови (96,3%). Задатке са кратким одговорима садржи око половине тестова обухваћених анализом (51,3%). Задаци допуњавања, задаци неке друге форме и задаци вишечланог избора заступљени су у, приближно, трећини анализираних тестова. Преостале форме су ређе заступљене (табела 72).

Табела 72. Заступљеност одговарајућих форми задатака на нивоу теста

Задаци	Заступљеност форме				укупно тестова	
	да		не			
	f	%	f	%	f	%
испуњавања налога	77	96,3	3	3,7	80	100,0
са кратким одговорима	41	51,3	39	48,8	80	100,0
допуњавања	27	33,8	53	66,3	80	100,0
друге форме	24	30,0	56	70,0	80	100,0
вишечланог избора	22	27,5	58	72,5	80	100,0
спаривања	16	20,0	64	80,0	80	100,0
двочланог избора	5	6,3	75	93,8	80	100,0
сређивања	2	2,5	78	97,5	80	100,0

Свега три теста обухваћена анализом (3,7%) не садрже задатке испуњавања налога. Готово половина анализираних тестова из српског језика и математике (44,2%), садржи између 4 и 5 задатака испуњавања налога. Према учесталости следе тестови који садрже између једног и два задатка испуњавања налога, тестови са два до три задатка испуњавања налога (табела 73).

Табела 73. Укупан број задатака испуњавања налога на нивоу теста

Број задатака	Број тестова	
	f	%
1-2	13	16,9
2-3	10	13,0
4-5	34	44,2
5-6	7	9,1
7-8	4	5,2
8-9	3	3,9
9-10	3	3,9
12-13	2	2,6
14	1	1,3
укупно	77	100,0

Укупно 39 анализираних тестова (48,8%) не садржи задатке са кратким одговорима. Међу тестовима који садрже наведену форму, најучесталији су тестови који садрже до једног задатака са кратким одговорима. Према учесталости следе тестови који садрже између једног и два задатка са кратким одговорима и тестови са 5 до 6 задатака са кратким одговорима (табела 74).

Табела 74. Укупан број задатака са кратким одговорима на нивоу теста

Број задатака	Број тестова	
	f	%
0-1	14	34,1
1-2	9	22,0
2-3	4	9,8
3-4	4	9,8
5-6	7	17,1
7	2	4,9
11	1	2,4
укупно	41	100,0

Када је реч о задацима допуњавања, укупно 53 (66,3%) анализирана теста не садрже задатке дате форме. Највећи број тестова који садрже задатке дате форме, обухвата до једног задатка допуњавања, а према учесталости следе тестови који садрже између једног и два задатка допуњавања (табела 75).

Табела 75. Укупан број задатака допуњавања на нивоу теста

Број задатака	Број тестова	
	f	%
0-1	16	59,3
1-2	7	25,9
2-3	2	7,4
4	2	7,4
укупно	27	100,0

Када је реч о задацима вишечланог избора, укупно 58 (72,5%) анализираних тестова не садржи задатке дате форме. Највећи број тестова који садрже задатке дате форме обухватају до једног задатка вишечланог избора, потом следе тестови који садрже између једног и два задатка и тестови са три задатка вишечланог избора (табела 76).

Табела 76. Укупан број задатака вишечланог избора на нивоу теста

Број задатака	Број тестова	
	f	%
0-1	8	36,4
1-2	5	22,7
3	5	22,7
4-5	2	9,1
6-7	1	4,6
8-9	1	4,6
укупно	22	100,0

Доминантно опредељење наших испитаника је да у задацима вишечланог избора понуде 3-4 алтернативе. Можемо рећи да изабрано решење налази подршку у методолошкој литератури (Buser, 1996; Rudner & Schafer, 2002; Frey *et al.*, 2005).

Укупно 64 (80,0%) теста обухваћена анализом не садрже задатке спаривања. Од преосталих 16 (20,0%) тестова, већина садржи један задатак спаривања (75,0%), а мањи број (25,0%) садржи два задатка спаривања. У конструкцији већине (70%) анализираних задатака спаривања уважена је препорука да број информација у колонама буде различит (Porham, 2003; Frey *et al.*, 2005), као и да број понуђених информација у појединачној колони не прелази 7 (Frey *et al.*, 2005).

Када је реч о задацима двочланог избора, већина анализираних тестова (93,8%) не садржи задатке дате форме. Преостали тестови (6,3%) садрже до једног задатка двочланог избора.

Већина анализираних тестова (97,5%) не садржи задатке сређивања. Само два теста (2,5%) (један из српског језика и један из математике) садрже задатке наведене форме. Дати тестови садрже по један задатак сређивања, што их чини најређе заступљеном формом у анализираним тестовима.

Укупно 24 (30,0%) теста садржала су задатке сврстане у „неку другу форму“ задатака. Реч је о задацима које најчешће (72,5%) карактеришу захтеви

да се у наведеном тексту подвуку одговарајуће речи (именице, придеви, заменице, погрешно написане речи и слично) или да се усправним цртама реченица подели на одговарајуће целине (комуникативне реченице, апозиције и слично) (27,5%).

У наставку анализе, избор форме задатка довешћемо у везу са нивоом знања који задатак испитује (табела 77).

Табела 77. Укрштање форме задатка и нивоа знања који се задатком испитује

Задаци	ниво познавања чињеница		ниво разумевања градива		ниво примене знања		укупно	
	f	%	f	%	f	%	f	%
двочланог избора	1	22,2	3,5	77,8	0	0,0	4,5	100,0
вишечланог избора	33	58,9	15	26,8	8	14,3	56	100,0
допуњавања	26,5	62,4	10	23,5	6	14,1	42,5	100,0
са кратким одговорима	49	40,8	43	35,8	28	23,3	120	100,0
испуњавања налога	122	32,7	139,5	37,3	112	30,0	373,5	100,0
спаривања	13	65,0	3	15,0	4	20,0	20	100,0
сређивања	1	50,0	1	50,0	0	0,0	2	100,0
друге форме	16,5	47,8	9	26,1	9	26,1	34,5	100,0

Задаци испуњавања налога у анализираним тестовима представљају доминантну форма задатка на свим нивоима испитивања знања. Заступљеност дате форме задатка на три издвојена нивоа знања готово да је уједначена. Добијени подаци су у извесном раскораку са препорукама које срећемо у литератури и према којима се примена дате форме задатака обично везује за испитивање виших нивоа знања (Popham, 2003; Roberson-Scott,2005). Наши

испитаници показују тенденцију да задатке са кратким одговорима користе за испитивање сва три нивоа знања, мада њихова заступљеност опада како се крећемо ка вишим нивоима знања. У методолошкој литератури, наглашава се да примена задатака са кратким одговорима може да одговара испитивању знања различите сложености, док Пофам (2003) упозорава да се њима најчешће испитује репродукција знања. Резултати нашег испитивања су у сагласности са Пофамовим полазиштем. Када је реч о задацима вишеструког избора, њихова примена се доводи у везу са испитивањем различитих нивоа знања (Buser, 1996; Popham, 2003; Roberson-Scott, 2005). Наши испитаници користе задатке вишеструког избора за испитивање сва три нивоа знања, мада њихова заступљеност опада како се крећемо ка вишим нивоима знања. Нешто већа заступљеност задатака вишеструког избора у домену испитивања познавања чињеница у сагласности је са резултатима појединих ранијих испитивања (Гојков, 2012). Анализа приложених тестова знања показује да задаци допуњавања прате сличну дистрибуцију као и задаци са вишеструким понуђеним одговорима. И у литератури, њихова примена се доминантно везује управо за испитивање нивоа познавања чињеница (Popham, 2003; Roberson-Scott, 2005). Фреквентност задатака спаривања и задатака других форми је ниска, али указује на тенденцију да се њихово коришћење чешће везује за испитивање нивоа познавања чињенице, него за испитивање виших нивоа знања. Када је реч о задацима спаривања, таква пракса је у сагласности са становиштем Пофама (2003), и делимично са становиштем Роберсон-Скот (2005). Испитивањем је регистрована слаба фреквенција задатака двочланог избора и задатака сређивања, тако да је могућност даље анализе била ограничена.

Као што смо претходно истакли, утврђена је тенденција да се лаким задацима означавају задаци који испитују ниво познавања чињеница, средње тешким, задаци који испитују разумевање градива, а тешким, задаци примене знања. Имајући то у виду, дистрибуција података у табели 78. очекивано је усаглашена са дистрибуцијом података у табели 77.

Табела 78. Укрштање форме и тежине задатака

Задаци	лаки		средње тежине		тешки		укупно	
	f	%	f	%	f	%	f	%
двочланог избора	4	88,9	0,5	11,1	0	0,0	4,5	100,0
вишечланог избора	29	51,8	19	33,9	8	14,3	56	100,0
допуњавања	19,5	45,9	14,5	34,1	8,5	20,0	42,5	100,0
са кратким одговорима	51,5	42,9	37,5	31,3	31	25,8	120	100,0
испуњавања налога	101	27,0	171,5	45,9	101	27,0	373,5	100,0
спаривања	12	60,0	5	25,0	3	15,0	20	100,0
сређивања	0	0,0	2	100,0	0	0,0	2	100,0
друге форме	16,5	47,8	11	31,9	7	20,3	34,5	100,0

У односу на указану тенденцију, приметна је разлика само у случају задатака испуњавања налога. Као што смо претходно видели, задаци испуњавања налога готово уједначено су заступљени на различитим нивоима знања, па би се очекивало да се слична дистрибуција понови и у случају расподеле задатака према тежини. Подаци, међутим, показују да су задаци испуњавања налога чешће сврставани у категорију задатака средње тежине, него у категорију лаких и тешких задатака.

Када је реч о заступљеност глагола у формулацији задатака различитих форми, код задатака двочланог избора доминира глагол „заокружити“ (100,0%), а исти глагол преовладава и у случају глагола вишечланог избора (53,6%). Код задатака допуњавања предњачи употреба глагола „допунити“ (42,9%). У случају задатака испуњавања налога најзаступљенији су глаголи „израчунати“ (24,2%), „одредити“ (20,6%), „написати“ (13,8), а потом следе

текстуални задаци (16,9%). Код задатака са кратким одговорима предњаче глаголи „одредити“ (38,9) и „написати“ (32,7%), а код задатака спаривања глагол „повезати“ (59,1%). Анализа је обухватила само два задатка сређивања, од којих се један састоји од глагола „обележити“, а други од глагола „поређати“. У случају задатака неке друге форме, најзаступљенији је глагол „подвући“ (66,7%).

Резултати показују да употреба глагола углавном одговара називу категорије задатка – формулација задатака бирања одговора најчешће обухвата глагол „заокружити“, формулација задатка допуњавања глагол „допунити“, формулација задатка спаривања глагол „повезати“ и тако даље. С једне стране, такав налаз може се посматрати као добар поступак у циљу постизања максималне директивности формулације задатка, а са друге стране може се тумачити као недостатак креативности у формулацијама задатака.

Разноврсност глагола употребљених у формулацији задатка испуњавања налога јесте значајно већа него у случају других форми задатака, што је донекле разумљиво ако се има у виду чињеница да је то најчешћа форма задатака у анализираним тестовима. Већа разноврсност формулација се може оправдати и чињеницом да се налог у задатку може свести на различите поступке: израчунавање, одређивање, записивање, решавање, набрајање, објашњење, конструисање и тако даље.

Начин утврђивања постигнућа на тесту знања

У наставку текста, приказаћемо и анализирати податке који се односе на различите аспекте утврђивања постигнућа на тесту знања. Реч је о следећим аспектима: израда кључа за оцењивање теста, бодовање појединачних задатака у зависности од форме и потпуности одговора, увођење негативних поена, утврђивање укупног скорa на тесту и превођење поена у оцену.

Када је реч о постојању кључа за оцењивање теста, установљено је да је већина интервјуисаних наставника (91%) припремила кључ за оцењивање анализираног теста. Укупно 7 наставника (9%) није израдило кључ. Распон

максималног броја поена који се може остварити на тесту креће се од 8 до 100 (табела 79). Дистрибуција података је доста разуђена, али се истиче податак да 20% анализираних тестова носи 100 поена. Анализом су обухваћени и тестови у којима нема бодова. Наиме, четири интервјуисана наставника не изводе поене којима бодују задатке, већ прегледани задатак обележавају одређеним ознакама. На пример, ознаком „+“ се обележава задатак који је у целини (или највећим делом) тачан. Ознаком „-“ се обележава задатак који је у потпуности (или највећим делом) нетачан. Делимично тачни задаци се обележавају тачком или ознаком „+ -“.

Табела 79. Максималан број поена који се може остварити на тесту

Број поена	Број тестова	
	f	%
0	4	5,0
8-12	10	12,5
13-17	7	8,8
18-22	12	15,0
23-27	6	7,5
28-32	1	1,3
33-37	4	5,0
38-42	4	5,0
43-47	3	3,8
48-52	2	2,5
53-57	1	1,3
58-62	5	6,3
63-67	3	3,8
73-77	1	1,3
83-87	1	1,3
100	16	20,0
укупно	80	100

У зависности од структуре и сложености задатка у тесту, просечан број поена по задатку у анализираним тестовима креће се од 1 до 34 поена (табела 80)

Табела 80. Просечан број поена по задатку³

Број поена	Број тестова	
	f	%
0-1	8	10,5
1-2	15	19,7
2-3	4	5,3
3-4	7	9,2
4-5	10	13,2
5-6	4	5,3
6-7	5	6,6
7-8	1	1,3
10-11	1	1,3
11-12	3	3,9
14-15	2	2,6
20-21	13	17,1
25	2	2,6
33-34	1	1,3
укупно	76	100,0

Иако је дистрибуција података доста разуђена, може се приметити да се три вредности издвајају по учесталости – просечна вредност поена између један и два (19,7%), 20-21 поена (17,1%) и 4-5 поена (13,2%).

Преко половине интервјуисаних наставника (55,0%) сваки задатак у тесту бодује и уколико је делимично тачан. То, ипак, значи да у анализираним тестовима постоје и задаци који се могу признати само уколико су у потпуности тачни (табела 81).

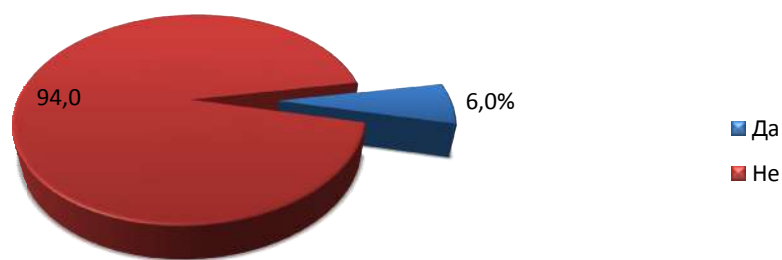
³ За сваки тест обухваћен анализом, утврђен је просечан број поена по задатку. Просечан број поена по задатку утврђен је тако што је укупан број поена који се на тесту може остварити подељен са укупним бројем задатака у тесту.

Табела 81. Број задатака на нивоу теста који се признају само уколико су у потпуности тачни

Број задатака	Број тестова	
	f	%
1	9	25,0
2	8	22,2
3	4	11,1
4	2	5,6
5	2	5,6
6	1	2,8
7	2	5,6
8	1	2,8
9	1	2,8
10	3	8,3
11	1	2,8
14	1	2,8
15	1	2,8
укупно	36	100,0

Како подаци показују, најчешћа околност коју смо сусретали у анализираним тестовима, јесте да тест садржи 1 или 2 задатка који се не могу признати као делимично тачни. Број задатака који у целини морају бити решени исправно креће се у делимично дисконтинуираном распону од 1 до 15.

Највећи број интервјуисаних наставника (94,0%) не предвиђа негативне поене на тесту (слика 11) што је у складу са одређеним теоријским становиштима (Taylor, Greer & Mussio, 1978; Буквић, 1996). Подсетимо се, задаци бирања одговора код којих се најчешће користе негативни поени (Petz, 1978; Мужич, 1968; Буквић, 1996) чине мање од 10% укупног броја анализираних задатака, што може бити додатно образложење зашто се негативни поени не срећу чешће у анализираним тестовима.



Слика 11. Заступљеност негативних поена на тесту

Укупан број поена на тесту већина наставника (90%) утврђује сумирањем свих поена оствареним на појединачним задацима. Уколико постоје негативни поени, њихова сума се одузима од укупног збира поена који су остварени давањем тачних одговора.

Табела 82. приказује просечан број поена за задатке различитих форми.⁴ Према приказаним подацима може се закључити да највише поена носе задаци испуњавања налога ($M = 7,1$), а најмање задаци двочланог избора ($M = 1,5$), сређивања ($M = 1,5$) и вишечланог избора ($M = 1,4$).

Табела 82. Просечан број поена по задатку одговарајуће форме

Форма	Просечан број поена
испуњавања налога	7,1
допуњавања	3,3
спаривања	2,7
са кратким одговорим	2,3
двочланог избора	1,5
сређивања	1,5
вишечланог избора	1,4
друге форме	6,8

⁴ Просечан број поена за задатке сваке појединачне форме, утврђен је тако што је укупан број поена за задатке одговарајуће форме на нивоу узрока тестова, подељен са укупним бројем задатака одговарајуће форме обухваћених анализом.

Следећи корак у анализи био је да се утврде гранични поени за сваку од школских оцена. Напоменимо да 10% испитаника није дало одговор на дато питање. Реч је о наставницима који су приложили примере тестова знања који се не оцењују.

Доњи праг за остваривање прелазне оцене креће се најчешће у распону од 25 до 40 % тачних одговора/решења на тесту. Најчесталије (41,7% одговора), доњи праг за остваривање прелазне оцене (2) јесте 30% тачних одговора/решења на тесту знања. Следећи одговор по учесталости је 40% (15,3% одговора), потом 25% (12,5% одговора) и 35% (11,1% одговора) (табела 83).

Табела 83. Доњи праг за довољну (2) оцену изражен у процентима тачних одговора/решења на тесту

Процент тачних одговора/решења	Број тестова	
	f	%
15	1	1,4
20	2	2,8
21	2	2,8
25	9	12,5
30	30	41,7
31	2	2,8
33	1	1,4
35	8	11,1
36	1	1,4
40	11	15,3
45	1	1,4
50	2	2,8
55	1	1,4
60	1	1,4
укупно	72	100,0

Резултати показују да је 50-55% тачних одговора/решења на тесту најчешће довољно за постизање добре оцене. Нешто више од половине испитаника (54,2%) сагласно је да је 50% тачних одговора/решења на тесту знања довољно за оцену добар (3). Следећи одговор по учесталости је 55%

тачних одговора/решења у тесту знања, док су остали одговори заступљени у мање од 10% случајева (табела 84).

Табела 84. Доњи праг за добру (3) оцену изражен у процентима тачних одговора/решења на тесту

Процент тачних одговора/решења	Број тестова	
	f	%
25	1	1,4
35	1	1,4
40	5	6,9
41	1	1,4
45	3	4,2
50	39	54,2
51	4	5,6
55	11	15,3
60	3	4,2
65	3	4,2
70	1	1,4
укупно	72	100,0

За остваривање врло добре оцене, најчешће (65,3% одговора) је потребно 65-70% тачних одговора/решења на тесту (табела 85)

Табела 85. Доњи праг за врло добру (4) оцену изражен у процентима тачних одговора/решења на тесту

Процент тачних одговора/решења	Број тестова	
	f	%
60	5	6,9
61	1	1,4
64	1	1,4
65	9	12,5
68	2	2,8
70	36	50,0
71	4	5,6
75	11	15,3
80	2	2,8
85	1	1,4
укупно	72	100,0

Како подаци показују (табела 86), доњи праг за остваривање одличне оцене (5) на тесту најчешће се креће у распону од 85 до 90% тачних одговора/решења на тесту, при чему је најчесталији одговор 90% (50,0% одговора).

Табела 86. Доњи праг за одличну (5) оцену изражен у процентима тачних одговора/решења на тесту

Процент тачних одговора/решења	Број тестова	
	f	%
80	9	12,5
81	1	1,4
85	18	25,0
86	1	1,4
90	36	50,0
91	4	5,6
95	3	4,2
укупно	72	100,0

Распон поена за довољну и добру оцену (45 јединица у оба случаја) већи је од распона за врло добру (25 јединица) и одличну оцену, при чему је најужи распон поена за највишу оцену (15 јединица). Највише разлика постоји у одговорима наставника у вези са доњим прагом за пролазну оцену – распон се креће од 15 до 60% тачних одговора/решења на тесту.

Провера квалитета теста знања

Пре непосредне примене теста знања, наставници још увек немају јасну представу о квалитету теста. Када за то постоје могућности, квалитет теста се може проверити пробним испитивањем или поступком ревизије теста од

стране искуснијих колега. То је посебно значајно уколико су сви задаци које је наставник саставио потпуно нови.

Већина испитаника (73,8%) барем једном током школске године израђује потпуно нов тест знања. Укупно 26,2% наставника у току школске године не израђује у потпуности нови тест знања, већ се у мањој или већој мери ослања на тестове који су израђени претходних година. Када је реч о потпуно новим тестовима знања, значајан број наставника (59,3%) израђује их скоро сваког месеца (табела 87).

Табела 87. Учесталост израде новог теста знања

Учесталост	f	%
једанпут годишње	1	1,7
два пута годишње	3	5,1
четири пута годишње	9	15,3
скоро сваког месеца	35	59,3
једном месечно	7	11,9
увек када правим тест	4	6,8
укупно	59	100,0

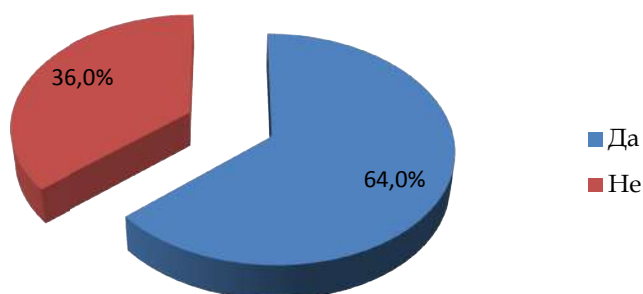
Мали број наставника (6,3%) у одговору на питање у вези са учесталошћу израде тестова знања, понудио је одговоре који су сврстани у категорију „осталих“ одговора. Реч је о следећим одговорима: *„Нови тест израђујем једном недељно, јер волим да израђујем петнаестоминутне тестове“*; *„Наизглед, тест је увек нови. У суштини, сви задаци су из претходних генерација, само направим нову комбинацију“*; *„Не могу да дам прецизан одговор. Колико често израђујем нови тест зависи од генерације – некада могу да искористим доста тога из претходних школских година, а некада све радим испочетка“*; *„Тек сам почела да радим, тако да стицајем околности сваки пут израђујем нови тест“*.

Уколико се у изради теста наставници ослањају на старе тестове, заступљеност раније израђених задатака у тестовима који се касније састављају креће се од 50% до 80% (табела 88).

Табела 88. Заступљеност раније израђених задатака у новим тестовима знања

Заступљеност	f	%
до 50%	5	31,3
до 60%	3	18,8
до 70%	5	31,3
до 80%	3	18,8
укупно	16	100,0

Већина наставника (64,0%) обухваћених испитивањем проверила је квалитет теста пре његове примене у редовном наставном раду (слика 12).



Слика 12. Провера квалитета теста пре финалне примене

Доминантни циљеви провере квалитета теста јесу провера тежине задатака (56,3%) и формулације задатака (32,5%). Мањи број наставника (6,3%) известио је да провера квалитета теста омогућава усаглашавање система бодовања међу наставницима. Остале категорије одговора ређе су заступљене (табела 89). У категорију „на неки други начин“ сврстани су одговори као што су: „Важно ми је да утврдим колико је градиво обухваћено тестом ученицима потребно у животу“ и „Због ученика са посебним потребама, увек консултујем стручног сарадника“.

Табела 89. Циљ провере квалитета теста

Сврха провере	f	%
провера тежине задатака	45	56,3
провера формулације задатака	26	32,5
усклађивање система бодовања међу колегама	5	6,3
процена колико тест оставља простора за креативно изражавање	3	3,8
процена усклађености захтева на тесту са захтевима Мале матуре	2	2,5
процена адекватности дужине теста	2	2,5
провера јасноће упутства	2	2,5
нешто друго	4	5,0

N = 80

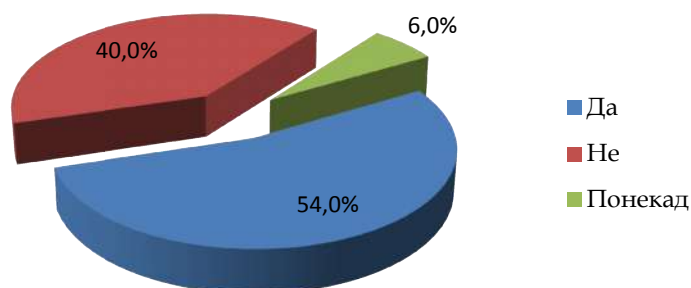
Наставници који су известили да су проверили квалитет теста, имали су могућност да укажу на начин/е како је провера реализована. Установили смо да су консултације са другим колегама уобичајени начин провере квалитета теста. Сва остала решења значајно су ређа (табела 90).

Табела 90. Начин провере квалитета теста

Начин	f	%
консултовањем са другим наставницима	50	87,7
консултовањем са стручним сарадником у школи	2	3,5
спровођењем пробног испитивања	2	3,5
на неки други начин	3	5,3
укупно	57	100,0

Непосредном применом теста и анализом постигнућа које ученици остварују на тесту добија се повратна информација о квалитету задатака садржаних у тесту. Како је показало истраживање Ал-Шара (Al-Shara'N, 2013), у домену анализе резултата добијених тестом наставици показују најмање компетенција када је реч о изради и примени тестова знања. Занимало нас је каква су искуства у том домену када је реч о нашим наставницима. Слика 13.

приказујемо одговоре испитаника на питање да ли анализирају постигнуће ученика на нивоу појединачних задатака.



Слика 13. Анализа постигнућа ученика на нивоу појединачних задатака у тесту

Дакле, око 60,0% испитаника анализира постигнуће ученика на нивоу појединачних задатака, увек или понекад. Анализа се најчешће спроводи са циљем планирања даљег рада у настави (79,2%). Неки од одговора који илуструју такву тежњу су: „Анализу радим да би се на часу појаснило оно што је стварало тешкоће“; „Све испод 50% успешности евидентирам и потрудим се да на то обратим пажњу“; „Ако су били лоши на неком задатку, поново радимо тај задатак на часу или на допунској“; „Тежак задатак решавамо на часу, а тај задатак следеће године дајем на пробном тесту“; „Ако је тежак задатак на напредном нивоу, вежбамо га на додатној. Ако је лоше урађен задатак на основном нивоу, вежбамо га на допунској и после поново имамо тест да ученици могу да поправе оцену“; „Пратим шта им је било тешко/лако. За следећу годину ми је то битно.“

Ипак, 40% наставника не анализира постигнуће ученика на нивоу појединачних задатака, умањујући тиме могућност да сазнају нешто више о квалитету самог теста и да пораде на његовом унапређивању. Неки од коментара који су пратили одговор да се анализа не ради били су: „Усмено прокоментаришем који је задатак тежак“; „Ученику се усмено скрене пажња у чему је мање успешан“; „Запамтим шта су лоше урадили“; „План то не предвиђа“.

Како подаци у табели 91. показују, мање од четвртине наставника користи резултате анализе постигнућа ученика на нивоу појединачних задатака у циљу унапређивања квалитета самог теста знања. Такав податак

упућује на закључак да се тест најчешће посматра као „добар“, односно да се постигнуће ученика на тесту посматра независно од квалитета самог теста.

Табела 91. Циљеви анализе постигнућа ученика на нивоу појединачних задатака

Циљеви анализе	f	%
планирање наставног рада	42	87,5
унапређивање квалитета задатака	11	22,9
вредновање сопственог рада	6	12,5
праћење напредовања ученика	2	4,2
нешто друго	3	6,3

N = 48

Мали број наставника је у вези са циљем анализе понудио одговоре које није било могуће сврстати у преостале издвојене категорије. Реч је о одговорима: *„Детаљна анализа се ради само на иницијалном тесту“*; *„Анализу радим само када им задатам тест на рачунару“* и *„(...) да видим ко је најбољи, да подстакнем такмичарски дух код ученика“*.

Може се очекивати да наставници чувају раније коришћене тестове знања, пре свега зато што их, у одговарајућим околностима, могу поново употребити, мењати и дорађивати. У томе им данас може помоћи употреба рачунара. Резултати интервјуа показали су да свега два наставника српског језика и математике не израђују базу задатака. Преосталих 78 испитаника то чини. Од тог броја, 79,5% наставника формира електронску базу тестова, а 20,5% то не чини. Укупно 71,8% наставника извештава да чува штампане верзије теста, док се 28,2% наставника не опредељује за то.

Већина наставника (92,5%) проналази начине да осигура безбедност тестова. Ипак, поједини наставници (7,5%) омогућавају ученицима да задрже оцењене тестове. Троје наставника је изјавило да ученици могу трајно да задрже тестове, док троје наставника практикује да ученицима уступе тестове како би родитељи стекли увид у њих, али да родитељи у одређеном тренутку ипак врате тестове наставнику.

Када је реч о конкретним начинима обезбеђивања тестова, установљено је да већина наставника (82,4%) то чини чувањем тестова на персоналном

рачунару. Следећи одговор по учесталости је чување штампане верзије тестова у школи, али у простору који се закључава (54,1%), односно чување штампане верзије тестова ван школе (27,0%) (табела 92).

Табела 92. Начини осигуравања базе тестова од евентуалних злоупотреба

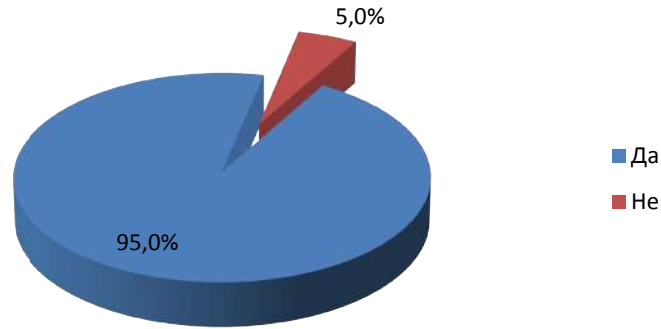
Начини	f	%
чувањем електронске верзије теста на персоналном рачунару	61	82,4
чувањем штампане верзије тестова у школи, али у простору који се закључава	40	54,1
чувањем штампане верзије тестова ван школе	20	27,0
чувањем електронске верзије теста на заједничком рачунару под шифром	5	6,8
остали начини	2	2,7

N = 74

Најређе заступљена категорија одговора је чување тестова на рачунару у школи (6,8%), док је двоје наставника известило о „осталим“ начинима поступања са тестовима („Све чувам у роковнику“; „Тестове чувам на USB-у који носим са собом“).

Припрема за тестирање и реализација тестирања у настави

У циљу стандардизације поступка тестирања, неопходно је организовати припрему ученика за тестирање. Испитивањем је утврђено да 95,0% наставника припрема ученике за тестирање, а само 5,0% то не чини (слика 14).



Слика 14. Припрема ученика за тестирање

Како показују подаци у табели 93, наставници најчешће планирају један (40,8%) или један до два часа припреме за тестирање (42,1%). Знатно ређе, припреми за тест посвећују се два или више часова.

Табела 93. Број часова припреме за тестирање

Број часова	f	%
1	31	40,8
1-2	32	42,1
2	6	7,9
2-3	5	6,6
3	1	1,3
5-6	1	1,3
укупно	76	100,0

Испитивањем је установљено да већина наставника (56,3%) на часовима припреме ученицима указује на садржај градива који ће бити обухваћен тестом. На такав закључак упућују следећи одговори: „Увек имамо један до два часа вежбања за тест. Ученици могу да виде сличне примере задатака, никада идентичне онима на тесту, али им је јасно шта ће се тражити“; „На часу задајем пробни тест који не оцењујем (мада ученици могу међусобно да се оцене), али је важно да обухвата градиво које се испитује на тесту. Такође, обнављамо све што није јасно и на допунској“; „Најавим наставне теме које ће тест обухватити. Упознају се са тим у ком смислу ће изгледати тест знања. Није нагласак на упознавању типова задатака

који ће се наћи на тесту колико на томе да раде задатке који их подстичу на размишљање.“

Уколико се наставници определиле да на часу припреме ученике упознају са формама задатака или начинима њиховог решавања, поступају на следеће начине: „Задаци на часу припреме су по типу 100% идентични онима на тесту знања“; „Све буде исто као на часу припреме само бројеве променим“; „На тесту променим само ред речи“; „Поделим им штампани материјал са решењима - то су слични задаци онима на тесту. На том материјалу радимо на часу или им дам за рад код куће“. Описани начин рада карактеристичан је за нешто више од трећине (38,8%) испитаника.

Како је раније истакнуто, ученици који се припремају за тест остварују боље резултате (Бјекић и Папић, 2005; Савовић и сар, 2006). Иако су тога свесни, наставници немају могућности да припреми посвете већи број часова, делимично зато што „час припреме за тест није предвиђен планом и програмом“. Припрема ученика са тестирање доприноси бољим условима стандардизације тестирања. Уколико су ученици упознати са садржајем и формом задатака који ће чинити садржај теста, ефикасније ће решавати задатке – неће се задржавати у решавању задатака чији садржај им није познат, или у тумачењу начина на који је потребно дати одговор на постављени задатак. Самим тим, процена ученичког знања биће поузданија.

Важан аспект стандардизације услова тестирања јесте просторно-временска организација. Истраживањем је установљено да просторна организација тестирања у школама обухваћеним испитивањем није нарочито наглашена. Наиме, у ситуацији тестирања наставници углавном не размештају ученике, већ ученици седе на својим уобичајеним местима. Мањи број наставника размешта ученике у свакој ситуацији тестирања (22,5%), односно понекад (18,8%) (табела 94).

Табела 94. Распоред седења у току тестирања

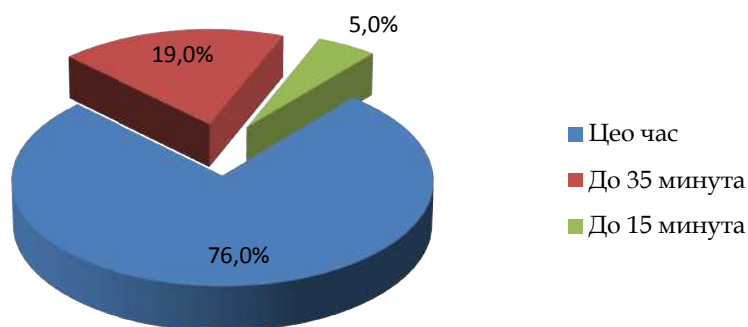
Распоред седења	f	%
не размештам ученике	47	58,8
понекад размештам ученике	15	18,8
увек размештам ученике	18	22,5
укупно	80	100,0

За потребе тестирања, већина наставника (80,0%) припрема две паралелне форме теста, како би се обезбедили минимални услови за поуздану проверу ученичког знања. Максималан број група се креће у распону од једне до 8 група (табела 95).

Табела 95. Максималан број група за рад на тесту

Број група	f	%
1	5	6,3
2	48	60,0
3	6	7,5
4	18	22,5
5	1	1,3
6	1	1,3
8	1	1,3
укупно	80	100,0

Када је реч о временској организацији тестирања, установили смо да решавање теста има различито трајање, у зависности од врсте теста. На пример, решавање иницијалног теста може бити ограничено на 30 минута, док решавање полугодишњег теста обично траје 45 минута. На основу одговора наставника, формиране су три категорије, а учесталост њиховог јављања упућује на закључак да решавање теста знања најчешће траје један школски час (слика 15).



Слика 15. Време предвиђено за решавање теста

Испитивање показује да наставници најчешће (76,3%) дозвољавају да ученик преда тест онда када заврши са радом (табела 96). Мањи број наставника (16,3%) захтева да ученик сачека крај часа, а да преостало време искористи да преконтролише своја решења, или да се позабави неким садржајем који не омета рад других ученика.

Табела 96. Предаја попуњеног теста

Тренутак у току часа у коме ученик предаје тест	f	%
када заврши са радом	61	76,3
на крају часа	13	16,3
када заврши са радом, али не пре истека половине времена предвиђеног за рад	6	7,5
укупно	80	100,0

Руднер и Шафер (Rudner & Schafer, 2002) истичу да је у датим околностима провера корисна за ученике, али јој ученици нерадо приступају. Како би читав поступак провере био мање стресан, пре почетка тестирања потребно је усмерити ученике да не проверавају сва решења, већ само она око којих имају дилеме.

Пре непосредне реализације тестирања, наставник доноси одлуку и у вези са пружањем могућности ученицима да преправљају решења/одговоре на тесту. Коришћење хемијске оловке обично асоцира на немогућност да се једном уписано решење/одговор исправи. Испитивање показује да у ситуацији

решавања теста, наставници најчешће не захтевају да се тест решава одређеним средством за писање (35% одговора), мада је присутна тенденција да се предност да хемијској оловци над графитном оловком (табела 97). Неколико интервјуисаних наставника је нагласило да се завршни тест (Мала матура) решава хемијском оловком и да је потребно навикавати ученике на дату околност кроз тестирања у редовној настави.

Табела 97. Дозвољена средства за писање у ситуацији тестирања

Средства за писање	f	%
свеједно је чиме ученик пише	28	35,0
хемијска оловка	22	27,5
графитна оловка	17	21,3
ученик може писати графитном оловком, али коначан одговор мора подебљати хемијском оловком	13	16,3
укупно	80	100

У току решавања теста, нешто мање од половине наставника (40,0%) дозвољава да ученици затраже било коју врсту помоћи (указивањем на начин саопштавања одговора или решења, област градива на коју се задатак односи и слично) односно да затраже помоћ која се односи искључиво на разјашњење поставке задатака (41,3%). Мањи број наставника (18,8%) ученицима не оставља могућност да затраже помоћ. Као што је раније истакнуто, након што започне решавање теста, било какво обраћање ученицима није препоручљиво (Taylor, Greer & Mussio, 1978; Мужих, 1986). Чињеница да је значајан број наших испитаника склон да ученицима указује помоћ у току решавања теста, може се довести у везу са недостатком одговарајућих упутстава за рад, јер 36,3% наставника пре почетка теста ученицима не пружа инструкције за решавање одговарајућих форми задатака.

Сличности и разлике између методолошких карактеристика тестова знања у настави српског језика и математике

С обзиром на специфичности школских предмета српског језика и математике, пошли смо од претпоставке да ће се методолошке карактеристике тестова израђених за потребе наставе значајно разликовати. Испитивањем је регистрован извештај број разлика које ћемо приказати и анализирати у наставку текста.

Најпре, испитивањем је утврђено да су наставници математике идентификовали више фаза у конструицији теста него наставници српског језика. Разлика у одговорима је статистички значајна на нивоу 0,05 (табела 98).

Табела 98. Предмет који наставник предаје и укупан број фаза конструиције теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	3,58	1,41	0,22
математика	40	4,85	1,73	0,27

$$t = - 3,06; df = 74,9; p = 0.01$$

Када је реч о одређивању дужине теста као једне од конструиционих фаза, установљено је да постоји статистички значајна разлика у одговорима наставника српског језика и математике. Наиме, наставници математике (86,7%) чешће од наставника српског језика (13,3%) одређивање дужине теста препознају као фазу конструиције теста у свом раду (табела 99). Статистичка значајност разлике у одговорима наставника постоји на нивоу 0,05 ($\chi^2 = 9,93$; $df = 1$; $p = 0.02$; $C = 0,33$).

Табела 99. Предмет који наставник предаје и одређивање дужине теста као фаза конструкције теста знања

Предмет	Одређивање дужине теста					
	не		да		укупно	
	f	%	f	%	f	%
српски језик	38	58,5	2	13,3	40	50,0
математика	27	41,5	13	86,7	40	50,0
укупно	65	100,0	15	100,0	80	100,0

Слично, установљено је да постоји статистички значајна разлика у одговорима наставника српског језика и математике и када је реч о изради кључа за оцењивање теста. Наиме, наставници математике (67,4%) чешће од наставника српског језика (32,6%) израду кључа препознају као фазу конструкције теста у свом раду (табела 100). Статистичка значајност разлике у одговорима наставника постоји на нивоу 0,05 ($\chi^2 = 11,31$; $df = 1$; $p = 0,01$; $C = 0,35$).

Табела 100. Предмет који наставник предаје и израда кључа за оцењивање теста знања као фаза конструкције теста знања

Предмет	Израда кључа за оцењивање теста знања					
	не		да		укупно	
	f	%	f	%	f	%
српски језик	26	70,3	14	32,6	40	50,0
математика	11	29,7	29	67,4	40	50,0
укупно	37	100,0	43	100,0	80	100,0

Препознавање конструкционих фаза које нису обухваћене предложеним теоријским моделом, означених као „остале фазе“, такође разликује начин рада наставника српског језика и математике. Установљено је да наставници математике (72,7%) чешће од наставника српског језика (27,3%) извештавају о постојању „осталих“ фаза у конструкцији теста (табела 101). Статистичка

значајност разлике у одговорима наставника постоји на нивоу 0,05 ($\chi^2 = 11,61$; $df = 1$; $p = 0.01$; $C = 0,36$).

Табела 101. Предмет који наставник предаје и заступљеност осталих фаза конструкције теста знања

Предмет	Заступљеност осталих фаза конструкције теста знања					
	не		да		укупно	
	f	%	f	%	f	%
српски језик	31	66,0	9	27,3	40	50,0
математика	16	34,0	24	72,7	40	50,0
укупно	47	100,0	33	100,0	80	100,0

Даља анализа је показала да тестови из српског језика садрже значајно већи број задатака у односу на тестове из математике. Разлика је статистички значајна на нивоу 0,05 (табела 102).

Табела 102. Предмет који наставник предаје и укупан број задатака у тесту

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	9,55	3,76	0.59
математика	40	6,78	3,05	0.48

$$t = 3,628; df = 74,9; p = 0.01$$

Статистички значајна разлика у одговорима наставника утврђена је када је реч о садржају уџбеника као основи за разраду теста знања. Наиме, наставници српског језика (75,0%) се више ослањају на садржај уџбеника него наставници математике (25,0%), а разлика у одговорима је статистички значајна на нивоу 0,05 ($\chi^2 = 10,77$; $df = 1$; $p = 0.00$; $C = 0,34$; Прилог 3, табела 15).

Тестови из математике најчешће само започињу лаким задатком (37,5%), а према учесталости следе тестови који у целини прате модел који се препоручује у литератури – тест почиње лаким задатком, наставља се средње тешким задацима, а завршава се тешким задатком (32,5%). Преостали тестови (30,0%) већ са првим задатком одступају од препорученог принципа. Тестови из српског језика најчешће само почињу лаким задатком (60,0%). Преостали анализирани тестови (32,5%) већ са првим задатком одступају од препорученог принципа. У односу на тестове из математике, код тестова из српског језика ређе је заступљен модел уређења задатака који се препоручује у литератури (7,5%)

На основу упоредне анализе тежине задатка и нивоа знања који се задатком испитује, регистрована је тенденција да се задаци који испитују ниво познавања чињеница означавају лаким, задаци који испитују ниво разумевања градива – средње тешким, а задаци који испитују примену знања – тешким задацима. Када је реч о тестовима из српског језика, укупно 133 задатка (35%) имала су другачију тежину него што су наставници иницијално очекивали. Од тог броја, 55% задатака показало се „тежим“ него што су то наставници очекивали. На пример, наставници извештавају да су поједини задаци који утврђују ниво познавања чињеница за већину ученика били средње тешки, иако се очекивало да ће то за већину ученика бити лаки задаци. Истовремено, 45% задатака се показало „лакшим“ него што се очекивало. У случају тестова из математике, регистровано је 69 (26%) задатка са наведеним одступањем. Од тог броја, 54% таквих задатака показало се, условно речено, „тежим“ него што су то наставници очекивали, а 46% „лакшим“. Занимљив је податак да је приближно 6% испитаника било неодлучно када је требало да одреде ниво знања који задатак мери, а свега 1% наставника је имало дилему код процене тежине задатка.

Тестови из српског језика су разноврснији према броју форми задатака које садрже (наставници српског језика – $M = 3,73$; наставници математике – $M = 1,58$; $t = 7,63$; $df = 64,7$; $p = 0.00$). Уколико се упореди заступљеност задатака одговарајуће форме у тестовима знања из математике и српског језика, може се

закључити да тестови из српског језика имају значајно више задатака: вишеструког избора (наставници српског језика - $M = 1,11$; наставници математике - $M = 0,29$; $t = 2,48$; $df = 56,1$; $p = 0.02$), допуњавања (наставници српског језика - $M = 0,85$; наставници математике - $M = 0,20$; $t = 3,40$; $df = 72,3$; $p = 0.00$), са кратким одговорима (наставници српског језика - $M = 2,65$; наставници математике - $M = 0,34$; $t = 5,51$; $df = 49,3$; $p = 0.00$), спаривања (наставници српског језика - $M = 0,45$; наставници математике - $M = 0,05$; $t = 3,55$; $df = 47,2$; $p = 0.00$) и „других“ форми (наставници српског језика - $M = 0,81$; наставници математике - $M = 0,05$; $t = 4,95$; $df = 43,3$; $p = 0.00$) (Прилог 3, табеле 16-21).

Када је реч о тестовима знања из српског језика, задаци вишечланог избора најчешће имају три понуђена одговора од којих је један тачан (регистровано је 12 таквих задатака). Следећи по учесталости су задаци у којима су понуђена 4 одговора од којих је један тачан (регистровано је 11 таквих задатака). Регистрована су 4 задатка у којима су од 5 понуђених одговора два тачна, односно три задатка са 5 понуђених одговора од којих су три тачна. Преостали, ређе заступљени задаци имају између 6 и 10 понуђених алтернатива, а број тачних одговора се креће у распону од 4 до 6. Када је реч о тестовима знања из математике, регистровано је 9 задатака вишечланог избора у којима су понуђене четири алтернативе од којих је једна тачан одговор. Преостали, ређе заступљени задаци имају између 3 или 5 понуђених алтернатива, а број тачних одговора се креће у распону од 2 до 3.

У тестовима знања из српског језика, заступљено је укупно 17 задатака спаривања. У анализираним тестовима из српског језика, задаци спаривања имају два низа понуђених информација. У сваком од низова понуђено је између две и 12 алтернатива које је потребно довести у везу. Број тачних одговора варира у распону од три до 12. Форма која се појављује нешто чешће од других (регистрована су четири таква задатка) је форма у којој су понуђене четири премисе и четири одговора, и предвиђена су четири тачна одговора. Све остале комбинације су заступљене са фреквенцијом 1 или 2. Интересантно је да од 17 задатака спаривања регистрованих у тестовима из српског језика, у

случају 6 задатака број алтернатива прелази границу препоручену у литератури (Frey *et al.*, 2005). У тестовима знања из математике, заступљена су само два задатка спаривања од којих један има четири премисе и четири понуђена одговора, а предвиђена су четири тачна одговора. Други задатак има четири премисе и 5 одговора, а предвиђена су четири тачна одговора.

Када је реч о нивоима знања који се задацима мере и тежини задатака, анализа је показала да је, у односу на тестове из математике, у тестовима из српског језика значајно више задатака који испитују: ниво познавања чињеница (наставници српског језика – $M = 3,93$; наставници математике – $M = 2,52$; $t = 2,88$; $df = 61,0$; $p < 0.01$), односно више лаких (наставници српског језика – $M = 3,54$; наставници математике – $M = 2,31$; $t = 3,02$; $df = 74,3$; $p = 0.00$) и средње тешких задатака (наставници српског језика – $M = 3,72$; наставници математике – $M = 2,78$; $t = 2,14$; $df = 74,6$; $p = 0.04$) (Прилог 3, табеле 22-24).

У тестовима из српског језика, у односу на тестове из математике, значајно је више (Прилог 3, табеле 24-45):

- задатака допуњавања који испитују ниво познавања чињеница (наставници српског језика – $M = 0,48$; наставници математике – $M = 0,13$; $t = 2,86$; $df = 75,7$; $p < 0.01$) и ниво примене знања (наставници српског језика – $M = 0,15$; наставници математике – $M = 0,00$; $t = 2,09$; $df = 39,0$; $p = 0.04$).
- задатака са кратким одговорима који испитују ниво познавања чињеница (наставници српског језика – $M = 0,90$; наставници математике – $M = 0,29$; $t = 2,70$; $df = 65,1$; $p < 0.01$), разумевања градива (наставници српског језика – $M = 1,00$; наставници математике – $M = 0,05$; $t = 4,82$; $df = 41,6$; $p = 0.00$) и примене знања (наставници српског језика – $M = 0,71$; наставници математике – $M = 0,00$; $t = 4,38$; $df = 39,0$; $p = 0.00$)
- задатака спаривања који испитују ниво познавања чињеница (наставници српског језика – $M = 0,33$; наставници математике – $M = 0,03$; $t = 3,00$; $df = 44,1$; $p < 0.01$), и ниво примене знања (наставници српског језика – $M = 0,10$; наставници математике – $M = 0,00$; $t = 2,08$; $df = 39,0$; $p = 0.04$)
- задатака вишеструког избора који испитују ниво разумевања градива (наставници српског језика – $M = 0,36$; наставници математике – $M = 0,03$; $t =$

3,44; $df = 44,6$; $p = 0.00$) и примене знања (наставници српског језика - $M = 0,20$; наставници математике - $M = 0,00$; $t = 2,45$; $df = 39,0$; $p = 0.02$)

- задатака „других“ форми који испитују ниво познавања чињеница (наставници српског језика - $M = 0,43$; наставници математике - $M = 0,00$; $t = 3,78$; $df = 39,0$; $p = 0.00$) и ниво примене знања (наставници српског језика - $M = 0,23$; наставници математике - $M = 0,00$; $t = 3,46$; $df = 39,0$; $p = 0.00$)
- лаких задатака који имају форму задатака допуњавања (наставници српског језика - $M = 0,39$; наставници математике - $M = 0,10$; $t = 2,49$; $df = 64,2$; $p = 0.02$), са кратким одговорима (наставници српског језика - $M = 0,93$; наставници математике - $M = 0,31$; $t = 2,63$; $df = 73,9$; $p = 0.01$), спаривања (наставници српског језика - $M = 0,33$; наставници математике - $M = 0,03$; $t = 2,81$; $df = 43,5$; $p < 0.01$) и задатака „друге“ форме (наставници српског језика - $M = 0,38$; наставници математике - $M = 0,03$; $t = 2,92$; $df = 42,6$; $p < 0.01$)
- средње тешких задатака који имају форму задатака вишеструког избора (наставници српског језика - $M = 0,44$; наставници математике - $M = 0,04$; $t = 3,24$; $df = 43,1$; $p = 0.00$), допуњавања (наставници српског језика - $M = 0,27$; наставници математике - $M = 0,08$; $t = 2,00$; $df = 56,1$; $p < 0.05$), са кратким одговорима (наставници српског језика - $M = 0,90$; наставници математике - $M = 0,00$; $t = 4,71$; $df = 39,0$; $p = 0.00$) и задатака „друге“ форме (наставници српског језика - $M = 0,26$; наставници математике - $M = 0,03$; $t = 2,85$; $df = 47,0$; $p < 0.01$)
- тешких задатака који имају форму задатака вишеструког избора (наставници српског језика - $M = 0,21$; наставници математике - $M = 0,00$; $t = 2,23$; $df = 39,0$; $p = 0.03$) и форму задатака са кратким одговорима (наставници српског језика - $M = 0,73$; наставници математике - $M = 0,03$; $t = 3,87$; $df = 40,5$; $p = 0.00$).

У тестовима из математике, у односу на тестове из српског језика, значајно је више (Прилог 3, табеле 46-52):

- задатака испуњавања налога (наставници српског језика - $M = 3,47$; наставници математике - $M = 5,75$; $t = -3,92$; $df = 78$; $p = 0.00$)

- задатака испуњавања налога који испитују ниво познавања чињеница (наставници српског језика - $M = 1,11$; наставници математике - $M = 1,79$; $t = - 2,46$; $df = 78$; $p = 0.02$), разумевања (наставници српског језика - $M = 1,38$; наставници математике - $M = 2,21$; $t = - 2,58$; $df = 67,4$; $p = 0.01$) и примене знања (наставници српског језика - $M = 1,13$; наставници математике - $M = 1,77$; $t = - 2,31$; $df = 75,2$; $p = 0.02$).
- лаких задатака који имају форму задатака испуњавања налога (наставници српског језика - $M = 0,88$; наставници математике - $M = 1,55$; $t = - 2,61$; $df = 78$; $p = 0.01$)
- средње тешких задатака који имају форму задатака испуњавања налога (наставници српског језика - $M = 1,82$; наставници математике - $M = 2,61$; $t = - 2,03$; $df = 78$; $p < 0.05$)
- тешких задатака који имају форму задатака испуњавања налога (наставници српског језика - $M = 1,00$; наставници математике - $M = 1,62$; $t = - 2,30$; $df = 67,4$; $p = 0.03$)

Околност да задатак носи, у просеку, један поен карактеристична је за анализиране тестове из српског језика, као и за полугодишње и годишње тестове који се конструишу по узору на завршни испит (где сваки задатак носи по један поен). Околност да задатак носи, у просеку, 20 поена карактеристична је за анализиране тестове из математике.

Даљом анализом, установљено је да је просечан број поена по задатку статистички значајно већи у тестовима из математике у односу на тестове из српског језика. Разлика је статистички значајна на нивоу 0,01 (тестови из српског језика - $M = 4,44$; тестови из математике - $M = 10,52$; $t = - 3,73$; $df = 78$; $p = 0.00$; Прилог 3, табела 53). Такође, максималан број поена који се на тесту може остварити већи је у случају тестова знања из математике него у случају тестова из српског језика. Разлика у броју поена је статистички значајна на нивоу 0,05 (тестови из српског језика - $M = 34,08$; тестови из математике - $M = 52,10$; $t = - 2,44$; $df = 78$; $p = 0.02$; Прилог 3, табела 54).

Испитивањем је регистровано да постоји статистички значајна разлика у одговорима наставника српског језика и математике када је реч о доњем прагу

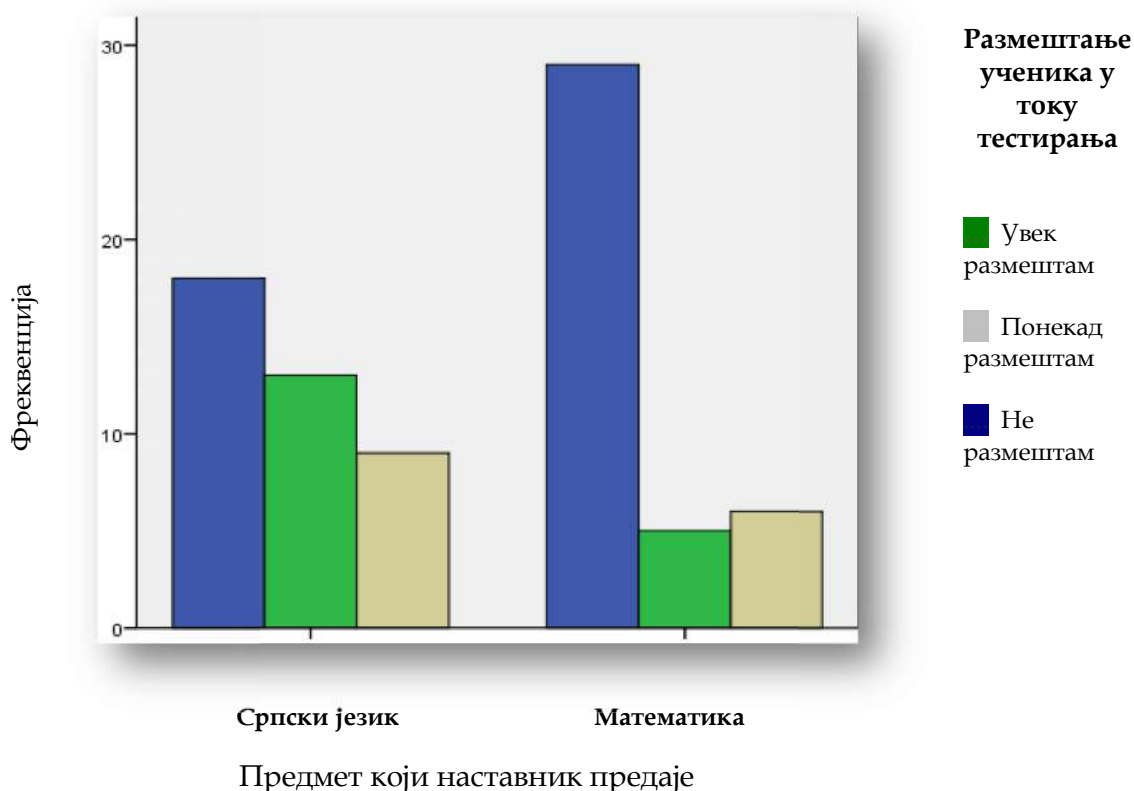
за пролазну оцену. Наиме, праг пролазности је виши у случају тестова из српског језика, а разлика у одговорима наставника је статистички значајна на нивоу 0,01 (тестови из српског језика – $M = 32,25$; тестови из математике – $M = 26,08$; $t = 2,92$; $df = 65,4$; $p < 0.01$; Прилог 3, табела 55).

Даља статистичка анализа показала је да су наставници српског језика склони да више задатака на нивоу теста признају као делимично тачне у односу на наставнике математике. Разлика у одговорима је статистички значајна на нивоу 0,05 (тестови из српског језика – $M = 6,75$; тестови из математике – $M = 5,08$; $t = 2,42$; $df = 78$; $p = 0.02$; Прилог 3, табела 56).

Даљом анализом је установљено да наставници математике израђују већи максималан број паралелних форми теста у односу на наставнике српског језика (тестови из српског језика – $M = 2,35$; тестови из математике – $M = 2,90$; $t = -2,10$; $df = 78$; $p = 0.04$; Прилог 3, табела 57).

Када су у питању разматрани аспекти реализације тестирања, статистички значајна разлика у одговорима наставника српског језика и математике установљена је у односу на саопштавање упутства за решавање теста и распоред седења. Наиме, статистички значајно чешће од наставника математике, наставници српског дају упутства за решавање теста ($\chi^2 = 9,14$; $df = 1$; $p = 0.00$; $C = 0,32$; Прилог 3, табела 58). У односу на наставнике математике, наставници српског језика ређе су бирали одговор „у ситуацији тестирања не размештам ученике“ ($\chi^2 = 6,73$; $df = 2$; $p = 0.04$; $C = 0,28$) (слика 16).

Када је реч о провери квалитета теста пре његове финалне примене, установљено је да статистички значајна разлика у одговорима наставника српског језика и математике постоји само када је реч о испитивању квалитета теста са циљем провере тежине задатака. Наиме, установљено је да су наставници математике чешће (67,5%) извештавали о испитивању теста са циљем провере тежине задатака него наставници српског језика (45%). Разлика у одговорима је статистички значајна на нивоу 0,05 ($\chi^2 = 4,11$; $df = 1$; $p = 0.04$; $C = 0,22$; Прилог 3, табела 60).



Слика 16. Предмет који наставник предаје и размештање ученика у току тестирања

У формулацијама задатака из српског језика најзаступљенији су глаголи: „написати“ (23,7%), „одредити“ (21,8%) и „подвући“ (14,1%). У случају математичких задатака, највише се користи глагол „израчунати“ (35,9%), а потом следе текстуални задаци (19,5%).

С обзиром на слабију разноврсност форми задатака математике у односу на задатке из српског језика, поређење заступљености глагола у формулацији задатака различитих форми, остварено је само у односу на задатке испуњавања налога. Када је реч о задацима испуњавања налога, у тестовима из српског језика, најфреквентнији су глаголи „одредити“ (34,3%) и „написати“ (30,1%). У случају математичких задатака испуњавања налога, најфреквентнија је употреба глагола „израчунати“ (40,6%), а потом следе текстуални задаци (19,8%). Иако је математичких задатака форме испуњавања налога било 20% више него задатака из српског језика (212 задатака из математике и 143 задатка

из српског језика), број различитих глагола употребљених у формулацијама приближно је једнака (28 различитих глагола у формулацији математичких задатака и 26 различитих глагола у формулацији задатака из српског језика). Може се закључити да је у случају задатака испуњавања налога, приметна тенденција да се наставници српског језика користе разноврснијим глаголима у формулацији задатака у односу на наставнике математике.

Испитивањем није утврђено постојање статистички значајне разлике у одговорима наставника српског језика и математике када је реч о: одређењу садржаја теста и задатака, одређивању нивоа знања који се задатком мери и тежине задатака, избору форме задатака, утврђивању начина оцењивања теста, и коришћењу примера задатака других аутора - као фаза конструкције теста; одређењу циља примене теста знања; наставном програму и образовним стандардима као основама за разраду садржаја теста знања; одређеним аспектима реализације тестирања; чувању и развијању базе задатака; коришћењу готових тестова знања; циљу анализе ученичких постигнућа на нивоу појединачних задатака; одређеним аспектима провере квалитета теста; минималном броју паралелних форми теста; одређивању доњег прага поена на тесту за школске оцене добар, врло добар и одличан; карактеристикама појединих задатака на нивоу теста (Прилог 3, табеле 90-100).

Испитивањем су регистроване значајне разлике у методолошким карактеристика тестова из области српског језика и математике. Резултати анализе упућују на закључак да су наставници математике више фокусирани на аспекте стандардизације услова тестирања као што су: препознавање израде кључа као фазе конструкције теста, израда паралелних форми теста, провера тежине задатака пре примене теста, одређивање дужине теста, утврђивање редоследа задатака у тесту. Наставници српског језика више наглашавају значај саопштавања упутства за решавање теста, што се може сматрати разумљивим с обзиром на већу разноврсност форми задатака у тесту. Из позиције ученика, решавање теста из српског језика траје дуже, али је лакше остварити праг за

пролазну оцену, с обзиром на већу заступљеност лаких задатака. Тест из математике садржи мање задатака, али је ученицима вероватно тежи од теста из српског језика, с обзиром на мању заступљеност лаких задатака. Уочене разлике у методолошким карактеристикама тестова знања из српског језика и математике подстицајне су за нова испитивања којима би се могле продубити наведене анализе. У извесној мери, разлике се могу објаснити разликама у природи градива које се изучава – у математици се више наглашава егзактност него у настави српског језика. Тежња као већој или мањој егзактности, одражава се и на разноврсност форми задатака у тесту. Традиционално, примена тестова више је карактеристична за наставу математике него за наставу српског језика. Наставници српског језика традиционално више примењују друге начине испитивања ученика (усмено испитивање и писане вежбе) (Ebel, 1979), тако да се и на тај начин може објаснити боље структурирање задатака у тестовима из математике.

Сходно добијеним налазима, може се закључити да је хипотеза која се односи на разлике у фазама и методолошким карактеристикама тестова знања које наставници српског језика и математике самостално израђују – делимично потврђена.

КОРИШЋЕЊЕ РЕЗУЛТАТА ДОБИЈЕНИХ ПРИМЕНОМ ТЕСТА ЗНАЊА ЗА УНАПРЕЂИВАЊЕ ВАСПИТНО-ОБРАЗОВНОГ РАДА ШКОЛЕ

У циљу утврђивања начина на које се могу користити резултати тестова знања у школској пракси размотрили смо начине коришћења резултата тестова које наставници самостално састављају, начине на које се на нивоу школе користе резултати Мале матуре, националних и међународних тестирања, и процену наставника о значају резултата Мале матуре, националних и међународних тестирања за унапређивање васпитно-образовног рада школе.

Табела 103. Циљеви примене резултата теста знања које наставници самостално израђују

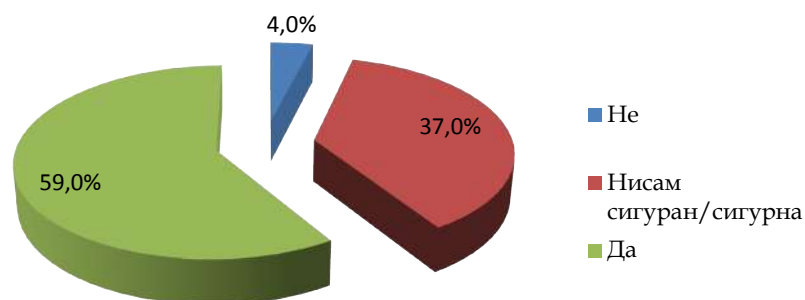
Циљеви примене	f	%
оцењивање ученика	286	79,0
утврђивање претходног знања које ученици имају у вези са одређеним делом градива	208	57,5
планирање и предузимање мера подршке ученицима у учењу	179	49,4
организовање рада у допунској настави	100	27,6
планирање садржаја и начина рада у настави	86	23,8
размена информација о напредовању ученика са родитељима и разредним старешином	84	23,2
организовање рада у додатној настави	82	22,7
унапређивање процеса израде теста знања	57	15,7
планирање стручног усавршавања	15	4,1
неки други циљ	7	1,9

N = 362

Навећи број наставника (79,0%) користи резултате тестова знања у циљу оцењивања ученика. Следећи одговори према учесталости су коришћење резултата тестирања у циљу утврђивања претходног знања ученика (57,5%) и планирања и предузимања мера подршке ученицима (49,4%). Домен организовања и планирања наставног рада, као и размена информација о напредовању ученика са родитељима и разредним старешином представљају следеће области примене резултата тестова знања према учесталости. Најређе, резултати тестова знања користе се у циљу унапређивања процеса

конструкције теста знања и планирања стручног усавршавања (табела 103) Дакле, тестови знања се нешто више користе у сврху сумативне евалуације (оцењивање ученика, испитивање претходног знања) него формативне (планирање и предузимања мера подршке ученицима). Мада је коришћење тестова знања у циљу оцењивања ученика тенденција која се у литератури сматра више традиционалном (Porham, 2001), она је употпуњена напорима наставника да тестове користе и за планирање и унапређивање рада са ученицима.

Када је реч о Малој матури, резултати показују да већина наставника (59,0%) сматра да се резултати Мале матуре користе на нивоу школе, значајан број наставника (37,0%) није сигуран да ли је то тачно, а мали број наставника (4,0%) сматра да се резултати не користе (слика 17). Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика на нивоу 0,01 ($\chi^2 = 161.350$; $df = 2$; $p = 0.00$).



Слика 17. Коришћење резултата Мале матуре на нивоу школе

Свега 38,7% наставника указало је на начине на које се резултати Мале матуре користе на нивоу школе. Анализом садржаја одговора, издвојиле су се две категорије одговора – допринос планирању наставног процеса у наредној школској години (47,1%) и анализа и вредновање остварених резултата (36,4%). Резултати хи-квадрат теста показују да у заступљености два наведена одговора не постоји статистички значајна разлика. Укупно 16 (11,4%) наставника сматра да се на нивоу школе резултати Мале матуре користе и за планирање

наставног процеса у наредној школској години и за анализу и вредновање остварених резултата. Мали број наставника (5,0%) понудио је одговоре који упућују на закључак да су резултати Мале матуре важни за рад школе, али нису специфичније одредили на који начин.

Како би се могло приступити разматрању начина примене резултата националних и међународних тестирања, било је важно утврдити колико су наставници информисани о реализацији појединих истраживања на нивоу школе. Тек мањи број наставника (23,2%) могао је да потврди да је у школи реализовано неко од националних и међународних испитивања у последњих 12 година, док се укупно 76,8% наставника определило за одговор „нисам сигуран/сигурна“. Табела 104. приказује о којим тестирањима је реч.

Табела 104. Тестирања реализована у школи

Назив тестирања	f	%
PISA 2012	19	16,7
национално тестирање 2005-2006	18	15,8
TIMSS 2011	16	14,0
TIMSS 2007	14	12,3
TIMSS 2003	11	9,6
национално тестирање 2003-2004	10	8,8
PISA 2003	3	2,6
PISA 2009	2	1,8
PISA 2006	0	0,0
неко друго	21	18,4
укупно	114	100,0

Око петине испитаника који су указивали на реализацију одговарајућих тестирања, известило је о учешћу школе у неким другим испитивањима као што су TALIS 2013, тестирање регионалног центра за таленте, пилот пројекат за национално тестирање и слично.

Слаба информисаност наставника о реализацији тестирања као што су национална и међународна тестирања, представља забрињавајући податак. Реч је о испитивањима која имају национални значај и могло би се очекивати да су наставници боље информисани о њиховој реализацији. Образложење

добијеног налаза може се потражити у чињеници да су разматрана тестирања могла бити реализована у време када наставници нису били запослени у школи. Детаљнијим испитивањем могло би се више сазнати о конкретним разлозима слабе информисаности наставника. У значајној мери, затечено стање ограничило је могућност детаљније анализе која се односила на процену значаја националних и међународних испитивања.

Када је реч о националним тестирањима, већина наставника није сигурна да ли се резултати националних испитивања користе на нивоу школе (61,0%). Око шестине испитаника (16,0%) сматра да се резултати националних испитивања не користе у школи, а још мањи број испитаника (11,9%) сматра да се резултати користе. Слична дистрибуција одговора добијена је и на питање у вези са резултатима међународних истраживања: већина наставника није сигурна да ли се резултати међународних тестирања користе на нивоу школе (61,0%), свега 5,2% наставника сматра да је то тачно, а 21,5% наставника сматра да то није пракса у њиховој школи. Резултати хи-квадрат теста показују да међу одговорима наставника постоји статистички значајна разлика на нивоу 0,01 и када је реч о коришћењу резултата националних ($\chi^2 = 181,61$; $df = 2$; $p = 0.00$; Прилог 3, табела 61) и међународних тестирања ($\chi^2 = 203,57$; $df = 2$; $p = 0.00$; Прилог 3, табела 62).

Резултати националних тестирања користе се у следеће сврхе: *„проверу знања ученика и проверу рада наставника“, „за планирање рада“, „планирање и унапређивање наставе“, „унапређивање садржаја наставе“, „унапређивање садржаја наставе и начина рада“, „побољшање наставног процеса и отклањање недостатака“, „утврђивање нивоа постигнућа по регионима и позиционирање саме школе“* и тако даље.

Резултати међународних тестирања користе се у следеће сврхе: *„унапређивање наставе и отклањање недостатака где је то могуће“, „планирање стручног усавршавања наставника и унапређивање процеса оцењивања“, „процену остварених резултата у настави (...)“, „да одредимо где смо и шта нам треба да побољшамо успех“, „унапређивање ваннаставног рада са ученицима“* и тако даље.

Резултати испитивања показују релативно подударну слику доприноса евалуативних испитивања различитим аспектима васпитно-образовног рада школе, при чему је акценат на доприносу резултата унапређивању рада у настави.

Применом дескриптивне скале, испитиван је значај резултата Мале матуре, националних и међународних тестирања за унапређивање васпитно-образовног рада школе. Добијене бодове вредности за сваки од разматраних аспеката, за свако од истраживања приказујемо у табели 105. Теоријски распон скале кретао се у вредностима од $M = 1$ до $M = 5$.

Табела 105. Значај резултата Мале матуре, националних и међународних тестирања за унапређивање васпитно-образовног рада школе

Коришћење резултата тестирања за:	Мала матура			национална тестирања			међународна теститања		
	N	M	SD	N	M	SD	N	M	SD
унапређивање садржаја наставе	350	4,01	0,935	327	3,85	0,853	320	3,73	0,881
унапређивање начина рада у настави	350	4,07	0,917	327	3,85	0,837	320	3,72	0,880
унапређивање процеса оцењивања	348	3,73	1,074	327	3,59	0,863	320	3,51	0,889
унапређивање ваннаставног рада са ученицима	349	3,43	1,058	325	3,43	0,820	320	3,43	0,857
планирање стручног усавршавања наставника	350	3,29	1,118	327	3,39	0,871	320	3,38	0,892
побољшање међуљудских односа у школи	350	2,56	1,207	327	2,77	0,997	320	2,84	0,950
побољшање опремљености школе	347	2,92	1,256	327	3,11	1,039	320	3,20	0,965
превазилажење разлика између деце различитог пола и породичног порекла	350	2,74	1,255	327	2,92	1,072	320	3,00	1,002

Наставници оцењују да је коришћење резултата Мале матуре, најзначајније у домену унапређивања рада у настави ($M = 4,07$), и унапређивања садржаја наставе ($M = 4,01$). Коришћење резултата Мале матуре најмање је значајно у домену побољшања међуљудских односа у школи ($M = 2,56$) и превазилажења разлика између деце различитог пола и породичног порекла ($M = 2,74$).

Аспекти унапређивања васпитно-образовног рада школе оцењени највишим и најнижим оценама у случају Мале матуре, истоветни су и у случају националних и међународних тестирања. Ипак, анализа добијених просечних вредности скорова показује да издвојена тестирања нису за сваки од разматраних домена подједнако значајна. Наиме, коришћење резултата Мале матуре у домену унапређивања рада у настави ($M_1 = 4,07$) и унапређивања садржаја наставе ($M_2 = 4,01$) оцењује се као значајније од коришћења резултата националних ($M_1 = 3,85$; $M_2 = 3,85$) и међународних тестирања ($M_1 = 3,72$; $M_2 = 3,73$). У тим доменима најмање повољну оцену добијају међународна тестирања. Истовремено, коришћење резултата Мале матуре у домену побољшања међуљудских односа у школи ($M_3 = 2,56$) и превазилажења разлика између деце различитог пола и породичног порекла ($M_4 = 2,74$), оцењено је као мање значајно од коришћења резултата националних ($M_3 = 2,77$; $M_4 = 2,92$) и међународних тестирања ($M_3 = 2,84$; $M_4 = 3,00$). Истина, у случају сва три тестирања просечан скор је релативно низак. У наведеним доменима благу предност у односу на друга тестирања имају међународна тестирања.

По свему судећи, резултати Мале матуре оцењују се више значајним са становишта наставе у односу на национална и међународна тестирања. Добијени налаз се може објаснити чињеницом да већина наставника обухваћених истраживањем предаје предмете чији садржај сачињава завршни испит. Несумњиво, наставницима су заинтересовани за садржај завршног испита и спремни да се у свом редовном раду фокусирају на садржаје који су његов саставни део.

Како би се добила потпунија слика о коришћењу тестова знања у настави, упоредили смо одговоре наставника који предају различите предмете и одговоре наставника различитих година радног стажа у вези: са циљевима израде тестова које самостално конструишу, коришћењу резултата Мале матуре, националних и међународних тестирања и значају коришћења резултата наведених тестирања у циљу унапређивања васпитно-образовног рада школе.

Када је реч о наставницима који предају различите школске предмете, одговори се на већину питања нису статистички значајно разликовали. Регистрована су три ужа аспекта проблема око којих су постојале разлике у одговорима. Најпре, испитивањем је утврђено да се одговори наставника различитих предметних области статистички значајно разликују када је реч о примени резултата теста у сврху планирања и предузимања мера подршке ученицима у учењу. Наиме, наставници друштвено-језичке групе предмета чешће примењују резултате тестова знања у сврху планирања и предузимања мера подршке ученицима у учењу у односу на наставнике природно-математичке групе предмета (табела 106). Утврђена разлика у одговорима наставника статистички је значајна на нивоу 0,01.

Табела 106. Предмет који наставник предаје и израда теста у циљу планирања и предузимања мера подршке ученицима у учењу

Предмет који наставник предаје	Израда теста у циљу планирања и предузимања мера подршке ученицима у учењу					
	не		да		укупно	
	f	%	f	%	f	%
друштвено-језичка	82	44,8	105	58,7	187	51,7
природно-математичка	101	52,2	74	41,3	175	48,3
укупно	183	100,0	179	100,0	362	100,0

$$\chi^2 = 6,95; df = 1; p < 0.01; C = 0,14$$

Такође, испитивањем је установљено да постоји статистички значајна разлика у одговорима наставника који предају различите предмете на питање о коришћењу резултата националних тестирања на нивоу школе. Наиме, наставници природно-математичке групе предмета чешће су извештавали о томе да се резултати националних тестирања не користе на нивоу школе ($\chi^2 = 6,15$; $df = 2$; $p < 0.05$; $C = 0,14$; Прилог 3, табела 63). Статистичка значајност разлике у одговорима наставника на дато питање постоји на нивоу 0,05.

Поред тога, регистровано је постојање статистички значајне разлике у проценама наставника различитих предмета када је реч о значају резултата Мале матуре за побољшање опремљености школе. Наиме, наставници природно-математичке групе предмета нешто повољније оцењују допринос резултата Мале матуре побољшању опремљености школе у односу на наставнике друштвено-језичке групе. Статистичка значајност разлике у ставу наставника постоји на нивоу 0,05 (табела 107).

Табела 107. Предмет који наставник предаје и коришћење резултата Мале матуре за побољшање опремљености школе

Предмет који наставник предаје	N	M	SD	SEM
друштвено-језичка група	179	2,77	1,19	0,09
природно-математичка група	166	3,08	1,31	0,10

$$t = - 2,33; df = 343; p = 0.02$$

Добијени податак се може протумачити са становишта нешто израженије потребе наставника природно-математичке групе предмета да се бољом опремљеношћу школских кабинета допринесе стварању оптималнијих услова за учење градива из биологије, хемије, физике, математике. Настава друштвено-језичке групе предмета ретко се реализује у специјализованим кабинетима, или је опремање таквих кабинета (за наставу историје, географије, српског и страних језика) мање сложено.

Међу одговорима наставника који предају различите предмете није утврђено постојање статистички значајне разлике када је реч о: изради теста знања са циљем испитивања претходног знања ученика, оцењивања ученика, планирања садржаја и начина рада у настави, организовања рада у додатној и допунској настави, размене информација о напредовању ученика са родитељима и разредним старешином, планирања стручног усавршавања, унапређивања процеса израде теста; коришћењу резултата Мале матуре и међународних тестирања на нивоу школе; значају резултата Мале матуре за унапређивање садржаја наставе, начина рада у настави, процеса оцењивања, ваннаставног рада са ученицима, планирање стручног усавршавања наставника, побољшање међуљудских односа у школи и превазилажење разлика између деце различитог пола и породичног порекла; значају резултата националних и међународних тестирања за унапређивање васпитно-образовног рада школе (Прилог 3, табеле 75-77).

Када је реч о наставницима различитих година радног стажа, одговори се на већину питања нису статистички значајно разликовали. Регистрована су два ужа аспекта проблема око којих су постојале разлике у одговорима. Наиме, утврђено је постојање статистички значајне разлике између наставника различитих година радног стажа у проценама значаја резултата: а) Мале матуре за унапређивање ваннаставног рада са ученицима и б) националних тестирања за унапређивање начина рада у настави.

Регистрована је тенденција да млађи наставници позитивније оцењују значај резултата Мале матуре за унапређивање ваннаставног рада са ученицима ($F = 4,16$; $df_1 = 2$; $df_2 = 346$; $p = 0.02$; Прилог 3, табела 64). Регистровану тенденцију није лако објаснити на основу расположивих података. Резултати Мале матуре могу се сматрати значајним са становишта организације ваннаставног рада у смислу допунског или додатног рада са ученицима. Могуће је да млађи наставници чешће посматрају ваннаставни рад као шансу да се надокнади оно што се не оствари у редовном раду, а што је важно са становишта завршног испита. Ипак, такав закључак би било неопходно проверити додатним испитивањима.

Наставници који имају између 10 и 20 година радног стажа повољније оцењују допринос резултата националних тестирања унапређивању начина рада у настави у односу на наставнике са мање од 10 година радног стажа и у односу на наставнике са више од 20 година радног стажа ($F = 4,03$; $df_1 = 2$; $df_2 = 324$; $p = 0.02$; Прилог 3, табела 65). Последња национална испитивања у нашој земљи реализована су пре 10 година. Могуће је да њихов значај боље могу да сагледају наставници који су у то време били запослени у школи. Постојање разлике у проценама најискуснијих наставника и њихових колега са нешто мање година радног стажа (10-20 година) није могуће једнозначно објаснити на основу постојећих података.

Међу одговорима наставника различитих година радног стажа није утврђено постојање статистички значајне разлике када је реч о: циљевима израде тестова; коришћењу резултата Мале матуре и националних тестирања на нивоу школе; значају резултата Мале матуре за унапређивање садржаја наставе, начина рада у настави, процеса оцењивања, планирање стручног усавршавања наставника, побољшање међуљудских односа у школи, побољшање опремљености школе и превазилажење разлика између деце различитог пола и породичног порекла; значају резултата националних тестирања за унапређивање садржаја наставе, процеса оцењивања, ваннаставног рада са ученицима, планирање стручног усавршавања наставника, побољшање међуљудских односа у школи, побољшање опремљености школе и превазилажење разлика између деце различитог пола и породичног порекла; значају резултата међународних тестирања за унапређивање васпитно-образовног рада школе (Прилог 3, табеле 87-89).

Ретке регистроване разлике могу бити показатељ недовољне информисаности наставника о томе коју врсту података евалуативне студије обезбеђују. Наочигледнији резултат свакако јесте постигнуће ученика на тесту, које наставници неретко директно доводе у везу са наставним процесом. Евалуативне студија, међутим, испитују и значајан број контекстуалних варијабли које омогућавају да се васпитно-образовни рад школе целовитије сагледа.

ЗАКЉУЧАК

Иако тестови знања нису нов истраживачки проблем, бројни чиниоци су допринели томе да већ деценијама не изгубе на својој актелности. Са напретком науке и технологије, мењао се, како теоријски, тако и практични приступ тестовима. Јавиле су се нове дилеме, а нека ранија питања изнова су отварања.

До половине XX века развијене су две теорије тестирања на којима се конципирају тестови знања. То су класична теорија тестирања (СТТ) и теорија ставског одговора (IRT). Теорије полазе од различитих претпоставки када је у питању: природа својства које се мери, начин на који испитаник решава задатке у тесту, стандардна грешка мерења, разумевање „правог“ скорa који испитаник може остварити на тесту, интерпретација скорa с обзиром на контекст, одређење тежине ајтема у тесту, однос према могућности погађања тачних одговора, испитивање односа квалитета дистрактора у задацима бирања одговора и резултата на тесту. Иако је конципирање тестова на темељу класичне теорије тестирања и даље заступљено, оцењује се да конципирање тестова на темељу теорије ставског одговора обезбеђује прецизније мерење. Последњих година, развија се и трећи модел мерења који би требало да превазиђе заједнички недостатак IRT и СТТ – немогућност моделовања процеса који стоје у основи испитаниковог понашања и ближе одређење природе мереног својства. Реч је о моделу мерења базираном на когнитивним процесима, чија ће се разрада и емпиријска провера реализовати у будућности.

У складу са развојем теорија тестирања, мењали су се и приступи конструкцији теста. Конвенционални приступ је развијен на основама класичне теорије тестирања, а ECD приступ на IRT моделима мерења. У класичном приступу, конструкција теста се описује низом фаза које се крећу од одређења циља теста, преко израде задатака (одређење нивоа знања који

задатак мери, тежине, избор форме), вишеетапних провера квалитета теста до обликовања финалне форме теста. Наведеним активностима, у ECD приступу придружују се активности усмерене на описивање везе ајтема и мереног својства. Битна новина у односу на конвенционални приступ је гарантовање валидности закључака о постигнућу ученика управо на основу јачине везе ајтема у тесту и мереног својства. Конвенционални приступ је тренутно више разрађен, тако да је у нашем истраживању послужио као методолошки оквир за разматрање фаза конструкције теста које наставници израђују.

У циљу конструкције доброг теста знања, проверавају се метријске карактеристике валидности, поузданости, објективности, осетљивости теста и дискриминативности задатака, тежине задатака и једнодимензионалности теста. Последњих деценија, појавила су се нова тумачења појединих метријских карактеристика и унапређивани су статистички поступци којима се остварује њихова провера. Промене су посебно изражене у одређењима валидности теста о којој се данас говори са становишта релевантности, вредносних импликација и друштвених последица одлука изведених на основу резултата тестирања (Messick, 1981; Shepard, 1993).

Упркос променама, анализа методолошке литературе указује на усаглашеност аутора по питању одређења значења теста. Мада разлике могу постојати на нивоу општости дефиниције, тестирање се обично сагледава као поступак који се припрема, реализује и чији се резултати вреднују према унапред одређеним правилима и принципима. Резултати тестирања омогућавају да се испитаници међусобно упореде у погледу оствареног постигнућа, или да се постигнуће испитаника упореди са постављеним критеријумом. Тест се, дакле, углавном дефинише на исти начин већ неколико деценија. Оно што се променило је то што су новије класификације тестова знања богатије за једну врсту теста. Реч је о врсти коју смо у раду означили као електронски тест знања. Њихова израда је постала могућа са напретком технологије и све смо ближе тренутку када ће компјутерски подржано тестирање постати део школске праксе.

Тестови знања су један од начина праћења и вредновања постигнућа ученика у школи. Примењују се паралелно са усменим испитивањем и писаним вежбама, у односу на које могу имати одређене предности и органичења. Резултати нашег испитивања показују да највећи број наставника указује на предности тестова знања над усменим испитивањем у доменима: веће економичности са становишта броја ученика који се тестирањем може обухватити, истоветности критеријума оцењивања и истоветности захтева који се постављају пред ученике. Добијени налаз обухвата неке од најчешће сагледаваних предности тестова знања у односу на усмено испитивање (Хавелка, 2000; Бјекић и Голубовић, 2003). Резултати показују да је број наставника који указују на предности тестови знања у односу на писане вежбе приближно уједначен броју наставника који то не чине. Налаз се може објаснити концепцијском сличношћу тестова и писаних вежби у школској пракси. Када је реч о компарацији класичних и електронских тестова знања, установљено је да тек мањи броју наставника указује на предности класичних тестова над електронским тестовима. Добијени налаз има подршку и у савременој литератури о тестовима знања (He & Tumm, 2005). Ипак, о наведеној компарацији известила је половина од укупног броја испитаника, што указује на недовољно познавање карактеристика електронских тестова на нивоу узорка испитаника.

Сходно резултатима, може се очекивати да постоје одређени недостаци тестова знања који их не чине увек конкурентним другим начинима вредновања ученичког постигнућа. Резултати показују да се недостаци тестова најчешће сагледавају у мањем квалитету добијеног одговора у односу на одговор који се може добити другачијим начином испитивања (у другим околностима одговор може бити образложен, указати на учење са разумевањем; ученик може исказати своје мишљење, креативност и слично). Наставници оцењују да тестови не пружају могућност да се помогне ученику приликом одговарања, што је од посебног значаја за ученике који су генерално мање успешни у саопштавању одговора писаним путем. Ипак, начин на који наши испитаници сагледавају недостатке тестова знања у односу на усмено

испитивање и писане вежбе, не обухвата сав потенцијал тестова знања као инструмената којима се могу испитивати различити нивои мишљења и који могу садржати задатке различите тежине и форме.

Упркос томе што имају значајне предности, тестови нису савршени инструменти и њихова критика заступљена је и у литератури. Критика тестова знања посебно је карактеристична за представнике аутентичног процењивања ученичких постигнућа који наглашавају немогућност тестова да испитују процес којим се стиже до одређеног резултата, ограничену могућност прилагођавања захтева у тесту индивидуалним карактеристикама ученика и ограничењима у испитивању активности као што су решавање проблемских ситуација, креативно мишљење, иницијатива, упорност. Може се очекивати да тек електронски тестови превазиђу неке од наведених недостатака класичних тестова знања.

Питање припремљености наставника за израду тестова знања представља важно питање у контексту израде и примене тестова знања. Ипак, у нашој средини врло је мало испитивања која се баве тим проблемом. Резултати истраживања које смо спровели, показују да већина наставника своју припремљеност за израду тестова знања оцењује високом оценом. При том, допринос припреме током иницијалног професионалног образовања је лоше процењен. Како резултати показују, припрема током иницијалног професионалног образовања повремено је потпуно изостајала, а повремено је била конципирана на начин који није омогућавао успешно овладавање знањима и вештинама потребним за израду тестова знања. Стручно усавршавање у домену израде тестова знања више је ослоњено на неформалне видове, као што су консултације са колегама, него на формалне видове усавршавања (семинаре, стручне скупове). Прецизније, највећи број наставника до тренутка испитивања није похађао ниједан семинар посвећен изради тестова знања. Један семинар у референтној области похађало је мање од трећине испитаника, а број наставника опада пропорционално са повећањем броја семинара. Ипак, наставници који су похваљали семинаре посвећене изради тестова знања, релативно повољно су оценили њихов

допринос. Резултати такође показују да лична иницијатива наставника за похађањем семинара доминира над иницијативом других, а немали број наставника самостално је финансирао своју обуку. Могло би се закључити да наставници јесу мотивисани за усавршавањем у домену израде тестова. Како наставници истичу, семинари стручног усавршавања представљају један од основних начина да се стекну релевантна сазнања и вештине потребне за израду тестова, али их је потребно учинити приступачнијим.

Укупно две трећине анкетираних наставника изјављује да се суочава са тешкоћама у изради тестова знања. Најчешће, у питању су потешкоће које се односе на: одређивање тежине задатака, броја поена које задатак носи, одређивање нивоа знања који задатак мери, дужине теста, избор врсте задатака и превођење поена на тесту у оцене. У циљу превазилажења тешкоћа, наставници ће се најпре обратити својим колегама или ће покушати самостално да реше проблем. Стручни сарадници у школи ретко су препознати као особе које могу да помогну наставницима да реше настале проблема. По свему судећи, значајном броју наставника неопходно је пружити подршку у процесу израде тестова знања. Када се оснаже њихове компетенције у домену израде тестова знања, може се очекивати и да тешкоће буду мање изражене. Даљом анализом је установљено да у околностима недовољне припреме наставника у току иницијалног професионалног образовања и слабе заступљености стручног усавршавања у домену израде тестова знања, професионално искуство наставника може бити ослонац у процесу конструкције добрих тестова знања. Чињеница да ће се мање искусне колеге радо обратити искуснијим колегама када имају потешкоће у изради тестова знања, додатно поткрепљује такав закључак.

Фазе конструкције теста које се разматрају у литератури, дефинисане су за потребе научних истраживања, па се може очекивати да фазе конструкције тестова у школској пракси не буду истоветне. Најпре, потребе праксе су другачије од потреба научног испитивања. Поред тога, наставници су у различитој мери упознати са методолошким захтевима конструкције теста. Испитивањем је утврђено да конструкција теста знања у школској пракси

најчешће обухвата фазе утврђивања наставног садржаја који ће се тестом обухватити, нивоа знања који се задатком испитује, начина оцењивања теста и израде кључа. Преостале фазе конструкције теста препознаје тек мањи број наставника. Око трећине наставника фазама конструкције теста означава активности које нису дефинисане методолошким оквиром конструкције теста. Такав резултат показује да појединим наставницима недостаје методолошко знање у вези са конструкцијом теста које би им омогућило да јасније разграниче техничка питања израде теста (на пример, умножавање тестова) од методолошких фаза конструкције теста. Такође, наставницима је потребно помоћи да увиде да, на пример, утврђивање „нивоа одељења“ није фаза конструкције теста, али је важно питање, које се може довести у везу са фазама конструкције теста као што су разматрање нивоа знања који се задатком мери или тежине задатка.

Оквири за разраду садржаја теста могу бити различити: теме, циљеви и задаци наставног програма, садржај уџбеника, исходи, компетенције, стандарди дефинисани за одговарајући наставни предмет, таксономије образовних циљева, али и лична процена наставника шта је оно што је ученицима битно за наставак школовања. Да је испитивање реализовано, на пример, пре 10 година, о одређеним оквирима за разраду садржаја теста не би се могло говорити, јер нису били доступни наставницима. Избор једног од наведених оквира у извесној мери детерминише процес конструкције теста знања. На пример, избор циљева или општих задатака наставног програма као оквира за разраду садржаја теста подразумева ослањање на општији и мање прецизан оквир. Избор образовних исхода и стандарда подразумева ослањање на оквир који је конкретнији и прецизира шта је то што би ученик требало да зна, које вештине и умења да стекне по завршетку учења. Резултати истраживања показују да највећи број наставника у разради садржаја теста узима у обзир и општији и конкретан оквир – наставни програм, стандарде знања, садржај уџбеника. Могућност избора одређена је информисаношћу наставника да избор постоји. Било би важно истражити у којој мери су наставници информисани о постојању и специфичностима других оквира за

разраду садржаја теста. Такође, отворено је питање колико су наставници упознати са вредносним утемељењем одговарајућих оквира и да ли имају критички однос према томе.

Иако постоје другачија схватања, у раду се прихвата становиште да су задаци основне јединице теста. Питање израде задатака представља најважније питање конструкције теста. Може се рећи да је то истовремено и најтеже питање у изради теста знања, из неколико разлога. Најпре, задаци су одраз нашег разумевања оквира од кога смо пошли у разради садржаја теста. Такође, конструкција задатака подразумева укрштање неколико чинилаца: форме задатка, нивоа знања који задатак испитује и тежине задатка.

Када је реч о форми задатака, подела од које смо пошли у раду је подела на задатке са бирањем одговора и на задатке са уписивањем одговора. Резултати испитивања показују да задаци испуњавања налога представљају најзаступљенију форму задатака у анализираним тестовима из српског језика и математике. Према учесталости следе задаци са кратким одговорима, док су остале форме значајно ређе заступљене. Када је реч о различитим нивоима знања који се задацима испитују, већина анализираних тестова садржи задатке који испитују три различита нивоа знања – ниво познавања чињеница, разумевања градива и примене знања. Ипак, у тестовима обухваћеним анализом најзаступљенији су задаци који утврђују ниво познавања чињеница. Добијени резултат је у сагласности са резултатима ранијих испитивања (Гргин, 2001; Баковљев, 1997; Хавелка, 2000; Бјекић и Голубовић, 2003; Girps, 2003; Дубљанин, 2010; Илић-Голубовић, 2011).

Већина тестова обухваћених анализом садржи задатке на три различита нивоа тежине: лаке, средње тешке и тешке задатке. Анализирани тестови садрже највише средње тешких задатака. Такав налаз се може сматрати оправданим са становишта класичне теорије тестирања (Буквић, 1996). Анализа показује да се може проблематизовати питање односа између нивоа знања који се задатком испитује и тежине задатка. Установљена је тенденција да наставници нивое знања поистовећују са нивоима дефинисаним

образовним стандардима и са тежином задатака. Утврђену тенденцију би било потребно детаљније истражити новим испитивањима.

Задаци испуњавања налога у анализираним тестовима представљају доминантну форма задатка на свим нивоима испитивања знања. Добијени подаци су у извесном раскораку са препорукама које срећемо у литератури, према којима се примена дате форме задатака обично везује за испитивање виших нивоа знања (Popham, 2003; Roberson-Scott, 2005). Избор осталих врста задатака обухваћених анализом (задатака са кратким одговорима, допуњавања, вишеструког избора и спаривања) за испитивање одговарајућих нивоа знања углавном је усаглашен са препорукама које срећемо у литератури.

Може се уочити тенденција да наставници најчешће израђују задатке који нису сложени за конструкцију. Међутим, конструкција задатака испуњавања налога или задатака са кратким одговорима је једноставна ако желимо да испитамо ниво познавања чињеница. У случају да желимо да испитамо сложеније нивое знања, израда наведених форми захтева више ангажовања од стране наставника. На основу налаза, можемо закључити да је наставнике потребно боље упознати са карактеристикама појединачних врста задатака, пре свега са нивоима знања који се њима испитују. Свакако, налаз да у тестовима доминирају задаци за испитивање нивоа познавања чињеница завређује пажњу. У извесном смислу, резултат није добар, јер показује да се у начину конструкције задатака није много тога променило већ 20 година.

Садржај теста обухвата и садржај упутства за решавање теста. Иако је испитивањем предвиђена анализа садржаја упутства, установљено је да упутства нису интегрални део тестова које наставници израђују. Уколико се упутство припрема, оно се саопштава усменим путем. Често, наставници не саопштавају ученицима било каква упутства пре почетка тестирања. Постојање упутства за израду теста предвиђа се у циљу боље стандардизације услова тестирања. Стандардизацији доприносе и обезбеђивање одговарајућих просторно-временских услова за реализацију тестирања и организовање припреме за тестирање. Истраживањем је утврђено да просторна и временска организација тестирања у нашим школама нису нарочито наглашене. Већина

наставника припрема ученике за тестирање у току једног до два наставна часа. На часовима припреме се најчешће указује на садржај градива који ће бити обухваћен тестом, док је упознавање са формама задатака или начинима њиховог решавања ретко заступљено.

Може се закључити да постоји простор да се у домену стандардизације услова тестирања унапреде сазнања наставника. Стандардизација услова тестирања једно је од суштинских одређења технике тестирања. Разумљиво је да се не могу сви захтеви стандардизације који важе за примену тестова у научне сврхе применити у школској пракси, али је важно предочити наставницима сврху стандардизације појединачних услова. Важно је указати на који начин се, на пример, недостатак упутства за решавање теста или изостанак припреме за тестирање може одразити на резултате тестирања.

У домену стандардизације поступка вредновања ученичких одговора, ситуација је нешто повољнија – готово сви наставници израђују кључ за оцењивање теста којим се дефинише шта се све може признати као тачан или делимично тачан одговор. Међутим, у погледу одређења доњег прага за остваривање одговарајуће школске оцене постоји доста разлика међу анализираним тестовима. Распон процента тачних одговора/решења највише варира у случају пролазне оцене – распон се креће од 15% до 60% тачних одговора/решења на тесту. Најужи распон је регистрован у случају највише оцене (од 80% до 95% тачних одговора/решења на тесту). Добијени налази упозоравају да је стари проблем „двојке код строгог наставника“ и „четворке код благог наставника“ тешко превазићи чак и применом тестова знања. У циљу решавања проблема потребно је радити на усаглашавању прага постигнућа за појединачне оцене на нивоу школске установе, а ако је могуће, и шире.

Када говоримо о провери квалитета тестова које наставници самостално израђују, испитивањем је установљено да већина интервјуисаних наставника проверава квалитет теста пре његове примене, и то уз консултације са другим колегама. Доминантни циљеви провере квалитета теста јесу провера тежине и формулације задатака. Задаци и тестови чијим квалитетом су наставници

задовољни, најчешће се чувају у електронским базама задатака, мада је присутна и тенденција да се паралелно чувају и штампане верзије теста. Иако постојање базе омогућава наставницима да исте задатке употребе више пута, испитивањем је установљено да већина наставника барем једном током школске године израђује потпуно нов тест знања. Истовремено, око половине интервјуисаних наставника барем једном годишње ће се одлучити за примену тестова других аутора, а трећина наставника ће из готових тестова преузети поједине задатке. Најчешће су у питању тестови других колега, мада значајан број наставника користи тестове различитих издавачких кућа, или са Интернета. По свему судећи, наставници уочавају важност израде тестова који су нови, али им је повремено лакше да се ослоне на тестове које је неко други припремио. Потреба да се употреби тест другог аутора, можда се може објаснити чињеницом да већина наставника тестове примењује четири пута годишње, а значајан број наставника то чини 5 и више пута током школске године. Регистрована учесталост примене тестова знања није мала, уколико се има у виду чињеница да је задавање теста одређено недовољно флексибилним планом тестирања на нивоу школске установе. Мада се може разумети опредељење наставника да се помогну већ израђеним тестовима, важно је имати критички однос према њиховом садржају.

Испитивањем су регистроване значајне разлике у појединим методолошким карактеристикама тестова из области српског језика и математике. Образложење се може потражити у специфичностима садржаја који се изучава, потом специфичности припреме током иницијалног професионалног образовања наставника, или у традицији израде тестова знања за потребе наставе. Уочене разлике у методолошким карактеристикама тестова знања из српског језика и математике подстицајне су за нова испитивања којима би се могле продубити наведене анализе. Исто тако, анализом би се могли обухватити и други школски предмети.

Налази испитивања које смо реализовали указују на тенденцију да се тест знања не схвата нужно као средство оцењивања ученика, што је било карактеристично за првобитну примену тестова знања. Наиме, иако је

најчешћи циљ задавања теста оцењивање ученика, врло често, наставници израђују тест како би се проверила савладаност градива од стране ученика која нема за циљ извођење оцене. Резултати тестова које наставници самостално израђују доминантно се користе у сврху сумативне евалуације. Усмерење ка унапређивању процеса учења применом тестова, које представља савремено становиште у педагошкој литератури, у школској пракси још увек није довољно препознато.

Поред тестова које наставници самостално конструишу, у школској пракси се користе и стандардизовани тестови знања. У нашем школском систему, њихова примена је карактеристична за реализацију завршног испита на крају обавезног школовања (Мале матуре), националних и међународних тестирања. Резултати завршног испита у нашем образовном систему обично се доводе у везу са диференцирањем ученика при упису у средњу школу. Ипак, то није једина функција завршног испита. Очекује се да резултати Мале матуре омогуће и процењивање ефеката основног васпитања и образовања, дају подстицај и усмерење у процесу унапређивања наставе и учења у основној школи. Резултати истраживања показују да већина наставника сматра да се резултати Мале матуре користе на нивоу школе. Око трећине наставника указало је на начине на које се то чини, као што су допринос планирању наставног процеса у наредној школској години, анализа и вредновање остварених резултата. Наставници оцењују да је коришћење резултата Мале матуре најзначајније у домену унапређивања рада у настави и унапређивања садржаја наставе, а најмање значајно у домену побољшања међуљудских односа у школи и превазилажења разлика између деце различитог пола и породичног порекла. Организација завршних испита има дугу традицију у развијеним земљама света, док је у наш образовни систем Мала матура уведена тек школске 2010/2011. године. Можемо очекивати да ће ефекти завршног испита потпуније бити сагледани у годинама пред нама.

У раду је представљена и концепција националних тестирања. За сва тестирања на националном нивоу карактеристично је да полазе од очекивања да се тестирањем ученика могу прикупити информације о ефикасности

образовног система. На темељу стечених сазнања, могу се доносити конкретне одлуке у домену образовне политике. Између тако постављеног циља и његове реализације стоје бројна нерешена питања у вези са могућностима преношења научних сазнања у свет праксе. Додатно, очекивање да резултати тестирања буду мерило успешности образовног система, има значајне последице по рад наставника и организацију наставног процеса. Када је реч о међународним испитивањима, у раду су приказани PISA и TIMSS испитивање. Оба испитивања се реализују применом тестова знања, мада су им концепције различите. PISA студија испитује у којој мери су ученици успели да развију одговарајуће компетенције (читалачку, математичку и научну писменост). TIMSS студија бави се анализом садржаја образовног програма једне земље (предвиђени курикулум), програма који је реализован у пракси (примењени курикулум), и програма које је усвојен од стране ученика (постигнути курикулум). Реализација међународних испитивања усмерена је ка компарацији образовних система. Полази се од уверења да су поређења образовних система могућа, односно да друштвене и културне разлике не умањују могућност компарације. Дато становиште вишеструко је оспоравано. Ограничења међународних студија односе се и на неусаглашеност наставних садржаја који се изучавају у школама различитих земаља и немогућност формирања узорка испитаника према унапред дефинисаном моделу.

У домену испитивања начина на који се на нивоу школе могу користити резултати националних и међународних испитивања, истраживањем су добијени релативно скромни подаци који ограничавају могућност извођења прецизнијих закључака. До тога је довела околност да око три четвртине наставника нису могли са сигурношћу да извештају о реализацији националних и међународних испитивања у школи у којој раде.

Стандардизовани тестови пружају корисне информације о наставном процесу које се не могу добити на други начин. Сходно Пофамовом становишту (2001), потенцијал стандардизованих тестова знања није у потпуности искоришћен и њихов квалитет се може даље унапређивати.

Примена тестова знања један је од уобичајених начина испитивања ученичких постигнућа. Иако има предности, тест знања се не може посматрати као искључиви начин праћења и вредновања постигнућа ученика. Знања и вештине које се испитују тестовима су сложена својства и пожељно је процењивати их применом различитих поступака. Одређена својства није погодно испитивати тестовима (на пример, сарадњу, труд, комуникацију ученика), док поједина сазнања не могу бити комплетна без примене теста (на пример, положај ученика у односу на друге ученике одељења према усвојености градива одређеног предмета). Највећи значај за даљи развој педагогије тестови могу имати ако се користе са мером и уколико се води рачуна о ограничењима њихове примене.

Реализовано истраживање пружа одређена сазнања у вези са применом тестова знања која могу бити оквир за даља испитивања или допринети променама у домену школске праксе. Ипак, испитивањем нису размотрена сва питања важна за примену тестова знања у школској пракси. Уколико би постојала могућност да се организује семинар стручног усавршавања за наставнике којим би се обухватила питања конструкције тестова знања, препоручујемо да се следећи проблеми обухвате обуком: фазе конструкције теста, оквири за разраду садржаја теста, карактеристике различитих типова задатака са нагласком на могућности испитивања различитих нивоа знања, вредновање ученичких одговора са нагласком на модел превођења скорана тесту у оцене, стандардизација услова тестирања, предности и ограничења тестова знања као поступка вредновања ученичких постигнућа, потенцијал тестова у домену формативне евалуације, електронски тестови знања. Наша препорука је и да се улога стручних сарадника у организацији сличних обука учини значајнијом. Обуке могу имати и интерни карактер, а стручним сарадницима у реализацији могу помоћи наставници са релевантним искуством. На тај начин би се испунили услови које поједини наставници истичу као предуслове успешног семинара – били би бесплатни, реализовани у месту у коме наставници раде, а реализатори би биле особе које поседују искуство у изради и примени тестова знања.

ЛИТЕРАТУРА

- Ackerman, T. A. (1992): A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective, *Journal of Educational Measurement*, Vol. 29, No. 1, 67-91.
- Algarabel, S. & C. Dasi (2001): The definition of achievement and the construction of tests for its measurement: a review of the main trends, *Psicológica*, Vol. 22, No. 1, 43-66.
- Ali Salmani Nodoushan, M. (2010): The interface of interim assesment and feedback: an opinion paper, *I-manager's Journal on Educational Psychology*, Vol. 4, No. 3, 1-8.
- Al-Shara'h, N. D. (2013): An investigation of Jordanian EFL teachers' procedures of achievement test construction, *Education*, Vol.133, No. 4, 422-430.
- Anastasi, A. (1984): Aptitude and Achievement Tests: The Curious Case of the Indestructible Strawperson; in B. S. Plake (Ed.): *Social and Technical Issues in Testing: Implications for Test Construction and Usage* (129-140). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Anastasi, A. (1986): Evolving concepts of test validation, *Annual Review of Psychology*, Vol. 37, No. 1, 1-15.
- Assessment in the primary school curriculum: guidelines for schools* (2007): Dublin: NCCA.
- Ashton, H. S. *et al.* (2005): Investigating the medium effect in computer-aided assessment of school chemistry and college computing national examinations, *British Journal of Educational Technology*, Vol. 36, No. 5, 771-787.
- Baker, F. B. (2001): *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation.
- Бакoвљeв, М. (1997): *Основи методологије педагошких истраживања*. Београд: Научна књига.
- Банђур, В. и Н. Поткоњак (1999): *Методологија педагогије*. Београд: Савез педагошких друштава Југославије.
- Bangert-Drowns, R. L., J. A. Kulik & C. C. Kulik (1991): Effects of frequent classroom testing, *Journal of Educational Research*, Vol. 85, No. 2, 89-99.
- Banks, R. E. (1980): Some problems of evaluation in large school districts; in C. B. Stalford (ed.): *Testing and evaluation in schools: Practitioners' views* (93-100). Washington DC: National Institute of Education; U.S. Department of Education.
- Бауцал, А. и сар. (2006): *Образовна постигнућа ученика трећег разреда - национално тестирање 2004*. Београд: Министарство просвете и спорта Републике Србије, Завод за вредновање квалитета образовања и васпитања.

- Bejar, I. I. (1980): A procedure for investigating the unidimensionality of achievement tests based on item parameter estimates, *Journal of Educational Measurement*, Vol. 17, No. 4, 283-296.
- Бјекић, Д. и Ж. Папић (2005): *Оцењивање - приручник за оцењивање у средњем стручном образовању*. Београд: Министарство просвете и спорта Републике Србије.
- Бјекић, Д. и Ж. М. Папић (2006): *Тестови знања: израда и примена у средњој школи*. Чачак: Агенција "Пап".
- Bloom, B. (1981): *Taksonomija ili klasifikacija obrazovnih i odgojnih ciljeva (kognitivno područje)*. Beograd: Republički zavod za unapređivanje vaspitanja i obrazovanja.
- Богавац, Т. и сар. (1972): *Тест у школској пракси*. Београд: Научна књига.
- Brennan, R. L. (2001): Some problems, pitfalls, and paradoxes in educational measurement, *Educational Measurement: Issues and Practice*, Vol. 20, No. 4, 6-18.
- Bridges, D., P. Smeyers & R. Smith (2008): Educational Research and the Practical Judgement of Policy Makers, *Journal of Philosophy of Education*, Vol. 42, No. 1, 5-14.
- Бујас, З. (1959): *Основе психофизиологије рада (увод у индустријску психологију)*. Загреб: Институт за хигијену рада југословенске академије знаности и умјетности.
- Буквић, А. (1996): *Начела израде психолошких тестова*. Београд: Завод за уџбенике и наставна средства.
- Буквић, А. (2007): *Начела израде психолошких тестова*. Београд: Завод за уџбенике.
- Buser, K. (1996): *Basic precepts in test construction*. Washington DC: ERIC Clearinghouse microfiches.
- Van der Linden, W. J. (1982): Criterion-referenced Measurement: Its Main Applications, Problems and Findings, *Evaluation in Education*, Vol. 5, 97-118.
- Van der Linden, W. J. (2005): *Linear models for optimal test design*. Heidelberg: Springer.
- Van der Linden, W. J. & M. Jeon (2012): Modeling Answer Changes on Test Items, *Journal of Educational and Behavioral Statistics*, Vol. 37, No. 1, 180-199.
- Vansickle, T. R. (2008): *A testing brief: types of tests*. Minneapolis: Questar Assessment, Inc.
- Vaughn, K. W. (1981): Planiranje objektivnog testa; u E. F. Lindquist (ur.): *Pripremanje testa i priroda merenja u pedagogiji (prevod izabranih poglavlja iz knjige E. F. Lindquist "Educational Measurement")* (7-33). Beograd: Republički zavod za unapređivanje vaspitanja i obrazovanja.
- Vojdanovska, M., J. Cranney & B. R. Newell (2010): The testing effect: the role of feedback and collaboration in a tertiary classroom setting, *Applied Cognitive Psychology*, Vol. 24, No. 8, 1183-1195.
- Vu, P. (2008): Do state tests make the grade? *The Pew Charitable Trusts*, Jan 17th, 2008 (online edition).

- Вучић, Л. (1979): Критеријски тестови – нов прилаз тестирању знања, *Психологија*, Вол. 12, бр. 3-4, 26-35.
- Вујисић-Живковић, Н. (2013): Улога педагошких истраживања у трансформисању образовне политике и праксе, *Зборник Института за педагошка истраживања*, Вол. 45, бр. 1, 7-23.
- Wainer, H. (1989): The Future of Item Analysis, *Journal of Educational Measurement*, Vol. 26, No. 2, 191-208.
- Walker, K. (2006): *Teacher made exam designs (research brief)*. Southfield, Michigan: Education Partnerships, Inc.
- Ward, W. C., S. B. Carlson & E. Woisetschlaeger (1983): *Ill-structured problems as multiple-choice items (GRE Board Professional Report GREB No. 81-18P; ETS Research Report 83-6)*. Princeton: Educational Testing Service.
- Weiss, D. J. & M. L. Davison (1981): Test theory and methods, *Annual Review of Psychology*, Vol. 32, No.1, 629-658.
- Weiss, D. J. (2004): Computerized Adaptive Testing for Effective and Efficient Measurement in Counseling and Education, *Measurement and Evaluation in Counseling and Development*, Vol. 37, No. 2, 70-84.
- Weiss, D. J. (2011): Item Banking, Test Development, and Test Delivery; in K.F. Geisinger (ed.): *The APA Handbook on Testing and Assessment (185-200)*. Washington DC: American Psychological Association.
- Whitty, G. (2006). Education(al) Research and Education Policy Making: Is Conflict Inevitable? *British Educational Research Journal*, Vol. 32, No. 2, 159-176.
- Gallagher, C. J. (2003): Reconciling a Tradition of Testing with a New Learning Paradigm, *Educational Psychology Review*, Vol. 15, No. 1, 83-99.
- Гашић-Павишић, С., Д. Станковић и Д. Малинић (2011): TIMSS 2007 у Србији: опис истраживања и главни резултати; у С. Гашић-Павишић и Д. Станковић (ур.): *TIMSS 2007 у Србији (11-37)*. Београд: Институт за педагошка истраживања.
- Гашић-Павишић, С. (2011): TIMSS 2007 у Србији: објашњење постигнућа ученика и препоруке за побољшање наставе и учења; у С. Гашић-Павишић и Д. Станковић (ур.): *TIMSS 2007 у Србији (307-334)*. Београд: Институт за педагошка истраживања.
- Gierl, M., S. N. Khaliq & K. Boughton (1999): *Gender Differential Item Functioning in Mathematics and Science: Prevalence and Policy Implications*. Retrieved from: http://www2.education.ualberta.ca/educ/psych/crame/files/dif_csse99.pdf, Sep 3th, 2013.
- Goldstein, H. (2004): International comparisons of student attainment: some issues arising from PISA study, *Assessment in Education: Principles, Policy & Practice*, Vol. 11, No. 3, 319-330.
- Гојков, Г. (2012): *Документологија*. Вршац: Висока школа струковних студија за образовање васпитача “Михаило Палов”.
- Green, D. B. (1976): *Reducing bias in achievement tests*. Washington DC: ERIC.
- Гргин, Т. (2001): *Школско оцјењивање знања*. Јастребарско: Наклада Слап.
- Guide to Assessment* (2015): Glasgow: Scottish Qualifications Authority.

- Davies, I. K. (1976): *Objectives in Curriculum Design*. London: McGraw-Hill Book Company (UK) Limited.
- Деспотовић, М. (2010): *Развој курикулума у стручном образовању*. Београд: Филозофски факултет.
- Dihoff, R. E., G. M. Brosvic & M. L. Epstein (2003): The Role of Feedback During Academic Testing: The Delay Retention Effect Revisited, *The Psychological Record*, Vol. 53, 533-548
- Di Nisio, R. (2010): Measure school learning through Rasch Analysis: the interpretation of results, *Procedia Social and Behavioral Sciences*, Vol. 9, 373-377.
- Дубљанин, С. (2010): *Врсте и функција задатака у настави (докторска дисертација)*. Београд: Филозофски факултет.
- Earl, L. M. (2003): *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks: Corwin Press, Inc.
- Ebel, R. L. (1979): *Essentials of Educational Measurement*. London: Prentice-Hall International, Inc.
- Ebel, R. L. (1981): Pisanje ajtema testa; u E. F. Lindquist (ur.): *Pripremanje testa i priroda merenja u pedagogiji (prevod izabranih poglavlja iz knjige E. F. Lindquist "Educational Measurement")* (34-102). Београд: Републички завод за унапређивање васпитања и образовања.
- Ebel, R. L. & D. A. Frisbie (1991): *Essentials of Educational Measurement*. Englewood Cliffs: Prentice-Hall International, Inc.
- Ediger, M. (2000): *Qualitative assessment versus measurement of student achievement*. Washington DC: ERIC.
- Жешић, М. (1941): *Рјечник латинскога и хрватскога језика*. Загреб: Школска књига.
- Zimmerman, B. J., & M. K. DiBenedetto (2008): Mastery learning and assessment: implications for students and teachers in an era of high-stakes testing, *Psychology in the Schools*, Vol. 45, No. 3, 206-216.
- Илић, М. Ђ. (2012): Циљеви, задаци и исходи грађанског образовања; у С. Маринковић (ур.): *Настава и учење – циљеви, стандарди, исходи* (201-214). Ужице: Учитељски факултет.
- Илић-Голубовић, И. (2011): Квалитет и трајност знања ученика о неживој природи, *Иновације у настави*, Вол. 24, бр. 3, 86-95.
- Izard, J. (2005): *Overview of test construction*. Paris: International Institute for Educational Planning /UNESCO.
- Jacob, B. A., & S. Levitt (2004): To catch a cheat, *Education Next*, Vol. 4, No.1, 68-75.
- Jacobs, L. C. (2002): *How to write better tests: A handbook for improving test construction skills*. Retrieved from: http://www.indiana.edu/~best/pdf_docs/better_tests.pdf, March 4th, 2014.
- Kane, M. T. (2001): Current Concerns in Validity Theory, *Journal of Educational Measurement*, Vol. 38, No. 4, 319-342.

- Kang, T. (2006): *Model Selection Methods for Unidimensional and Multidimensional IRT Models (doctoral dissertation)*. Madison: University of Wisconsin - Madison.
- Каталог програма сталног стручног усавршавања наставника, васпитача и стручних сарадника за школску 2014/2015 и 2015/2016. Београд: Завод за унапређивање образовања и васпитања.
- Kean, M. H. (1980): Responding to conflicting evaluation demands; in C. B. Stalford (ed.): *Testing and evaluation in schools: Practitioners' views (77-86)*. Washington DC: National Institute of Education; U.S. Department of Education.
- Kennedy, D. (2009): *Pisanje i upotreba ishoda učenja (praktični vodič)*. Beograd: Savet Evrope, kancelarija u Beogradu.
- Куриасоу, С. (2001): *Темељна наставна умијећа*. Загреб: Educa.
- Ковач-Церовић, Т. и Љ. Левков (прир.) (2002): *Квалитетно образовање за све – пут ка развијеном друштву*. Београд: Министарство просвете и спорта Републике Србије.
- Konrad, H. S. (1981): Eksperimentalna provera testovnog materijala; u E. F. Lindquist (ur.): *Pripremanje testa i priroda merenja u pedagogiji (prevod izabranih poglavlja iz knjige E. F. Lindquist "Educational Measurement")*(103-120). Beograd: Republički zavod za unapređivanje vaspitanja i obrazovanja.
- Krathwohl, D. R. (2002): A Revision of Bloom's Taxonomy: An Overview, *Theory Into Practice*, Vol. 41, No. 4, 212-218.
- Кузмановић, Д. И Д. Павловић Бабић (2011): Приступи процењивању образовних постигнућа ученика: критички осврт, *Зборник Института за педагошка истраживања*, Вол. 43, бр. 1, 63-85.
- Кундачина, М. и В. Банђур (1997): Израда и употреба микротестова знања у разредној настави, *Педагогија*, Вол. 30, бр. 1-2, 53-64.
- Кундачина, М. (2012): Циљеви, задаци и исходи еколошког васпитања; у С. Маринковић (ур.): *Настава и учење – циљеви, стандарди, исходи* (191-200). Ужице: Учитељски факултет.
- Кундачина, М. и Ј. Стаматовић (2012): Акредитовани програми усавршавања наставника – стање и потребе, *Иновације у настави*, Вол. 25, бр. 1, 68-78.
- Lipnevich, A. A. & J. K. Smith (2008): *Response to Assessment Feedback: The Effects of Grades, Praise, and Source of Information (Research Report)*. Lawrence Township: Educational Testing Service.
- Lissitz, R. W & K. Samuelsen (2007): A Suggested Change in Terminology and Emphasis regarding Validity and Education, *Educational Researcher*, Vol. 36, No. 8, 437-448.
- Livingston, S. A. (2009): Constructed-Response Test Questions: Why We Use Them; How We Score Them, *R&D Connections*. Retrieved from: http://www.ets.org/Media/Research/pdf/RD_Connections11.pdf, Sep 3th, 2013.
- Lord, F. M. (1977): Optimal number of choices per item - a comparison of four approaches, *Journal of Educational Measurement*, Vol. 14, No. 1, 33-38.

- Максимовић, Ј. (2012): Методолошка образованост наставника – основа за прихватање и примену иновација. Преузето са: www.sao.org.rs, 2. 04.2016.
- Максимовић, А. (2013): *Конкретизација циљева васпитања и образовања кроз наставни програм и процес (докторска дисертација)*. Београд: Филозофски факултет.
- Максимовић, А. и М. Марковић (2012): Стандарди постигнућа ученика у васпитно-образовном процесу; у Н. Пантић и Ј. Чекић Марковић (ур): *Наставници у Србији: ставови о професији и о реформама у образовању*. Београд: Центар за образовне политике.
- Maravić Ćisar, S. (2011): *Evaluacija znanja programskih jezika C++ i Java primenom računarskih adaptivnih testova (doktorska disertacija)*. Zrenjanin: Tehnički fakultet "Mihajlo Pupin".
- Marzano, R. J. (2006): *Classroom assessment and grading that work*. Alexandria: Association for Supervision and Curriculum Development.
- Mathews, J. (2006): Just Whose Idea Was All This Testing? *The Washington Post*, Nov 14th, 2006 (online edition).
- Матовић, Н. (2007): *Мерење у педагошким истраживањима*. Београд: Институт за педагогију и андрагогију Филозофског Факултета.
- McDaniel, M. A., H. L. Roediger III & K. B. McDermott (2007): Generalizing test-enhanced learning from the laboratory to the classroom, *Psychonomic Bulletin & Review*, Vol. 14, No. 2, 200-206.
- Messick, S. (1975): The standard problem: Meaning and values in measurement and evaluation, *American Psychologist*, Vol. 30, No.10, 955-966.
- Messick, S. (1980): Test validity and the ethics of assessment, *American Psychologist*, Vol. 35, No. 11, 1012-1027.
- Messick, S. (1981): Evidence and ethics in the evaluation of tests, *Educational Researcher*, Vol. 10, No. 9, 9-20.
- Милановић-Наход, С. и Д. Јањетовић (2004): Наставни програми и успех ученика; у С. Милановић-Наход и Н. Шарановић-Божановић (прир.): *Знање и постигнуће (79-100)*. Београд: Институт за педагошка истраживања.
- Mislevy, R. J. & G.D. Haertel (2006): Implications of Evidence-Centered Design for Educational Testing, *Educational Measurement: Issues and Practice*, Vol. 2, No. 4, 6-20.
- Момировић, К., Б. Волф и Д. Поповић (1999): *Увод у теорију мерења и интерне метријске карактеристике композитних инструмената*. Приштина: Факултет за физичку културу.
- Moss, P. A. (2007): Reconstructing Validity, *Educational Researcher*, Vol. 36, No. 8, 470-476.
- Мужић, В. (1968): *Тестови знања*. Загреб: Школска књига.
- Мужић, В. (1986): *Методологија педагошког истраживања*. Сарајево: Завод за издавање уџбеника.
- Мужић, В. (2004): *Увод у методологију истраживања одгоја и образовања*. Загреб: Educa.

- Nastava usmerena na ishode, kompetencije i standarde* (2015). Belgrade: European Consulting Group & Human dynamics; Reading: CfBT.
- Natarajan, V. (2009): *Basic Principles of IRT And Application to Practical Testing & Assessment*. Retrieved from: www.merittrac.com, Aug 30th 2013.
- National Testing of Pupils in Europe: Objectives, Organisation and Use of Results* (2009). Brussels: Education, Audiovisual and Culture Executive Agency; Eurydice.
- Nikandrov, N. D. (2014): Reforming education since the year 2000: the case of Russia, *International Dialogues on Education: Past and Present*, Vol. 1, No. 1 (online journal).
- Nowak, S. (1977): *Methodology of Sociological Research*. Warszawa: PWN – Polish Scientific Publishers; Dordrecht/Boston: D. Reidel Publishing Company.
- Образовни стандарди за крај обавезног образовања* (2010). Београд: Министарство просвете Републике Србије и Завод за вредновање квалитета образовања и васпитања.
- Опште основе школског програма* (2003). Београд: Министарство просвете и спорта Републике Србије.
- Павловић-Бабић, Д. и сар. (2003): *Оцењивање оријентисано на исходе*. Београд: Министарство просвете и спорта Републике Србије; Центар за евалуацију.
- Пејић, А. и О. Годоровић (2007): *Национално тестирање ученика IV разреда*. Београд: Министарство просвете и спорта Републике Србије; Завод за вредновање квалитета образовања и васпитања.
- Petz, V. (1978): Penalizirati ili ne penalizirati pogrešne odgovore u testovima znanja alternativnog tipa, *Revija za psihologiju*, Vol. 8, br. 1-2, 303-310.
- Phelps, R. P. (2006): Characteristics of an Effective Student Testing System, *Educational Horizons*, Vol. 85, No.1, 19-29.
- Popham, W. J. (1999): Why standardized tests don't measure educational quality, *Educational Leadership: Using Standards and Assessments*, Vol. 56, No. 6, 8-15.
- Popham, W. J. (2001): *The Truth About Testing: An Educator's Call to Action*. Alexandria: Association for Supervision and Curriculum Development.
- Popham, W. J. (2003): *Test Better, Teach Better: The Instructional Role of Assessment*. Alexandria: ASCD.
- Postlethwaite, T. N. (2005): *Educational research: some basic concepts and terminology*. Paris: International Institute for Educational Planning / UNESCO.
- Правилник о оцењивању (2013). *Службени гласник РС*, бр. 72/2009, 52/11 и 55/13.
- Правилник о програму завршног испита у основном образовању и васпитању (2014). *Службени гласник РС – Просветни гласник*, бр. 1/2011, 1/2012, 1/2014 и 12/2014.
- Правилник о сталном стручном усавршавању и стицању звања наставника, васпитача и стручних сарадника (2012). *Службени гласник РС*, бр. 13/2012 и 31/2012.
- Професионални развој на нивоу школе/вртића* (2013). Подгорица: Завод за школство и Британски савјет.

- Рајстон, Ц. В, Ц. Цастман и И. Робинс (1966): *Вредновање у савременом образовању*. Београд: Вук Караџић.
- Ravitch, D. (2014): The myth of Chinese super schools (Who's Afraid of the Big Bad Dragon? Why China Has the Best (and Worst) Education System in the World – book review), *The New York Review of Books*, Vol. 61, No. 18 (online edition).
- Razzouk, R. (2011): *Using Evidence-Centered Design for Developing Valid Assessments of 21st Century Skills*. Retrieved from: www.edvation.com, Feb 5th, 2016.
- Reiner, C. M. et al. (2002): *Preparing Effective Essay Questions (A Self-Directed Workbook For Educators)*. Stillwater: New Forums Press.
- Реформа образовања у Републици Србији: школски програм – концепција, стратегија, имплементација (2003). Београд: Министарство просвете и спорта Републике Србије.
- Rivera, J. E. (2007): *Test Item Construction And Validation: Developing A Statewide Assessment For Agricultural Science Education (doctoral disertation)*. Ithaca: Cornell University.
- Roberson-Scott, K. (2005): *Evaluation of student learning: test construction and other practical strategies*. Retrieved from: www.docfoc.com/dr-kristi-roberson-scott, Aug 30th, 2013.
- Rochex, J. (2006): Social, methodological, and theoretical issues regarding assessment: lessons from a secondary analysis of PISA 2000 literacy tests, *Review of Research in Education (Special Issue on Rethinking Learning: What Counts as Learning and What Learning Counts)*, Vol. 30, 163-212.
- Rotberg, I. C. (2006): Assessment around the world, *Educational Leadership*, Vol. 64, No. 3, 58-63.
- Рубинштајн, С. Л. (1981): *О мишљењу и путевима његовог истраживања*. Београд: Завод за уџбенике и наставна средства.
- Rudner, L. & W. Schafer (2002): *What Teachers Need to Know About Assessment*. Washington, DC: National Education Association.
- Савовић, Б. и сар. (2006): *Примена тестова знања у основној и средњој школи*. Београд: Завод за вредновање квалитета образовања и васпитања.
- Sax, G. (1989): *Principles of Educational and Psychological Measurement and Evaluation*. Belmont: Wadsworth Publishing Company.
- Slideplayer.com/slide/271053, Jun 14th, 2016.
- Стаматовић, Ј. (2006): Стручно усавршавање као сегмент професионалног развоја наставника – процес и потребе, *Настава и васпитање*, Вол. 55, бр. 4, 473-482.
- Стојановић, А. (2011): Аморефлексивност васпитача као детерминанта методичких иновација у предшколском васпитању, *Иновације у настави*, Вол. 24, бр. 3, 114-121.
- Стојановић, Б. (2012): Наставник између циљева и задатака наставе књижевности и стандарда постигнућа ученика; у С. Маринковић (ур.): *Настава и учење – циљеви, стандарди, исходи* (395-408). Ужице: Учитељски факултет.

- Sharkness, J. & L. DeAngelo (2011): Measuring Student Involvement: A Comparison of Classical Test Theory and Item Response Theory in the Construction of Scales from Student Surveys, *Research in Higher Education*, Vol. 52, No. 5, 480-507.
- Slideplayer.com/slide/27105053, Mar 3th, 2016.
- Shepard, L. (1980): Technical issues in minimum competency testing, *Review of Research in Education*, Vol. 8, 30-82.
- Shepard, L. (1993): Evaluating test validity, *Review of Research in Education*, Vol. 19, 405-450.
- Spahiu, B. (1991): Razlika u uspehu učenika izražena školskim ocenama i testovima znanja, *Psihologija*, Vol. 24, br.1-2, 200-210.
- Stiggins, R. J. & N. J. Bridgeford (1985): The Ecology of Classroom Assessment, *Journal of Educational Measurement*, Vol. 22, No. 4, 271-286.
- Стратегија развоја образовања у Србији до 2020. године, *Службени гласник РС*, бр. 107/2012.
- Supporting Assessment in Schools – 3 (Standardised Testing in Compulsory Schooling)* (2005). Dublin: NCCA.
- Taylor, H, R. N. Greer & J. Mussio (1978): *Construction and use of classroom tests: a resource book for teachers*. Victoria: Learning Assessment Branch; Ministry of Education; Province of British Columbia.
- Thissen, D., L. Steinberg & A. R. Fitzpatrick (1989): Multiple-Choice Models: The Distractors Are Also Part of the Item, *Journal of Educational Measurement*, Vol. 26, No. 2, 161-176.
- Testing and Assessment (Third Report Of Session 2007-08)* (2008). London: House of Commons; Children, Schools and Families Committee.
- Testing in American Schools: Asking the Right Questions* (1992). Washington, DC: Government Printing Office.
- Traub, R. E. & R. G. Wolfe (1981): Latent trait theories and the assessment of educational achievement, *Review of Research in Education*, Vol.9, 377-435.
- Трој, Ф. (1957): *Прилог питању проверавања знања и оцењивања ученика у нашим школама*. Београд: Завод за унапређење школства у нас; Полит.
- Turner, S. M. et al. (2001): APA's guidelines for test user qualifications, *American Psychologist*, Vol. 56, No. 12, 1099-1113.
- Tyumeneva, J. (2013): *Disseminating and Using Student Assessment Information in Russia*. Washington DC: The International Bank for Reconstruction and Development.
- Фајгељ, С. и Б. Косановић (2001): Нова и стара ајтем анализа – поређење, *Психологија*, Вол. 34, бр. 1-2, 83-110.
- Фајгељ, С. (2003): *Психометрија: метод и теорија психолошког мерења*. Београд: Центар за примењену психологију.
- Фајгељ, С. (2004): *Метод истраживања понашања*. Београд: Центар за примењену психологију.
- Фајгељ, С. и Ј. Кнебл (2004): Да ли се може елиминисати погађање на пријемним испитима, *Психологија*, Вол. 37, бр. 2, 195-207.
- Фајгељ, С. (2005): *Психометрија: метод и теорија психолошког мерења*. Београд: Центар за примењену психологију.

- Fan, X. (1998): Item response theory and classical test theory: an empirical comparison of their item/person statistics, *Educational and Psychological Measurement*, Vol. 58, No. 3, 357-381.
- Fives, H. & N. DiDonato-Barnes (2013): Classroom Test Construction: The Power of a Table of Specifications, *Practical Assessment, Research & Evaluation*, Volume 18, No. 3, 1-7.
- Fluckiger, J. (2004): Teaching for Learning or Teaching for Testing: Classroom Teachers' Moral and Ethical Dilemma, *The Delta Kappa Gamma Buletin*, Summer 2004, 41-44.
- Frey, B. B. *et al.* (2005): Item-writing rules: collective wisdom, *Teaching & Teacher Education*, Vol. 21, No. 4, 357-364.
- Frey, B. B., & V. L. Schmitt (2007): Coming to terms with classroom assessment, *Journal of Advanced Academics*, Vol. 18, No. 3, 402-423.
- Haladyna, T. M. & S. M. Downing (1989): A Taxonomy of Multiple-Choice Item-Writing Rules, *Applied Measurement in Education*, Vol. 2, No. 1, 37-50.
- Haladyna, T. M. & G. H. Roid (1981). The role of instructional sensitivity in the empirical review of criterion-referenced test items, *Journal of Educational Measurement*, Vol. 18, No. 1, 39-53.
- Hambleton, R. K. & L. L. Cook (1977): Latent trait models and their use in the analysis of educational test data, *Journal of Educational Measurement*, Vol. 14, No. 2, 75-96.
- Hambleton, R. K. & S. Jirka (2004): How to do your best on standardized tests: some sugestions for adult learners, *Adventures in Assessment*, Vol.16, 5-11.
- Hamzah, M. S. G. & S. K. Abdullah (2011): Test item analysis: an educator professionalism approach, *US-China Education Review*, Vol. 11, No. 3, 307-322.
- Hargreaves, A. *et al.* (2001): *Learning to change: teaching beyond education and standards*. San Francisco: Jossey-Bass.
- Hattie, J., R. M. Jaeger & L. Bond (1999): Persistent methodological questions in educational testing, *Review of Research in Education*, Vol. 24, 393-446.
- Хавелка, Н. (2000): *Ученик и наставник у образовном процесу*. Београд: Завод за уџбенике и наставна средства.
- He, Q. & P. Tymms (2005): A computer-assisted test design and diagnosis system for use by classroom teachers, *Journal of Computer Assisted Learning*, Vol. 22, No. 6, 419-429.
- Hendrickson, A., K. Huff & R. Luecht (2010): Claims, evidence, and achievement-level descriptors as a foundation for item design and test specifications, *Applied Measurement in Education*, Vol. 23, No. 4, 358-377.
- Herman, J. L. & S. Golan (1991): *Effects of Standardized Testing on Teachers and Learning – Another Look (CSE Technical Report 334)*. Los Angeles: UCLA Graduate School of Education; National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Howe, K. R. (2004), A Critique of Experimentalism, *Qualitative Inquiry*, Vol. 10, No. 1, 42-61.
- Hubley, A. M. & B. D. Zumbo (2011): Validity and the consequences of test interpretation and use, *Social Indicators Research*, Vol. 103, No. 2, 219-230.

- <http://www.johnbiggs.com.au/academic/solo-taxonomy/>, Feb 11th, 2016.
- Catterall, J. S. (1989): Standards and school dropouts: A national study of tests required for high school graduation, *American Journal of Education*, Vol. 98, No. 1, 1-34.
- Chatterji, M. (2013): *Validity and Test Use: An International Dialogue on Educational Assessment, Accountability and Equity*. Retrieved from: http://www.emeraldinsight.com/promo/pdf/validity_foreword.pdf, Sep 19th, 2013.
- Cizek, G. J., D. Bowen & K. Church (2010): Sources of validity evidence for educational and psychological tests: a follow-up study, *Educational and Psychological Measurement*, Vol. 70, No. 5, 732-743.
- Cizek, G. J. (1998): *Filling in the Blanks: Putting Standardized Tests to the Test*. Washington DC: Thomas B. Fordham Institute.
- Clarke, M. M. *et al.* (2000): Retrospective on educational testing and assessment in the 20th century, *Journal of Curriculum Studies*, Vol. 32, No. 2, 159-181.
- Cohen, A. S. & J. A. Wollack (2012): *Handbook on test development: helpful tips for creating reliable and valid classroom tests*. Retrieved from: [testing.wisc.edu/Handbook on Test Construction.pdf](http://testing.wisc.edu/Handbook%20on%20Test%20Construction.pdf), July 27th, 2015.
- Cooper, M. & M. Leiter (1980): Teachers on testing; in C. B. Stalford (ed.): *Testing and evaluation in schools: Practitioners' views* (33-41). Washington DC: National Institute of Education; U.S. Department of Education.
- Crocker L. & J. Algina (1986): *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.
- Чапрић, Г. и сар. (2007): *Национално тестирање ученика IV разреда*. Београд: Завод за вредновање квалитета образовања и васпитања.
- Џордан, А. М. (1966): *Мерење у педагогији*. Београд: Вук Караџић.

ПРИЛОЗИ

ПРИЛОГ 1. ИНСТРУМЕНТИ ИСТРАЖИВАЊА

FILOZOFSKI FAKULTET
UNIVERZITET U BEOGRADU
Institut za pedagogiju i andragogiju

INSTRUMENT ZA NASTAVNIKE

Poštovani nastavnici,

U toku je istraživanje kojim se ispituje primena testova znanja u školskoj praksi. Testovi su postali deo naše prakse, bilo da ih sami konstruišemo, bilo da su deo eksterne evaluacije rada škole – završni ispiti, nacionalna i različita međunarodna ispitivanja učeničkih postignuća. Kao neposredni realizatori nastavnog procesa, smatramo da Vi najbolje možete da procenite koje su mogućnosti i ograničenja u primeni testova znanja u našoj školskoj praksi, i posebno u nastavi predmeta koji predajete.

Podaci dobijeni ispitivanjem biće korišćeni isključivo u naučne svrhe. Svi Vaši odgovori su nam podjednako važni.

Kod pitanja u kojima su ponuđeni odgovori, potrebno je da zaokružite odgovor koji najobjektivnije izražava Vaše mišljenje. Kod pitanja u kojima odgovori nisu ponuđeni, potrebno je da svoje mišljenje upišete u prostor predviđen za odgovor. Kod nekih pitanja, postoji dodatno uputstvo za davanje odgovora.

Upitnik je anoniman.

Zahvaljujemo se na saradnji!

1.

Pol: Naziv i mesto završene više škole/fakulteta:

a) ženski

b) muški

a) fakultet/viša škola: _____

b) mesto: _____

2. Navedite nazive predmeta koje predajete:

3. Navedite koliko imate godina radnog staža u prosveti:

4. Koje su prednosti testova znanja u odnosu na usmeno ispitivanje učenika?
(možete zaokružiti više odgovora)
- a) zahtevi su isti za sve učenike
 - b) kriterijum ocenjivanja je isti za sve učenike
 - c) postupak provere jednostavno se priprema
 - d) postupak provere jednostavno se realizuje na času
 - e) ocenjivanje odgovora je lako
 - f) obim gradiva koji se postupkom može obuhvatiti je veliki
 - g) broj učenika koji se postupkom može obuhvatiti je veliki
 - h) postupak se može više puta ponoviti sa istim pitanjima
 - i) učenici ne mogu da "varaju"
5. Koje su prednosti testova znanja u odnosu na pisane vežbe (vežbe koje sadrže pitanja esejskog tipa)?
(možete zaokružiti više odgovora)
- a) zahtevi su isti za sve učenike
 - b) kriterijum ocenjivanja je isti za sve učenike
 - c) postupak provere jednostavno se priprema
 - d) postupak provere jednostavno se realizuje na času
 - e) ocenjivanje odgovora je lako
 - f) obim gradiva koji se postupkom može obuhvatiti je veliki
 - g) broj učenika koji se postupkom može obuhvatiti je veliki
 - h) postupak se može više puta ponoviti sa istim pitanjima
 - i) učenici ne mogu da "varaju"
6. Koje su prednosti testova znanja u odnosu na elektronske testove?
(možete zaokružiti više odgovora)
- a) zahtevi su isti za sve učenike
 - b) kriterijum ocenjivanja je isti za sve učenike
 - c) postupak provere jednostavno se priprema
 - d) postupak provere jednostavno se realizuje na času
 - e) ocenjivanje odgovora je lako
 - f) obim gradiva koji se postupkom može obuhvatiti je veliki
 - g) broj učenika koji se postupkom može obuhvatiti je veliki
 - h) postupak se može više puta ponoviti sa istim pitanjima
 - i) učenici ne mogu da "varaju"
 - j) nisam koristio/la elektronske testove

7. Prema Vašem mišljenju, koji su osnovni nedostaci testova znanja u odnosu na:

a) usmeno ispitivanje _____

b) pisane vežbe _____

c) elektronske testove _____

8. Koliko često u toku školske godine koristite testove znanja?

- a) jednom u toku školske godine
- b) dva puta u toku školske godine
- c) tri puta u toku školske godine
- d) četiri puta u toku školske godine
- e) pet i više puta u toku školske godine
- f) ne koristim testove znanja u nastavi

9. Kada u toku školske godine zadajete test znanja učenicima?

(možete zaokružiti više odgovora)

- a) nakon obrade veće nastavne celine
- b) na kraju tromesečja
- c) na kraju polugodišta
- d) u nekom drugom trenutku: _____
- e) ne koristim testove znanja u nastavi

10. Procenite svoju pripremljenost za izradu testova znanja:

- a) potpuno sam pripremljen/a
- b) dosta sam pripremljen/a
- c) osrednje sam pripremljen/a
- d) malo sam pripremljen/a
- e) nisam pripremljen/a

11. Kako procenjujete sopstvenu pripremljenost za izradu testova znanja sa stanovišta potrebnih znanja i veština?

znanja potrebna za izradu testa znanja

veštine potrebne za izradu testa znanja

- a) potpuno sam pripremljen/a
- b) dosta sam pripremljen/a
- c) osrednje sam pripremljen/a
- d) malo sam pripremljen/a
- e) nisam pripremljen/a

- a) potpuno sam pripremljen/a
- b) dosta sam pripremljen/a
- c) osrednje sam pripremljen/a
- d) malo sam pripremljen/a
- e) nisam pripremljen/a

12. Da li ste se tokom studija pripremali za izradu testova znanja?

- a) ne
- b) da
- c) ne sećam se

13. Kakav status je imao predmet/predmeti u okviru koga/kojih je razmatrana tematika posvećena testovima znanja?

- a) obavezni studijski predmet
- b) izborni studijski predmet
- c) i obavezni i izborni studijski predmet
- d) ne sećam se
- e) nije bilo takvog predmeta

14. Koji od navedenih sadržaja su bili razmatrani u okviru tog/tih predmeta?
(možete zaokružiti više odgovora)

- a) pojam i vrste testova znanja
- b) mogućnosti primene testova znanja u nastavi
- c) proces izrade testa znanja
- d) vrste i karakteristike zadataka u testu znanja
- e) nešto drugo: _____
- f) ne sećam se
- g) nije bilo takvih sadržaja

15. Priprema za izradu testova znanja na studijama, doprinela je:

- a) usvajanju teorijskih znanja iz oblasti testiranja
- b) razvijanju veština potrebnih za izradu testova znanja
- c) usklađivanju teorijskih i praktičnih znanja
- d) nečemu drugom: _____
- e) ne sećam se
- f) nije bilo takve pripreme

16. Koliko je priprema tokom studija doprinela Vašoj sadašnjoj osposobljenosti za izradu testova znanja?

- a) veoma mnogo
- b) mnogo
- c) osrednje
- d) malo
- e) nimalo

17. Šta biste menjali kada je reč o pripremi za izradu testova znanja tokom studija na Vašem fakultetu?

18. U koje oblike stručnog usavršavanja ste bili uključeni kada je reč o izradi testova znanja?

(možete zaokružiti više odgovora)

- a) konsultovanje sa iskusnijim kolegama
- b) saradnja sa stručnim saradnikom u školi
- c) seminari stručnog usavršavanja
- d) čitanje relevantne metodološke literature
- e) učešće na stručnom skupu
- f) učešće na letnjoj/zimskoj školi
- g) boravak na stručnom/studijskom putovanju
- h) neki drugi oblik: _____
- i) nemam takvo iskustvo

19. Koliko akreditovanih programa stručnog usavršavanja (seminara) ste pohađali iz domena izrade testova znanja?

- a) nijedan
- b) jedan
- c) dva
- d) tri
- e) više od tri

20. Koji sadržaji su bili obuhvaćeni seminarom/seminarima?

(možete zaokružiti više odgovora)

- a) pojam i vrste testova znanja
- b) mogućnosti primene testova znanja u nastavi
- c) proces izrade testa znanja (razrada sadržaja testa, sastavljanje zadataka, ocenjivanje testa i slično)
- d) vrste i karakteristike zadataka u testu znanja
- e) nešto drugo: _____
- f) nisam pohađao/la takav seminar

21. Pohađanje seminara doprinelo je:

(možete zaokružiti više odgovora)

- a) usvajanju teorijskih znanja iz oblasti testiranja
- b) razvijanju veština potrebnih za izradu testova znanja
- c) usklađivanju teorijskih i praktičnih znanja
- d) nečemu drugom: _____
- e) nisam pohađao/la takav seminar

22. Na čiju inicijativu ste pristupili pohađanju seminara?

- a) na ličnu inicijativu
- b) stručnog aktiva nastavnika
- c) stručnih saradnika škole
- d) direktora škole
- e) Zavoda za unapređivanje obrazovanja i vaspitanja/Ministarstva prosvete
- f) nekog drugog: _____
- g) nisam pohađao/la takav seminar

23. Ko je finansirao seminar/e?

(možete zaokružiti više odgovora)

- a) Ministarstvo prosvete
- b) škola
- c) opština
- d) lokalna zajednica
- e) zainteresovani nastavnici
- f) neko drugi: _____
- g) nisam pohađao/la takav seminar

24. Koliko su navedeni oblici stručnog usavršavanja doprineli Vašem usavršavanju u procesu izrade testova znanja?

oblik stručnog usavršavanja		doprinosi usavršavanju nastavnika u procesu izrade testova znanja					
		veoma mnogo	mного	osrednje	malo	nimalo	nemam takvo iskustvo
1	pohađanje seminara	A	B	C	D	E	F
2	učešće na stručnom skupu	A	B	C	D	E	F
3	saradnja sa kolegama	A	B	C	D	E	F
4	saradnja sa stručnim saradnikom	A	B	C	D	E	F
5	praćenje stručne literature	A	B	C	D	E	F

Upišite redni broj oblika koji je najviše doprineo Vašem usavršavanju u oblasti izrade testova znanja: _____

25. Na koji način se, prema Vašem mišljenju, može unaprediti stručno usavršavanje nastavnika u domenu izrade testova znanja?

26. Da li ste do sada samostalno izrađivali testove za proveru znanja učenika?

- a) da
- b) ne

Ukoliko je Vaš odgovor “da”, koliko često ste to činili?

- a) jednom u toku školske godine
- b) dva puta u toku školske godine
- c) tri puta u toku školske godine
- d) četiri puta u toku školske godine
- e) pet i više puta u toku školske godine

Ukoliko je Vaš odgovor “ne”, zbog čega to niste činili?

- a) test nije mera znanja
- b) u nastavi predmeta koji predajem nije pogodno koristiti test znanja
- c) koristim testove koje je neko drugi pripremio
- d) pravljenje testa je komplikovano i oduzima puno vremena
- e) nemam iskustvo u pravljenju testova znanja
- f) nešto drugo: _____

27. Prilikom sastavljanja testova znanja, teško mi je da:
(možete zaokružiti više odgovora)

- a) razradim sadržaj testa
- b) odredim sadržaj zadataka u odnosu na ishode, standarde
- c) odredim nivo znanja koji zadatak meri (poznavanje činjenica, razumevanje, primena i drugi)
- d) odredim težinu zadatka
- e) izaberem adekvatnu vrstu zadataka za određeni sadržaj
- f) odredim dužinu testa
- g) odredim redosled zadataka u testu
- h) odredim broj poena koje zadatak nosi
- i) prevedem poene na testu u ocene
- j) sastavim uputstvo za rešavanje testa
- k) realizujem probno ispitivanje
- l) razvijam i čuvam bazu zadataka
- m) nešto drugo: _____
- n) nemam teškoća

28. Na koji način rešavate navedene teškoće?
(možete zaokružiti više odgovora)

- a) konsultovanjem sa iskusnijim kolegama
- b) u saradnji sa pedagogom škole
- c) u saradnji sa psihologom škole
- d) samostalno, putem “pokušaja i pogrešaka”
- e) čitanjem relevantne metodološke literature
- f) na neki drugi način: _____
- g) nemam teškoća

29. Sa kojim ciljem ste Vi lično do sada koristili rezultate testova koje pravite za potrebe nastave?

(možete zaokružiti više odgovora)

- a) utvrđivanje predznanja koje učenici imaju u vezi sa određenim delom gradiva
- b) planiranje i preduzimanje mera podrške učenicima u učenju
- c) ocenjivanje učenika
- d) planiranje sadržaja i načina rada u nastavi
- e) organizovanje rada u dopunskoj nastavi
- f) organizovanje rada u dodatnoj nastavi
- g) razmena informacija o napredovanju učenika sa roditeljima i razrednim starešinom
- h) planiranje stručnog usavršavanja
- i) unapređivanje procesa izrade testa znanja
- j) neki drugi cilj: _____
- k) nemam takvo iskustvo

30. Koja od navedenih testiranja su realizovana u Vašoj školi?

(možete zaokružiti više odgovora)

- a) PISA 2003
- b) PISA 2006
- c) PISA 2009
- d) PISA 2012
- e) TIMSS 2003
- f) TIMSS 2007
- g) TIMSS 2011
- h) Nacionalno testiranje učenika 2003-2004
- i) Nacionalno testiranje učenika 2005-2006
- j) neko drugo: _____
- k) nisam siguran/sigurna

31. Da li se i na koji način u Vašoj školi koriste rezultati dobijeni primenom određenih testova znanja?

vrsta testiranja	korišćenje rezultata na nivou škole
Mala matura	a) ne
	b) nisam siguran/sigurna
	c) da, koriste se za:
nacionalna testiranja	a) ne
	b) nisam siguran/sigurna
	c) da, koriste se za:
međunarodna testiranja (PISA, TIMSS)	a) ne
	b) nisam siguran/sigurna
	c) da, koriste se za:

32. Prema Vašem mišljenju, koliko su rezultati Male mature značajni za unapređivanje rada škole?

rezultati Male mature					
korišćenje rezultata testiranja za:	procena značajnosti				
	veoma je značajno	značajno je	neodlučan/ neodlučna sam	malo je značajno	nije značajno
unapređivanje sadržaja nastave	A	B	C	D	E
unapređivanje načina rada u nastavi	A	B	C	D	E
unapređivanje procesa ocenjivanja	A	B	C	D	E
unapređivanje vannastavnog rada sa učenicima	A	B	C	D	E
planiranje stručnog usavršavanja nastavnika	A	B	C	D	E
poboljšanje međuljudskih odnosa u školi	A	B	C	D	E
poboljšanje opremljenosti škole	A	B	C	D	E
prevazilaženje razlika između dece različitog pola i porodičnog porekla	A	B	C	D	E

33. Prema Vašem mišljenju, u kojoj meri rezultati nacionalnih testiranja mogu biti značajni za unapređivanje rada škole? Ukoliko Vam nije poznata koncepcija nacionalnih testiranja, molimo Vas da taj podatak upišete ispod tabele.

rezultati nacionalnih testiranja					
korišćenje rezultata testiranja za:	procena značajnosti				
	veoma je značajno	značajno je	neodlučan/ neodlučna sam	malo je značajno	nije značajno
unapređivanje sadržaja nastave	A	B	C	D	E
unapređivanje načina rada u nastavi	A	B	C	D	E
unapređivanje procesa ocenjivanja	A	B	C	D	E
unapređivanje vanmnastavnog rada sa učenicima	A	B	C	D	E
planiranje stručnog usavršavanja nastavnika	A	B	C	D	E
poboljšanje međuljudskih odnosa u školi	A	B	C	D	E
poboljšanje opremljenosti škole	A	B	C	D	E
prevazilaženje razlika između dece različitog pola i porodičnog porekla	A	B	C	D	E

34. Prema Vašem mišljenju, u kojoj meri rezultati međunarodnih testiranja mogu biti značajni za unapređivanje rada škole? Ukoliko Vam nije poznata koncepcija međunarodnih testiranja, molimo Vas da taj podatak upišete ispod tabele.

rezultati međunarodnih testiranja					
korišćenje rezultata testiranja za:	procena značajnosti				
	veoma je značajno	značajno je	neodlučan/ neodlučna sam	malo je značajno	nije značajno
unapređivanje sadržaja nastave	A	B	C	D	E
unapređivanje načina rada u nastavi	A	B	C	D	E
unapređivanje procesa ocenjivanja	A	B	C	D	E
unapređivanje van nastavnog rada sa učenicima	A	B	C	D	E
planiranje stručnog usavršavanja nastavnika	A	B	C	D	E
poboljšanje međuljudskih odnosa u školi	A	B	C	D	E
poboljšanje opremljenosti škole	A	B	C	D	E
prevazilaženje razlika između dece različitog pola i porodičnog porekla	A	B	C	D	E

Prostor za Vaša pitanja i sugestije

Hvala!

ПРОТОКОЛ СТРУКТУРИРАНОГ ИНТЕРВЈУА СА НАСТАВНИЦИМА

Датум: _____ 2015.

Време: _____

Место: _____

Школа: _____

Предмет који наставник предаје:

Назив и место завршене више школе/факултета:

1) српски језик

факултет/виша школа: _____

2) математика

место: _____

Пол:

Године радног стажа у просвети:

a) женски

b) мушки

1. Шта је циљ теста?

- утврђивање предзнања које ученици имају у вези са одређеним делом градива
- праћење ученичких постигнућа
- оцењивање ученика
- праћење и вредновање ефикасности сопствених наставних метода/облика рада
- утврђивање који задаци у тесту су добри, а који лоши
- нешто друго: _____

2. Шта је била полазна основа за разраду садржаја конкретног теста знања? (могуће је више одговора)

- садржај наставног програма
- садржај уџбеника
- стандарди знања дефинисани за наставни предмет
- нешто друго: _____

3. Колики број задатака је изведен из дате основе? На који део градива се задатак односи?

бр. задатка	основа за разраду садржаја конкретних задатака				
	садржај наставног програма	садржај уџбеника	стандарди знања дефинисани за наставни предмет	нешто друго	на који део градива се задатак односи?
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
укупно					

4. Који ниво знања мери конкретан задатак?

бр. задатка	ниво знања који задатак мери			
	чињенице	разумевање	примена	неки други ниво
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
укупно				

5. Коју тежину има конкретан задатак?

бр.зadatка	степен тежине задатка		
	лак задатак	средње тежак задатак	тежак задатак
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
укупно			

6. Да ли ученике припремате за тестирање?

- да
- не

У случају потврдног одговора, утврдити на који начин се то остварује.

7. Да ли су предвиђена упутства за решавање теста намењена ученицима?

- да
- не

У случају потврдног одговора, утврдити:

- Каква је форма упутства за ученике?
 - усмена
 - писана
 - и писана и усмена

У случају да је предвиђено упутство за ученике, утврдити елементе његовог садржаја:

- *просторна и временска организација тестирања*
 - распоред седења
 - средства која ученик може користити у току рада
 - време израде теста
 - договор око начина предаје попуњеног теста – у било ком тренутку у току часа када се рад заврши; окретањем листа и чекањем да се тестирање оконча
- *упутство у вези са понашањем ученика у току тестирања*
 - недозвољена понашања
 - могућност да се затражи помоћ
- *указивање на начин решавања појединих типова задатака*
- *нешто друго:* _____

8. Да ли постоји кључ за оцењивање теста?

- да
- не

9. Колико поена носе појединачни задаци?

бр.здатка	број поена које задатак носи
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

10. Да ли је предвиђено да се бодују и делимично тачни одговори? Колико поена носи делимично тачан задатак?

бр. задатка	да ли задатак може бити делимично тачан?	број поена за делимично тачан одговор
1	да не	
2	да не	
3	да не	
4	да не	
5	да не	
6	да не	
7	да не	
8	да не	
9	да не	
10	да не	
11	да не	
12	да не	
13	да не	
14	да не	
15	да не	
16	да не	
17	да не	
18	да не	
19	да не	
20	да не	

11. Да ли су предвиђени негативни поени?

- да
- не

У случају потврдног одговора, евидентирати колики је износ негативних поена за поједине задатке.

бр.зadatка	број негативних поена
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

12. Како се утврђује укупан број поена на тесту?

- сумирањем свих поена (уколико нема негативних поена)
- сумирањем свих позитивних поена и одузимањем суме негативних поена од укупног збира
- на неки други начин: _____

13. Колико се поена максимално може остварити на тесту?
(самостално израчунати на основу питања бр. 9)

14. Скала превођења поена у оцене:

распон поена	оцена
	5
	4
	3
	2

15. Да ли сте, пре финалне примене теста, проверили квалитет теста?

- да
- не

У случају потврдног одговора:

- Описати на који начин је реализована дата провера квалитета теста:
 - спровођењем пробног испитивања
 - консултовањем са другим наставницима
 - консултовањем са стручним сарадником у школи
 - на неки други начин: _____

- Са којим циљем?
 - провера тежине задатака
 - провера јасноће упутства
 - провера формулације задатака
 - неки други циљ: _____

16. Да ли формирате своју базу задатака?

- да, формирам електронску базу задатака
- да, чувам штампане верзије тестова
- не

У случају потврдног одговора, утврдити:

- Колико често базу допуњујете новим задацима/састављате нови тест?
 - једанпут годишње
 - на крају полугодишта
 - на тромесечју
 - скоро сваког месеца
 - једном месечно
 - _____

- Како базу осигуравате од евентуалних злоупотреба?
 - чувањем електронске верзије теста на личном рачунару
 - чувањем електронске верзије теста на заједничком рачунару под шифром
 - чувањем штампане верзије тестова ван школе
 - чувањем штампане верзије тестова у школи, али у простору који се закључава
 - остали начини: _____

17. Да ли користите тестове или задатке које је саставио неко други?

- да
- не

У случају потврдног одговора, утврдити:

- Ко саставља те задатке/тестове?
 - друге колеге
 - тимови стручњака које окупља издавачка кућа _____
 - чланови интернет форума _____
 - неко други: _____

- Колико често користите тестове/задатке које саставља неко други?
 - једном годишње
 - једном у току полугодишта
 - сваког тромесечја
 - скоро сваког месеца
 - једном месечно
 - увек
 - _____

18. Да ли анализирате постигнуће ученика на нивоу појединачних задатака?

- да
- не

У случају потврдног одговора, утврдити:

- Шта је сврха таквих анализа?
 - похрањивање у базу оних задатака који су се показали јасним и дискриминативним
 - преправљање задатака
 - искључивање лоших задатака из базе – нејасних, превише тешких или превише лаких задатака
 - нешто друго: _____

ФИЛОЗОФСКИ ФАКУЛТЕТ
УНИВЕРЗИТЕТ У БЕОГРАДУ
Институт за педагогију и андрагогију

ФАЗЕ У ИЗРАДИ ТЕСТА ЗНАЊА

Опишите процес израде теста знања који састављате за потребе наставе.

ФИЛОЗОФСКИ ФАКУЛТЕТ
УНИВЕРЗИТЕТ У БЕОГРАДУ
Институт за педагогију и андрагогију

ПРОТОКОЛ АНАЛИЗЕ САДРЖАЈА ТЕСТА ЗНАЊА

Датум: _____ 2015.

Време: _____

Место: _____

Школа: _____

Наставник: _____

Тест знања из:

- 1) српског језика
- 2) математике

1. Редослед и укупан број задатака у тесту:

- редослед задатака према форми и њихов укупан број

број задатка	форма задатка							
	двочлани избор	вишечлани избор	допуњавање	испуњавање налога	кратки одговори	спаривање	сређивање	нека друга форма
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
укупно								

Нека друга форма задатка: ред.број задатка: _____ форма: _____

ред.број задатка: _____ форма: _____

ред.број задатка: _____ форма: _____

- редослед задатака према нивоу знања који задатак мери и њихов укупан број (утврђује се на основу интервјуа)
- редослед задатака према тежини и њихов укупан број (утврђује се на основу интервјуа)

2. Уколико постоје задаци вишечланог избора, утврдити колико је у њима понуђено одговора, и колико је предвиђено тачних одговора:

редни број задатка	број понуђених одговора	број тачних одговора

3. Уколико постоје задаци спаривања, утврдити колико је понуђено одговора у одговарајућим колонама у задатку и колики је број тражених парова:

редни број задатка	број понуђених одговора у једној колони	број понуђених одговора у другој колони	број тражених парова

Више од две колоне: ред.број задатка: ____ број колоне: ____
 ред.број задатка: ____ број колоне: ____
 ред.број задатка: ____ број колоне: ____

5. Форма задатака којима се испитује одређени ниво усвојености знања:

форма задатка	ниво знања који се испитује			
	ниво познавања чињеница	ниво разумевања градива	ниво примене знања	неки други ниво
задаци двочланог избора				
задаци вишечланог избора				
задаци допуњавања				
задаци испуњавања налога				
задаци са кратким одговорима				
задаци спаривања				
задаци сређивања				
нека друга форма				

6. Форма задатака одређене тежине:

форма задатка	тежина задатка		
	лаки задаци	средње тешки задаци	тешки задаци
задаци двочланог избора			
задаци вишечланог избора			
задаци допуњавања			
задаци испуњавања налога			
задаци са кратким одговорима			
задаци спаривања			
задаци сређивања			
нека друга форма			

7. Форма задатака и глаголи и глаголски облици примењени у њиховој формулацији:

форма задатка	глаголи и глаголски облици		
задаци двочланог избора			
задаци вишечланог избора			
задаци допуњавања			
задаци испуњавања налога			
задаци са кратким одговорима			
задаци спаривања			
задаци сређивања			
нека друга форма			

ПРИЛОГ 2. ЛИСТА ГЛАГОЛА КОЈИМА СЕ ДЕФИНИШУ ОБРАЗОВНИ ЦИЉЕВИ У БЛУМОВОЈ ТАКСОНОМИЈИ¹

Знање	Разумевање	Примена	Анализа	Синтеза	Евалуација
описати	превести	уопштити	разликовати	креирати	судити
навести	организовати	решити	потврдити	дизајнирати	доказивати
изрећи	скратити	персонализовати	скицирати	измислити	оповрћи
поновити	симболички изразити	мислити на другачији начин	направити дијаграм	елаборирати	проценити
препознати	пренети	ревидирати	посматрати	иницирати	дебатовати
идентификовати	дефинисати	урадити на другачији начин	категорисати	створити	проценити тежину
именовати	интерпретирати	пројектовати	истраживати	представити сликом	дискутовати
регистровати	закључити	проширити	разграничити	замислити	разрешити нејасноћу
цитирати	предвидети	кодирати	организовати	модификовати	оценити
лоцирати	објаснити		класификовати	кодирати	вредновати
	демонстрирати			дефинисати	мерити
	појаснити			претпоставку	
				предвидети	направити приоритете
				комбиновати	истраживати
				минимизирати	
				максимизирати	
				адаптирати	

¹ Извор: Павловић-Бабић и сар., 2003

ПРИЛОГ 3. ТАБЕЛАРНИ ПРИКАЗ РЕЗУЛТАТА ИСТРАЖИВАЊА

Табела 1. Глобална процена припремљености за израду тестова знања

Процена	f	%
потпуно сам припремљен/а	98	27,5
доста сам припремљен/а	188	52,7
осредње сам припремљен/а	55	15,4
мало сам сам припремљен/а	11	3,1
нисам сам припремљен/а	5	1,4
укупно	357	100,0
$\chi^2 = 342,97; df = 4; p = 0.00$		

Табела 2. Похађање припреме за израду тестова знања током студија

Похађање припреме	f	%
да	186	56,2
не	145	43,8
укупно	331	100,0
$\chi^2 = 5,08; df = 1; p = 0.02$		

Табела 3. Статус студијског предмета у оквиру кога је организована припрема за израду тестова знања

Статус предмета	f	%
обавезни	117	88,6
изборни	11	8,3
и обавезни и изборни	4	3,0
укупно	132	100,0
$\chi^2 = 182,23; df = 2; p = 0.00$		

Табела 4. Самостална израда тестова знања у досадашњем раду наставника

Самостална израда	f	%
да	341	94,5
не	20	5,5
укупно	361	100,0
$\chi^2 = 285,43; df = 1; p = 0.00$		

Табела 5. Године радног стажа наставника у просвети и предности тестова знања у односу на електронске тестове с обзиром на истоветност захтева који се постављају ученицима

Предности тестова знања у односу на електронске тестове : истоветност захтева који се постављају ученицима						
Године радног стажа	не		да		укупно	
	f	%	f	%	f	%
0-10	134	87,1	20	13,0	154	100,0
10-20	118	83,1	24	16,9	142	100,0
20 и више	48	72,7	18	27,3	66	100,0
укупно	300	82,9	62	17,1	362	100,0

$\chi^2 = 6,65; df = 2; p = 0.04; C = 0,13$

Табела 6. Године радног стажа наставника у просвети и предности тестова знања у односу на електронске тестове с обзиром на истоветност критеријума оцењивања

Предности тестова знања у односу на електронске тестове: истоветност критеријума оцењивања						
Године радног стажа	не		да		укупно	
	f	%	f	%	f	%
0-10	128	83,1	26	16,9	154	100,0
10-20	119	83,8	23	16,2	142	100,0
20 и више	45	68,2	21	31,9	66	100,0
укупно	292	80,7	70	19,3	362	100,0

$\chi^2 = 8,08; df = 2; p = 0.02; C = 0,15$

Табела 7. Године радног стажа наставника у просвети и разматрање садржаја који се односе на појам и врсте тестова знања током припреме на студијама

Разматрање садржаја који се односе на појам и врсте тестова знања током припреме на студијама						
Године радног стажа	не		да		укупно	
	f	%	f	%	f	%
0-10	96	62,3	58	37,7	154	100,0
10-20	112	78,9	30	21,1	142	100,0
20 и више	55	83,3	11	16,7	66	100,0
укупно	263	72,7	99	27,3	362	100,0

$$\chi^2 = 14,80; df = 2; p = 0.01; C = 0,20$$

Табела 8. Године радног стажа наставника у просвети и разматрање садржаја који се односе на врсте и карактеристике задатака у тесту знања

Разматрање садржаја који се односе на врсте и карактеристике задатака у тесту знања						
Године радног стажа	не		да		укупно	
	f	%	f	%	f	%
0-10	113	73,4	41	26,6	154	100,0
10-20	121	85,2	21	14,8	142	100,0
20 и више	61	92,4	5	7,6	66	100,0
укупно	295	81,5	67	18,5	362	100,0

$$\chi^2 = 13,26; df = 2; p = 0.01; C = 0,19$$

Табела 9. Године радног стажа наставника у просвети и број семинара посвећених изради тестова знања које је наставник похађао

Године радног стажа	Број семинара											
	ниједан		један		два		три		више од три		укупно	
	f	%	f	%	f	%	f	%	f	%	f	%
0-10	84	54,9	41	26,8	12	7,8	6	3,9	10	6,5	153	100,0
10-20	66	46,5	39	27,5	22	15,5	5	3,5	10	7,0	142	100,0
20 и више	18	27,7	24	36,9	9	13,8	2	3,1	12	18,5	65	100,0
укупно	168	46,7	104	28,9	43	11,9	13	3,6	32	8,9	360	100,0

$$\chi^2 = 21,17; df = 8; p < 0.01; C = 0,24$$

Табела 10. Године радног стажа наставника у просвети и разматрање садржаја који се односе на појам и врсте тестова знања у оквиру семинара

Године радног стажа	Разматрање садржаја који се односе на појам и врсте тестова знања у оквиру семинара					
	не		да		укупно	
	f	%	f	%	f	%
0-10	128	83,1	26	16,9	154	100,0
10-20	102	71,8	40	28,2	142	100,0
20 и више	41	62,1	25	37,9	66	100,0
укупно	271	74,9	91	25,1	362	100,0

$$\chi^2 = 12,00; df = 2; p = 0.00; C = 0,18$$

Табела 11. Године радног стажа наставника у просвети и разматрање садржаја који се односе на процес израде теста знања у оквиру семинара

		Разматрање садржаја који се односе на процес израде теста знања у оквиру семинара					
Године радног стажа	не		да		укупно		
	f	%	f	%	f	%	
0-10	115	74,7	39	25,3	154	100,0	
10-20	92	64,8	50	35,2	142	100,0	
20 и више	37	56,1	29	43,9	66	100,0	
укупно	244	67,4	118	32,6	362	100,0	

$$\chi^2 = 8,01; df = 2; p = 0.02; C = 0,15$$

Табела 12. Године радног стажа наставника у просвети и учешће на летњој/зимској школи на којој је разматрана проблематика тестова знања

		Учешће на летњој/зимској школи на којој је разматрана проблематика тестова знања					
Године радног стажа	не		да		укупно		
	f	%	f	%	f	%	
0-10	151	98,1	31	2,0	182	100,0	
10-20	131	92,3	11	7,7	142	100,0	
20 и више	59	89,4	7	10,6	66	100,0	
укупно	341	94,2	21	5,8	362	100,0	

$$\chi^2 = 8,00; df = 2; p = 0.02; C = 0,15$$

Табела 13. Године радног стажа наставника у просвети и решавање тешкоћа у изради тестова знања путем консултација са искуснијим колегама

		Решавање тешкоћа у изради тестова знања путем консултација са искуснијим колегама					
		не		да		укупно	
Године радног стажа		f	%	f	%	f	%
0-10		74	48,1	80	51,9	154	100,0
10-20		75	52,8	67	47,2	142	100,0
20 и више		45	68,2	21	31,8	66	100,0
укупно		194	53,6	168	46,4	362	100,0

$\chi^2 = 7,58$; $df = 2$; $p = 0.02$; $C = 0,14$

Табела 14. Године радног стажа наставника у просвети и избор одговора „не сећам се“ на питање које се односи на садржаје посвећене тестовима знања који су разматрани током припреме на студијама

		Избор одговора „не сећам се“					
		не		да		укупно	
Године радног стажа		f	%	f	%	f	%
0-10		140	90,9	14	9,1	154	100,0
10-20		124	87,3	18	12,7	142	100,0
20 и више		48	72,7	18	27,3	66	100,0
укупно		312	86,2	50	13,8	362	100,0

$\chi^2 = 13,08$; $df = 2$; $p = 0.01$; $C = 0,19$

Табела 15. Предмет који наставник предаје и садржај уџбеника као основа за разраду садржаја теста знања

Предмет	Садржај уџбеника као основа за разраду садржаја теста знања					
	не		да		укупно	
	f	%	f	%	f	%
српски језик	19	47,5	21	52,5	40	50,0
математика	33	82,5	7	17,5	40	50,0
укупно	52	65,0	28	35,0	80	100,0

$\chi^2 = 10,77; df = 1; p = 0.00; C = 0,34$

Табела 16. Предмет који наставник предаје и укупан број форми задатака на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	3,73	1,519	0,240
математика	40	1,58	0,931	0,147

$t = 7,63; df = 64,7; p = 0.00$

Табела 17. Предмет који наставник предаје и укупан број задатака вишеструког избора на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	1,11	1,886	0,298
математика	40	0,29	1,905	0,143

$t = 2,48; df = 56,1; p = 0.02$

Табела 18. Предмет који наставник предаје и укупан број задатака допуњавања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,85	0,965	0,153
математика	40	0,20	0,723	0,114
$t = 3,40; df = 72,3; p = 0.00$				

Табела 19. Предмет који наставник предаје и укупан број задатака са кратким одговорима на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	2,65	2,495	0,394
математика	40	0,34	0,916	0,145
$t = 5,51; df = 49,3; p = 0.00$				

Табела 20. Предмет који наставник предаје и укупан број задатака спаривања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,45	0,677	0,107
математика	40	0,05	0,221	0,035
$t = 3,55; df = 47,2; p = 0.00$				

Табела 21. Предмет који наставник предаје и укупан број задатака „других“ форми на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,81	0,943	0,149
математика	40	0,05	0,221	0,035
$t = 4,95; df = 43,3; p = 0.00$				

Табела 22. Предмет који наставник предаје и укупан број задатака који испитују ниво познавања чињеница на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	3,93	2,710	0,429
математика	40	2,52	1,506	0,238
$t = 2,88; df = 61,0; p < 0.01$				

Табела 23. Предмет који наставник предаје и укупан број лаких задатака на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	3,54	2,011	0,318
математика	40	2,31	1,602	0,253
$t = 3,02; df = 74,3; p = 0.00$				

Табела 24. Предмет који наставник предаје и укупан број средње тешких задатака на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	3,72	2,174	0,344
математика	40	2,78	1,748	0,276
$t = 2,14; df = 74,6; p = 0.04$				

Табела 25. Предмет који наставник предаје и укупан број задатака допуњавања који испитују ниво познавања чињеница на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,48	0,615	0,097
математика	40	0,13	0,516	0,082
$t = 2,86; df = 75,7; p < 0.01$				

Табела 26. Предмет који наставник предаје и укупан број задатака допуњавања који испитују ниво примене знања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,15	0,438	0,069
математика	40	0,00	0,000	0,000
$t = 2,09; df = 39,0; p = 0.04$				

Табела 27. Предмет који наставник предаје и укупан број задатака са кратким одговорима који испитују ниво познавања чињеница на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,90	1,212	0,192
математика	40	0,29	0,750	0,119
$t = 2,70; df = 65,1; p < 0.01$				

Табела 28. Предмет који наставник предаје и укупан број задатака са кратким одговорима који испитују ниво разумевања градива на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	1,00	1,212	0,192
математика	40	0,05	0,221	0,035
$t = 4,82; df = 41,6; p = 0.00$				

Табела 29. Предмет који наставник предаје и укупан број задатака са кратким одговорима који испитују ниво примене знања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,71	1,021	0,161
математика	40	0,00	0,000	0,000
$t = 4,38; df = 39,0; p = 0.00$				

Табела 30. Предмет који наставник предаје и укупан број задатака спаривања који испитују ниво познавања чињеница на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,33	0,616	0,097
математика	40	0,03	0,158	0,025
$t = 3,00; df = 44,1; p < 0.01$				

Табела 31. Предмет који наставник предаје и укупан број задатака спаривања који испитују ниво примене знања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,10	0,304	0,048
математика	40	0,00	0,000	0,000
$t = 2,08; df = 39,0; p = 0.04$				

Табела 32. Предмет који наставник предаје и укупан број задатака вишеструког избора који испитују ниво разумевања градива на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,36	0,590	0,093
математика	40	0,03	0,158	0,025
$t = 3,44; df = 44,6; p = 0.00$				

Табела 33. Предмет који наставник предаје и укупан број задатака вишеструког избора који испитују ниво примене знања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,20	0,516	0,082
математика	40	0,00	0,000	0,000
$t = 2,45; df = 39,0; p = 0.02$				

Табела 34. Предмет који наставник предаје и укупан број задатака „других“ форми који испитују ниво познавања чињеница на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,43	0,712	0,113
математика	40	0,00	0,000	0,000
$t = 3,78; df = 39,0; p = 0.00$				

Табела 35. Предмет који наставник предаје и укупан број задатака „других“ форми који испитују ниво примене знања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,23	0,424	0,67
математика	40	0,00	0,000	0,000
$t = 3,46; df = 39,0; p = 0.00$				

Табела 36. Предмет који наставник предаје и укупан број лаких задатака који имају форму задатака допуњавања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,39	0,625	0,099
математика	40	0,10	0,379	0,060
$t = 2,49; df = 64,2; p = 0.02$				

Табела 37. Предмет који наставник предаје и укупан број лаких задатака који имају форму задатака са кратким одговорима на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,93	1,158	0,183
математика	40	0,31	0,911	0,144
$t = 2,63; df = 73,9; p = 0.01$				

Табела 38. Предмет који наставник предаје и укупан број лаких задатака који имају форму задатака спаривања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,33	0,656	0,104
математика	40	0,03	0,158	0,025
$t = 2,81; df = 43,5; p < 0.01$				

Табела 39. Предмет који наставник предаје и укупан број лаких задатака који имају форму задатака „друге“ форме на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,38	0,740	0,117
математика	40	0,03	0,158	0,025
$t = 2,92; df = 42,6; p < 0.01$				

Табела 40. Предмет који наставник предаје и укупан број средње тешких задатака који имају форму задатака вишеструког избора на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,44	0,761	0,120
математика	40	0,04	0,175	0,028
$t = 3,24; df = 43,1; p = 0.00$				

Табела 41. Предмет који наставник предаје и укупан број средње тешких задатака који имају форму задатака допуњавања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,27	0,555	0,088
математика	40	0,08	0,267	0,042
$t = 2,00; df = 56,1; p < 0.05$				

Табела 42. Предмет који наставник предаје и укупан број средње тешких задатака који имају форму задатака са кратким одговорима на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,90	1,201	0,190
математика	40	0,00	0,000	0,000
$t = 4,71; df = 39,0; p = 0.00$				

Табела 43. Предмет који наставник предаје и укупан број средње тешких задатака који имају форму задатака „друге“ форме на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,26	0,492	0,078
математика	40	0,03	0,158	0,025
$t = 2,85; df = 47,0; p < 0.01$				

Табела 44. Предмет који наставник предаје и укупан број тешких задатака који имају форму задатака вишеструког избора на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,21	0,590	0,093
математика	40	0,00	0,000	0,000
$t = 2,23; df = 39,0; p = 0.03$				

Табела 45. Предмет који наставник предаје и укупан број тешких задатака који имају форму задатака са кратким одговорима на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,73	1,132	0,179
математика	40	0,03	0,158	0,025
$t = 3,87; df = 40,5; p = 0.00$				

Табела 46. Предмет који наставник предаје и укупан број задатака испуњавања налога на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	3,47	2,627	0,415
математика	40	5,75	2,582	0,408

$t = - 3,92; df = 78; p = 0.00$

Табела 47. Предмет који наставник предаје и укупан број задатака испуњавања налога који испитују ниво познавања чињеница на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	1,11	1,238	0,196
математика	40	1,79	1,236	0,195

$t = - 2,46; df = 78; p = 0.02$

Табела 48. Предмет који наставник предаје и укупан број задатака испуњавања налога који испитују ниво разумевања градива на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	1,38	1,699	0,269
математика	40	2,21	1,116	0,177

$t = - 2,58; df = 67,4; p = 0.01$

Табела 49. Предмет који наставник предаје и укупан број задатака испуњавања налога који испитују ниво примене знања на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	1,13	1,101	0,174
математика	40	1,77	1,337	0,211

$t = - 2,31; df = 75,2; p = 0.02$

Табела 50. Предмет који наставник предаје и укупан број лаких задатака који имају форму испуњавања налога на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	0,88	1,249	0,198
математика	40	1,55	1,044	0,165

$t = - 2,61; df = 78; p = 0.01$

Табела 51. Предмет који наставник предаје и укупан број средње тешких задатака који имају форму задатака испуњавања налога на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	1,82	1,782	0,282
математика	40	2,61	1,704	0,269

$t = - 2,03; df = 78; p < 0.05$

Табела 52. Предмет који наставник предаје и укупан број тешких задатака који имају форму задатака испуњавања налога на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	1,00	0,926	0,146
математика	40	1,62	1,409	0,223

$t = - 2,30; df = 67,4; p = 0.03$

Табела 53. Предмет који наставник предаје и просечан број поена по задатаку на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	4,44	3,684	0,583
математика	40	10,52	9,621	1,521

$t = - 3,73; df = 78; p = 0.00$

Табела 54. Предмет који наставник предаје и укупан број поена који се може остварити на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	34,08	20,427	3,230
математика	40	52,10	41,929	6,630

$t = - 2,44; df = 78; p = 0.02$

Табела 55. Предмет који наставник предаје и доњи праг за пролазну оцену на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	32,25	12,349	1,953
математика	40	26,08	11,483	1,816

$t = 2,92; df = 65,4; p < 0.01$

Табела 56. Предмет који наставник предаје и број задатака који се могу признати као делимично тачни на нивоу теста

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	6,75	3,536	0,559
математика	40	5,08	2,596	0,410

$t = 2,42; df = 78; p = 0.02$

Табела 57. Предмет који наставник предаје и укупан број паралелних форми теста које наставник израђује

Предмет који наставник предаје	N	M	SD	SEM
српски језик	40	2,35	0,802	0,127
математика	40	2,90	1,446	0,229

$t = - 2,10; df = 78; p = 0.04$

Табела 58. Предмет који наставник предаје и саопштавање упутства за решавање теста

Предмет	Саопштавање упутства					
	не		да		укупно	
	f	%	f	%	f	%
српски језик	8	20,0	32	80,0	40	100,0
математика	21	52,5	19	47,5	40	100,0
укупно	29	36,3	51	63,8	80	100,0
$\chi^2 = 9,14; df = 1; p = 0.00; C = 0,32$						

Табела 59. Предмет који наставник предаје и садржај уџбеника као основа за разраду садржаја теста знања

Предмет	Садржај уџбеника као основа за разраду садржаја теста знања					
	не		да		укупно	
	f	%	f	%	f	%
српски језик	22	55,0	18	45,0	40	50,0
математика	13	32,5	27	67,5	40	50,0
укупно	35	43,8	45	56,3	80	100,0
$\chi^2 = 4,11; df = 1; p = 0.04; C = 0,22$						

Табела 60. Предмет који наставник предаје и провера тежине задатака као циљ испитивања квалитета теста

Предмет	Провера тежине задатака као циљ испитивања квалитета теста					
	не		да		укупно	
	f	%	f	%	f	%
српски језик	22	55,0	18	45,0	40	50,0
математика	13	32,5	27	67,5	40	50,0
укупно	35	43,8	45	56,3	80	100,0
$\chi^2 = 4,11; df = 1; p = 0.04; C = 0,22$						

Табела 61. Коришћење резултата националних тестирања на нивоу школе

Коришћење резултата	f	%
не	58	18,0
нисам сигуран/ сигурна	221	68,6
да	43	13,4
укупно	322	100,0

$\chi^2 = 181,61; df = 2; p = 0.00$

Табела 62. Коришћење резултата међународних тестирања на нивоу школе

Коришћење резултата	f	%
не	78	24,5
нисам сигуран/ сигурна	221	69,5
да	19	6,0
укупно	318	100,0

$\chi^2 = 203,57; df = 2; p = 0.00$

Табела 63. Предмет који наставник предаје и коришћење резултата националних тестирања на нивоу школе

Предмет	Коришћење резултата националних тестирања на нивоу школе							
	не		нисам сигуран/ сигурна		да		укупно	
	f	%	f	%	f	%	f	%
друштвено-језичка група	21	120	23	164	21	120	23	164
природно-математичка група	37	101	20	158	37	101	20	158
укупно	58	221	43	322	58	221	43	322

$\chi^2 = 6,15; df = 2; p < 0.05; C = 0,14$

Табела 64. Године радног стажа наставника у просвети и коришћење резултата Мале матуре за унапређивање ваннаставног рада са ученицима

Коришћење резултата Мале матуре за унапређивање ваннаставног рада са ученицима			
Године радног стажа	N	M	SD
0-10	149	3,60	0,979
10-20	137	3,38	1,099
20 и више	63	3,16	1,096
укупно	349	3,43	1,058
F = 4,16; df1 = 2; df2 = 346; p = 0.02			

Табела 65. Године радног стажа наставника у просвети и коришћење резултата националних тестирања за унапређивање начина рада у настави

Коришћење резултата националних тестирања за унапређивање начина рада у настави			
Године радног стажа	N	M	SD
0-10	141	3,76	0,827
10-20	128	4,01	0,828
20 и више	58	3,71	0,838
укупно	327	3,85	0,837
F = 4,03; df1 = 2; df2 = 324; p = 0.02			

Табела 66. Предмет који наставник предаје и предности тестова знања у односу на усмено испитивање

Критеријум поређења	χ^2	df	p	C
истоветност захтева	1,51	1	0.22	0,07
реализација поступка испитивања	0,78	1	0.38	0,05
оцењивање одговора	2,10	1	0.15	0,08
обухват градива испитивањем	0,44	1	0.51	0,04
обухват ученика испитивањем	2,50	1	0.11	0,08
поновљивост поступка испитивања	0,06	1	0.81	0,01
могућност ученика да „варају“	1,93	1	0.17	0,07

Табела 67. Предмет који наставник предаје и предности тестова знања у односу на писане вежбе

Критеријум поређења	χ^2	df	p	C
истоветност захтева	0,02	1	0.90	0,01
истоветност критеријума оцењивања	0,35	1	0.55	0,03
припрема поступка испитивања	2,09	1	0.15	0,08
реализација поступка испитивања	1,28	1	0.26	0,06
оцењивање одговора	0,87	1	0.35	0,05
обухват градива испитивањем	0,34	1	0.56	0,03
обухват ученика испитивањем	0,01	1	0.94	0,00
поновљивост поступка испитивања	0,31	1	0.58	0,03
могућност ученика да „варају“	0,16	1	0.69	0,02

Табела 68. Предмет који наставник предаје и предности тестова знања у односу на електронске тестове

Критеријум поређења	χ^2	df	p	C
истоветност захтева	3,85	1	0.05	0,10
припрема поступка испитивања	1,47	1	0.23	0,06
реализација поступка испитивања	0,30	1	0.59	0,03
оцењивање одговора	0,31	1	0.58	0,03
обухват градива испитивањем	2,21	1	0.14	0,08
обухват ученика испитивањем	0,09	1	0.77	0,02
поновљивост поступка испитивања	2,06	1	0.15	0,08
могућност ученика да „варају“	0,46	1	0.50	0,04

Табела 69. Предмет који наставник предаје и процена припремљености за израду тестова знања

Процена припремљености за израду теста знања	t	df	p
укупна процена припремљености за израду тестова знања	0,93	345,1	0.35
процена припремљености са становишта потребних знања	1,16	357	0.25
процена припремљености са становишта потребних вештина	- 0,14	355	0.89

Табела 70. Предмет који наставник предаје и припрема за израду тестова знања током иницијалног професионалног образовања

Припрема током иницијалног професионалног образовања	χ^2	df	p	C
похађање припреме за израду тестова током студија	2,43	1	0.62	0,03
статус предмета у оквиру кога је разматрана тематика посвећена тестовима знања	2,32	1	0.13	0,13
разматрање тематике о појму и врстама тестова знања током студија	0,41	1	0.84	0,01
разматрање тематике о могућностима примене тестова знања у настави током студија	0,48	1	0.49	0,04
разматрање тематике о процесу израде теста током студија	0,02	1	0.90	0,01
разматрање тематике о врстама и карактеристикама задатака у тесту током студија	0,42	1	0.52	0,03
допринос припреме на студијама усвајању теоријских знања из области тестирања	0,73	1	0.39	0,05
допринос припреме на студијама развијању вештина потребних за израду тестова знања	1,25	1	0.26	0,06

Табела 71. Предмет који наставник предаје и процена доприноса одређених облика стручног усавршавања припремљености за израду тестова

Допринос облика стручног усавршавања	t	df	p
допринос семинара оспособљености за израду тестова знања	0,12	190	0.90
допринос сарадње са колегама усавршавању у процесу израде тестова знања	-0,62	316,0	0.54
допринос консултовања стручне литературе усавршавању у процесу израде тестова знања	-1,43	300	0.15

Табела 72. Предмет који наставник предаје и стручно усавршавање у домену израде тестова знања

Стручно усавршавање у домену израде тестова знања	χ^2	df	p	C
консултовање са искуснијим колегама као облик стручног усавршавања у домену израде тестова знања	0,71	1	0.40	0,04
сарадња са стручним сарадником у школи као облик стручног усавршавања у домену израде тестова знања	0,62	1	0.43	0,04
похађање семинара стручног усавршавања као облик стручног усавршавања у домену израде тестова знања	0,15	1	0.70	0,02
консултовање релевантне методолошке литературе као облик стручног усавршавања у домену израде тестова знања	0,14	1	0.71	0,02
учешће на стручном скупу као облик стручног усавршавања у домену израде тестова знања	0,08	1	0.79	0,02
број семинара које је наставник похађао разматрање тематике о појму и врстама тестова знања на семинару	1,80	4	0.77	0,07
разматрање тематике о могућностима примене тестова знања у настави на семинару	0,94	1	0.33	0,05
разматрање тематике о процесу израде теста знања на семинару	1,95	1	0.16	0,07
разматрање тематике о процесу израде теста знања на семинару	2,44	1	0.12	0,08
разматрање тематике о врстама и карактеристикама задатака у тесту на семинару	0,12	1	0.73	0,02
допринос семинара усвајању теоријских знања из области тестирања	0,18	1	0.67	0,02
допринос семинара развијању вештина потребних за израду тестова знања	0,63	1	0.43	0,04
допринос семинара усклађивању теоријских и практичних знања	0,76	1	0.38	0,05
похађање семинара на личну иницијативу	0,20	1	0.65	0,02
похађање семинара на иницијативу директора школе	0,03	1	0.87	0,01
похађање семинара на иницијативу Завода/Министарства просвете	0,14	1	0.70	0,02
Министарство просвете као финансијер семинара	0,51	1	0.48	0,04
школа као финансијер семинара	2,44	1	0.12	0,08
наставници као финансијери семинара	0,85	1	0.36	0,05
нисам похађала/похађао такав семинар	0,14	1	0.71	0,02

Табела 73. Предмет који наставник предаје и тешкоће у изради тестова знања и начини њиховог превазилажења

Тешкоће у изради тестова знања и начини њиховог превазилажења	χ^2	df	p	C
тешкоће у разради садржаја теста	0,44	1	0.51	0,04
тешкоће у одређивању садржаја задатака	1,95	1	0.16	0,07
тешкоће у одређивању нивоа знања који задатак мери	0,02	1	0.89	0,01
тешкоће у одређивању тежине задатка	1,06	1	0.30	0,05
тешкоће у избору врсте задатка	1,01	1	0.32	0,05
тешкоће у одређивању дужине теста	0,02	1	0.89	0,01
тешкоће у одређивању редоследа задатака	2,79	1	0.10	0,09
тешкоће у одређивању броја поена који задатак носи	0,84	1	0.36	0,05
тешкоће у превођењу поена на тесту у оцену	1,18	1	0.28	0,06
тешкоће у развијању и чувању базе задатака	1,51	1	0.22	0,06
решавање тешкоћа уз консултације са искуснијим колегама	1,72	1	0.19	0,07
решавање тешкоћа у сарадњи са педагогом школе	0,01	1	0.92	0,01
решавање тешкоћа у сарадњи са психологом школе	0,69	1	0.41	0,44
решавање тешкоћа путем „покушаја и погрешака“	0,94	1	0.33	0,05
решавање тешкоћа уз помоћ методолошке литературе	2,01	1	0.16	0,07
немам тешкоћа у изради тестова знања	1,78	1	0.18	0,07

Табела 74. Предмет који наставник предаје и самостална израда тестова знања

Самостална израда тестова	χ^2	df	p	C
заступљеност самосталне израде тестова знања	0,61	1	0.44	0,04
учесталост самосталне израде тестова знања током школске године	6,90	4	0.14	0,14

Табела 75. Предмет који наставник предаје и циљ израде теста знања

Циљ израде теста	χ^2	df	p	C
утврђивање претходног знања код ученика као циљ израде теста	1,54	1	0.21	0,07
оцењивање ученика као циљ израде теста	0,34	1	0.56	0,03
планирање садржаја и начина рада у настави као циљ израде теста	0,41	1	0.53	0,03
организовање рада у допунској настави као циљ израде теста	2,98	1	0.08	0,09
организовање рада у додатној настави као циљ израде теста	2,10	1	0.16	0,07
размена информација о напредовању ученика као циљ израде теста	0,81	1	0.37	0,05
планирање стручног усавршавања као циљ израде теста	0,16	1	0.69	0,02
унапређивање процеса израде теста као циљ израде теста	0,03	1	0.87	0,01

Табела 76. Предмет који наставник предаје и коришћење резултата Мале матуре на нивоу школске установе

Коришћења резултата Мале матуре	χ^2	df	p	C
заступљеност коришћења резултата Мале матуре на нивоу школе	2,03	2	0.36	0,08
начин коришћења резултате Мале матуре на нивоу школе	2,20	1	0.14	0,14

Табела 77. Предмет који наставник предаје и могућности коришћења резултата добијених применом стандардизованих тестова знања

Могућности коришћења резултата добијених применом стандардизованих тестова знања	t	df	p
коришћење резултата Мале матуре за унапређивање садржаја наставе	0,97	348	0.34
коришћење резултата Мале матуре за унапређивање начина рада у настави	0,57	343,7	0.57
коришћење резултата Мале матуре за унапређивање процеса оцењивања	-0,26	346	0.80
коришћење резултата Мале матуре за унапређивање ваннаставног рада са ученицима	0,59	347	0.56
коришћење резултата Мале матуре за планирање стручног усавршавања наставника	-0,70	340,0	0.49
коришћење резултата Мале матуре за побољшање међуљудских односа у школи	0,62	348	0.55
коришћење резултата Мале матуре за превазилажење разлика између деце различитог пола и породичног порекла	0,56	336,7	0.58
коришћење резултата националних тестирања за унапређивање садржаја наставе	1,34	324,5	0.18
коришћење резултата националних тестирања за унапређивање начина рада у настави	0,57	324,9	0.57
коришћење резултата националних тестирања за унапређивање процеса оцењивања	0,12	325	0.90
коришћење резултата националних тестирања за унапређивање ваннаставног рада са ученицима	1,40	322,7	0.16
коришћење резултата националних тестирања за планирање стручног усавршавања наставника	0,22	325	0.83
коришћење резултата националних тестирања за побољшање међуљудских односа у школи	-0,44	298,7	0.66
коришћење резултата националних тестирања за побољшање опремљености школе	-1,17	321,2	0.25
коришћење резултата националних тестирања за превазилажење разлика између деце различитог пола и породичног порекла	-1,64	325	0.10
коришћење резултата међународних тестирања за унапређивање садржаја наставе	0,65	318	0.52
коришћење резултата међународних тестирања за унапређивање начина рада у настави	0,28	318	0.78
коришћење резултата међународних тестирања за унапређивање процеса оцењивања	0,56	318	0.58
коришћење резултата међународних тестирања за унапређивање ваннаставног рада са ученицима	1,43	318	0.15
коришћење резултата међународних тестирања за планирање стручног усавршавања наставника	1,15	305,5	0.25
коришћење резултата међународних тестирања за побољшање међуљудских односа у школи	0,58	293,7	0.57
коришћење резултата међународних тестирања за побољшање опремљености школе	-0,14	307,5	0.89
коришћење резултата међународних тестирања за превазилажење разлика између деце различитог пола и породичног порекла	0,00	296,0	1.00

Табела 78. Године радног стажа наставника и предности тестова знања у односу на усмено испитивање

Критеријум поређења	χ^2	df	p	C
	истоветност захтева	5,17	2	0.08
истоветност критеријума оцењивања	2,20	2	0.33	0,08
припрема поступка испитивања	4,56	2	0.10	0,11
реализација поступка испитивања	2,23	2	0.33	0,08
оцењивање одговора	1,30	2	0.52	0,06
обухват градива испитивањем	1,58	2	0.46	0,07
обухват ученика испитивањем	5,39	2	0.07	0,12
поновљивост поступка испитивања	1,52	2	0.47	0,07
могућност ученика да „варају“	1,72	2	0.42	0,07

Табела 79. Године радног стажа наставника и предности тестова знања у односу на писане вежбе

Критеријум поређења	χ^2	df	p	C
	истоветност захтева	0,41	2	0.82
истоветност критеријума оцењивања	3,99	2	0.14	0,10
припрема поступка испитивања	0,68	2	0.71	0,04
реализација поступка испитивања	0,70	2	0.71	0,04
оцењивање одговора	1,28	2	0.53	0,06
обухват градива испитивањем	0,06	2	0.97	0,01
обухват ученика испитивањем	0,72	2	0.70	0,04
поновљивост поступка испитивања	0,38	2	0.83	0,03
могућност ученика да „варају“	0,23	2	0.89	0,03

Табела 80. Године радног стажа наставника и предности класичних тестова у односу на електронске тестове

Критеријум поређења	χ^2	df	p	C
припрема поступка испитивања	3,84	2	0.15	0,10
реализација поступка испитивања	0,48	2	0.79	0,04
оцењивање одговора	0,68	2	0.71	0,04
обухват градива испитивањем	4,17	2	0.12	0,11
обухват ученика испитивањем	1,44	2	0.49	0,06
поновљивост поступка испитивања	0,12	2	0.94	0,02

Табела 81. Године радног стажа наставника и процена припремљености за израду тестова знања

Процена припремљености	F	df1	df2	p
укупна процена припремљености за израду тестова знања	0,10	357	359	0.90
процена припремљености са становишта потребних знања	0,37	356	358	0.69
процена припремљености са становишта потребних вештина	0,12	354	356	0.89

Табела 82. Године радног стажа наставника и припрема за израду тестова знања током иницијалног професионалног образовања

Припрема током иницијалног професионалног образовања	χ^2	df	p	C
похађање припреме за израду тестова током студија	3,20	2	0.20	0,10
статус предмета у оквиру кога је разматрана тематика посвећена тестовима знања	7,00	4	0.14	0,22
разматрање тематике о могућностима примене тестова знања у настави током студија	2,48	2	0.29	0,08
разматрање тематике о процесу израде теста током студија	3,23	2	0.20	0,09
допринос припреме на студијама усвајању теоријских знања из области тестирања	4,00	2	0.14	0,11

Табела 83. Године радног стажа наставника и и стручно усавршавање у домену израде тестова знања

Стручно усавршавање у домену израде тестова знања	χ^2	df	p	C
консултовање са искуснијим колегама као облик стручног усавршавања у домену израде тестова знања	3,36	2	0.19	0,10
сарадња са стручним сарадником у школи као облик стручног усавршавања у домену израде тестова знања	3,77	2	0.15	0,10
учешће на стручном скупу као облик стручног усавршавања у домену израде тестова знања	0,12	2	0.94	0,02
разматрање тематике о могућностима примене тестова знања у настави на семинару	2,89	2	0.24	0,09
разматрање тематике о врстама и карактеристикама задатака у тесту на семинару	4,31	2	0.12	0,11
допринос семинара усвајању теоријских знања из области тестирања	5,00	2	0.08	0,12
допринос семинара усклађивању теоријских и практичних знања	2,49	2	0.29	0,08
похађање семинара на личну иницијативу	1,84	2	0.40	0,07
Министарство просвете као финансијер семинара	3,70	2	0.16	0,10
наставници као финансијери семинара	0,14	2	0.93	0,02

Табела 84. Године радног стажа наставника и процена доприноса одређених облика стручног усавршавања припремљености за израду тестова

Доприноса облика стручног усавршавања	F	df1	df2	p
допринос семинара оспособљености за израду тестова знања	0,50	189	191	0.61
допринос сарадње са колегама усавршавању у процесу израде тестова знања	2,02	315	317	0.14
допринос консултовања стручне литературе усавршавању у процесу израде тестова знања	0,192	299	301	0.83

Табела 85. Године радног стажа наставника и тешкоће у изради тестова знања и начини њиховог превазилажења

Тешкоће у изради тестова знања и начини њиховог превазилажења	χ^2	df	p	C
тешкоће у разради садржаја теста	0,77	2	0.68	0,05
тешкоће у одређивању садржаја задатака	5,45	2	0.07	0,12
тешкоће у одређивању нивоа знања који задатак мери	0,70	2	0.70	0,04
тешкоће у одређивању тежине задатка	0,84	2	0.66	0,05
тешкоће у избору врсте задатка	0,20	2	0.91	0,02
тешкоће у одређивању редоследа задатака	0,70	2	0.71	0,04
тешкоће у одређивању броја поена који задатак носи	3,17	2	0.21	0,09
тешкоће у превођењу поена на тесту у оцену	0,71	2	0.70	0,04
тешкоће у развијању и чувању базе задатака	2,70	2	0.26	0,09
решавање тешкоћа у сарадњи са педагогом школе	0,87	2	0.65	0,05
решавање тешкоћа у сарадњи са психологом школе	0,58	2	0.75	0,04
решавање тешкоћа путем „покушаја и погрешака“	0,80	2	0.67	0,05
решавање тешкоћа уз помоћ методолошке литературе	1,86	2	0.40	0,07
немам тешкоћа у изради тестова знања	2,37	2	0.31	0,08

Табела 86. Године радног стажа наставника и учесталост коришћења тестова знања током школске године и самосталност у изради тестова

Учесталост коришћења и самосталност у изради тестова	χ^2	df	p	C
учесталост коришћења тестова током школске године	12,07	8	0.15	0,18
самосталност у изради тестова знања	0,48	2	0.79	0,04

Табела 87. Године радног стажа наставника и и циљ израде теста знања

Циљ израде теста	χ^2	df	p	C
утврђивање претходног знања код ученика као циљ израде теста	4,46	2	0.11	0,11
планирање и предузимање мера подршке ученицима у учењу	0,01	2	1.00	0,01
оцењивање ученика као циљ израде теста	0,13	2	0.94	0,02
планирање садржаја и начина рада у настави као циљ израде теста	3,09	2	0.21	0,09
организовање рада у допунској настави као циљ израде теста	0,61	2	0.74	0,04
организовање рада у додатној настави као циљ израде теста	1,98	2	0.37	0,07
размена информација о напредовању ученика као циљ израде теста	1,77	2	0.41	0,07
планирање стручног усавршавања као циљ израде теста	5,57	2	0.06	0,12
унапређивање процеса израде теста као циљ израде теста	3,32	2	0.19	0,10

Табела 88. Године радног стажа наставника и коришћење резултата Мале матуре и националних тестирања на нивоу школске установе

Коришћења резултата Мале матуре и националних тестирања	χ^2	df	p	C
коришћење резултата Мале матуре на нивоу школе	3,87	4	0.42	0,10
начин коришћења резултате Мале матуре на нивоу школе	0,06	2	0.97	0,02
коришћење резултата националних тестирања на нивоу школе	4,47	4	0.35	0,12

Табела 89. Године радног стажа наставника и могућности коришћења резултата добијених применом стандардизованих тестова знања

Могућности коришћења резултата добијених применом стандардизованих тестова знања	F	df1	df2	p
коришћење резултата Мале матуре за унапређивање садржаја наставе	1,01	347	349	0.37
коришћење резултата Мале матуре за унапређивање начина рада у настави	0,70	347	349	0.50
коришћење резултата Мале матуре за унапређивање процеса оцењивања	0,05	345	347	0.95
коришћење резултата Мале матуре за планирање стручног усавршавања наставника	1,28	347	349	0.28
коришћење резултата Мале матуре за побољшање међуљудских односа у школи	0,48	347	349	0.62
коришћење резултата Мале матуре за побољшање опремљености школе	0,45	342	344	0.64
коришћење резултата Мале матуре за превазилажење разлика између деце различитог пола и породичног порекла	1,04	347	349	0.35
коришћење резултата националних тестирања за унапређивање садржаја наставе	0,54	324	326	0.59
коришћење резултата националних тестирања за унапређивање процеса оцењивања	2,73	324	326	0.07
коришћење резултата националних тестирања за унапређивање ваннаставног рада са ученицима	0,95	322	324	0.39
коришћење резултата националних тестирања за планирање стручног усавршавања наставника	1,49	324	326	0.23
коришћење резултата националних тестирања за побољшање међуљудских односа у школи	0,85	324	326	0.43
коришћење резултата националних тестирања за побољшање опремљености школе	0,32	324	326	0.73
коришћење резултата националних тестирања за превазилажење разлика између деце различитог пола и породичног порекла	0,43	324	326	0.65
коришћење резултата међународних тестирања за унапређивање садржаја наставе	0,95	317	319	0.39
коришћење резултата међународних тестирања за унапређивање начина рада у настави	0,70	317	319	0.50
коришћење резултата међународних тестирања за унапређивање процеса оцењивања	1,36	317	319	0.26
коришћење резултата међународних тестирања за унапређивање ваннаставног рада са ученицима	0,73	317	319	0.48
коришћење резултата међународних тестирања за планирање стручног усавршавања наставника	1,15	317	319	0.32
коришћење резултата међународних тестирања за побољшање међуљудских односа у школи	0,09	317	319	0.91
коришћење резултата међународних тестирања за побољшање опремљености школе	0,18	317	319	0.84
коришћење резултата међународних тестирања за превазилажење разлика између деце различитог пола и породичног порекла	0,32	317	319	0.73

Табела 90. Предмет који наставник предаје и фазе у конструкцији теста знања

Фазе у конструкцији теста	χ^2	df	p	C
одређивање садржаја теста	0,25	1	0.62	0,06
одређивање садржаја задатака	0,08	1	0.78	0,03
одређивање нивоа знања који се задатком испитује	1,15	1	0.28	0,12
одређивање тежине задатка	0,00	1	1.00	0,00
избор форме задатка	0,95	1	0.33	0,11
утврђивање начина оцењивања теста	0,05	1	0.82	0,03
коришћење задатака других аутора	2,14	1	0.14	0,16

Табела 91. Предмет који наставник предаје и циљ израде теста знања

Циљ израде	χ^2	df	p	C
провера савладаности градива	0,25	1	0.62	0,06
оцењивање ученика	0,35	1	0.56	0,67

Табела 92. Предмет који наставник предаје и основа за разраду садржаја теста

Основа за разраду садржаја теста	χ^2	df	p	C
садржај наставног програма	0,39	1	0.53	0,07
стандарди знања	0,25	1	0.62	0,06

Табела 93. Предмет који наставник предаје и реализација тестирања

Аспекти реализације	χ^2	df	p	C
средства за писање	1,01	3	0.80	0,11
трајање тестирања	0,75	2	0.69	0,10
пружање помоћи ученицима током решавања теста	1,13	2	0.57	0,12

Табела 94. Предмет који наставник предаје и провера квалитета теста

Провера квалитета теста	χ^2	df	p	C
заступљеност провере квалитета теста	1,35	1	0.25	0,13
провера квалитета теста уз консултације са другим колегама	1,92	1	0.17	0,15
провера формулације задатака као циљ провере	0,00	1	1.00	0,00

Табела 95. Предмет који наставник предаје и формирање и чување база задатака

Формирање и чување базе задатака	χ^2	df	p	C
формирање електронске базе задатака	0,29	1	0.59	0,06
чување штампаних верзија теста	0,95	1	0.33	0,11
чување електронских верзија теста на персоналном рачунару као начин осигуравања базе	0,62	1	0.43	0,09
чување штампаних верзија теста ван школе као начин осигуравања базе	1,07	1	0.30	0,12
чување штампаних верзија теста у школи, у простору који се закључава, као начин осигуравања базе	0,80	1	0.37	0,10

Табела 96. Предмет који наставник предаје и коришћење тестова других аутора

Коришћење тестова других аутора	χ^2	df	p	C
заступљеност коришћења тестова других аутора	0,27	2	0.88	0,06
тестови других колега	0,31	1	0.58	0,06
тестови које припремају тимови стручњака које окупља издавачких кућа	0,11	1	0.75	0,04
тестови са интернета	1,87	1	0.17	0,15

Табела 97. Предмет који наставник предаје и циљ анализе ученичких постигнућа на тесту

Планирање наставног рада као циљ анализе ученичких постигнућа	χ^2	df	p	C
	0,20	1	0.65	0,05

Табела 98. Предмет који наставник предаје и доњи праг поена за одговарајућу оцену

Доњи праг поена	t	df	p
доњи праг поена за добру (3) оцену	1,78	69,9	0.08
доњи праг поена за врло добру (4) оцену	0,52	70	0.61
доњи праг поена за одличну (5) оцену	-0,64	66,0	0.53

Табела 99. Предмет који наставник предаје и минималан број паралелних форми теста

Минималан број паралелних форми теста	t	df	p
	-0,86	78	0.40

Табела 100. Предмет који наставник предаје и карактеристике задатака у тесту

Карактеристике задатака у тесту	t	df	p
укупан број нивоа знања који се испитује задацима у тесту	-1,29	62,1	0.20
укупан број задатака који испитују ниво разумевања градива	1,99	54,5	0.05
укупан број задатака који испитују ниво примене знања	1,93	76,7	0.06
укупан број нивоа тежине задатака који се тестом испитује	0,00	78,0	1.00
укупан број тешких задатака	1,87	76,3	0.07
укупан број задатака који се не могу признати као делимично тачни на нивоу теста	1,56	71,3	0.12
укупан број задатака двочланог избора	0,25	78	0.81
укупан број задатака сређивања	0,00	78,0	1.00
укупан број задатака двочланог избора који испитују ниво познавања чињеница	1,00	39,0	0.32
укупан број задатака вишечланог избора који испитују ниво познавања чињеница	1,17	64,1	0.25
укупан број задатака сређивања који испитују ниво познавања чињеница	1,00	39,0	0.32
укупан број задатака двочланог избора који испитују ниво разумевања градива	-0,58	70,7	0.56
укупан број задатака допуњавања који испитују ниво разумевања градива	1,52	54,1	0.13
укупан број задатака спаривања који испитују ниво разумевања градива	1,02	63,4	0.31
укупан број задатака сређивања који испитују ниво разумевања градива	-1,00	39,0	0.32
укупан број задатака „друге“ форме који испитују ниво разумевања градива	1,78	62,2	0.08
укупан број лаких задатака са формом задатака двочланог избора	0,00	78	1.00
укупан број лаких задатака са формом задатака вишечланог избора	0,96	75,5	0.34
укупан број средње тешких задатака са формом задатака двочланог избора	1,00	39,0	0.32
укупан број средње тешких задатака са формом задатака спаривања	1,39	58,7	0.17
укупан број средње тешких задатака са формом задатака сређивања	0,00	78,0	1.00
укупан број тешких задатака са формом задатака допуњавања	1,86	46,7	0.07
укупан број тешких задатака са формом задатака спаривања	1,78	39,0	0.08

БИОГРАФИЈА

Ивана Луковић је рођена 1979. године у Београду. Основне студије педагогије на Филозофском факултету у Београду завршила је 2003. године. Професионално искуство стицала је радећи као школски педагог, истраживач-приправник на Институту за педагошка истраживања, потом као сарадник у настави и асистент на Одељењу за педагогију и андрагогију, Филозофског факултета у Београду. Тренутно је ангажована као асистент у настави предмета Увод у педагошка истраживања, Увод у андрагошка истраживања и Методе и технике истраживања у образовању на основним студијама педагогије и андрагогије Филозофског факултета у Београду. Од 2004. године ангажована је на научно-истраживачким пројектима Института за педагошка истраживања и, касније, Института за педагогију и андрагогију. Објавила је више радова у научним часописима, научним монографијама и зборницима радова са научних скупова. Од 2005. године аутор је и реализатор семинара за стручно усавршавање запослених у образовању.

Изјава о ауторству

Име и презиме аутора _____ Ивана Луковић _____
Број индекса _____ 2P130006 _____

Изјављујем

да је докторска дисертација под насловом

«Тестови знања у школској пракси»

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, 11.07.2016.

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора _____ Ивана Луковић _____
Број индекса _____ 2P130006 _____
Студијски програм _____ педагогија _____
Наслов рада _____ «Тестови знања у школској пракси» _____
Ментор _____ проф. др Наташа Матовић _____

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањена у **Дигиталном репозиторијуму Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада. Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

У Београду, 11.07.2016.

Потпис аутора

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

«Тестови знања у школској пракси»

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)
2. Ауторство – некомерцијално (CC BY-NC)
- 3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)**
4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)
5. Ауторство – без прерада (CC BY-ND)
6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, 11.07.2016.
