Софийски университет „Св. Климент Охридски”
Факултет по математика и информатика
Катедра „Обучение по математика и информатика”

Доц. д-р Кирил Банков

ШИРОКО МАЩАБНИ ОЦЕНЪЧНО-ДИАГНОСТИЧНИ ПЕДАГОГИЧЕСКИ ИЗСЛЕДВАНИЯ

Хабилитационен труд
за академичната длъжност *професор*

София, 2012

СЪДЪРЖАНИЕ

ИЗПОЛЗВАНИ СЪКРАЩЕНИЯ

БЕЛ – Български език и литература
ВП – Високи постижения
ВУ – Висше училище
ДОИ – Държавни образователни изисквания
ЗСООМУП – Закон за степента на образование, образователния минимум и учебен план
ЗНП – Закон за народната просвета
МВП – Много високи постижения
МНП – Много ниски постижения
МО – Министерство на образованието (използва се като общо название на държавно ведомство, отговорно за провеждане на образователната политика в определена образователна система)
МОМН – Министерство на образованието, младежта и науката (в България)
НИО – Научноизследователски институт по образование
НП – Ниски постижения
ОДИ – Оценъчно-диагностично изследване (използва се и в мн. ч.)
СО – Система за оценявана
СУ – Софийски университет
УС – Учебно съдържание
ЦКОКУО – Център за контрол и оценка на качеството на училищното образование

ACER – Australian Center for Educational research
CITO – Инициалите на една от водещите в света компании за тестово оценяване в образованието
EAP – Expected A-Posterior Estimate
ETS – Educational Testing Service
HLM – Hierarchical Linear Modeling
IEA – International Association for Evaluation of Educational Achievement
IRT – Item Response Theory
ISCED – International Standard Classification of Education
MLE – Maximum Likelihood Estimation
MSU – Michigan State University
MT21 – Mathematics Teaching in the 21st Century
NAEP – National Assessment of Educational Progress
NSF – National Science Foundation
OECD – Organization for Economic Co-operation and Development
OLS – Ordinary Least Square
PIRLS – Progress in International Reading Literacy Study
PISA – Progress in International Student Achievement
PPS – Probabilities Proportional to their Size
SES – Social Economic Status
TALIS – Teaching and Learning International Survey
TEDS-M – Teacher Education and Development Study – Mathematics
TIMSS – Trends in International Mathematics and Science Study
UNESCO – United Nations Educational, Scientific and Cultural Organization
WLE – Weighted Maximum Likelihood Estimation

УВОД

Изследванията, свързани с оценяване в образованието, са своеобразен връх в развитието на педагогиката. От теоретична гледна точка те обединяват знанията от различни области, като психология, социология, философия, статистика. От практическа гледна точка те са най-добрият начин за мониторинг на образованието, тъй като информират и подпомагат управленските органи, стимулират подготовката и повишаването на квалификацията на учителите, обогатяват и осъвременяват учебните планове, методите на преподаване и цялостната учебно-възпитателна дейност.

В редица страни проектираните и функциониращите национални програми за оценяване в образованието синтезират най-новите постижения в областта на измерването, диагностиката, тестологията и психометрията. Създадените тестове и други измерителни методи и средства позволяват пълно, обективно и надеждно да се оценят действителните резултати от учебно-възпитателната работа, напредъка в развитието, съществуващите трудности и грешки.

Оценката означава, най-общо казано, "стойността" на даден продукт, идея, дейност или резултат по отношение на определена цел или област от действителността и познанието. Тази оценка може да бъде субективна, изразена като мнение на отделно лице, но тя може да стане по-обективна чрез прилагането на система от научни методи и подходи. Продукти, които могат в по-голяма степен да се измерят в количествено отношение, се оценяват значително по-обективно в сравнение с други, при които прилагането на количествени измервания е ограничено. Образованието и неговите "продукти" се отнасят към тези области, за които няма нито еднозначни измерители, нито достатъчно обективни критерии, въз основа на които да се направи напълно обективно оценяване. Поради това през последните десетилетия възникна и се разви ново научноизследователско и практическо направление, наречено "оценяване на образованието" (Educational Assessment).

Оценяването в образованието се прави въз основа на резултатите от специално подготвени и проведени изследвания, наречени **оценъчно-диагностични изследвания** (ОДИ). В този труд често се използва думата **оценяване** вместо „оценъчно-диагностично изследване". Тези изследвания са интердисциплинарни. С тях се изучават, за да бъдат оценени и диагностицирани, опреде-

лени процеси, продукти, програми, експерименти, както и свързаните с тях ефекти. Методите, с които се провеждат оценъчните изследвания, отговарят на диагностичните изисквания за обективност, надеждност и валидност.

Оценяването на образованието чрез оценъчно-диагностични изследвания може да обхваща както крупни проекти, експерименти и иновации в областта, така и вече проведени изследвания или проекти за учебни планове и програми. Могат да се оценяват още и последиците от планирани и проведени изменения в образователната система. Получените резултатите се използват многопосочно за вземане на важни управленски решения в образованието, за подобряване на учебно-възпитателната работа в училище, за изменения в учебната документация и управлението на образованието, квалификацията и подготовката на учители и ръководни кадри и др. Анализът на специализираната литература показва, че в световен мащаб на оценяването и на оценъчно-диагностичните изследвания се отделя заслужено голямо внимание, натрупан е значителен опит, разработени са теоретични подходи и експериментални модели.

Голяма част от оценъчно-диагностичните изследвания са **широко мащабни** (Large Scale Assessment). Според интернет страницата на Mathematical Association of America (http://www.maa.org/SAUM/faq.html) това са оценявания на голям брой ученици върху цяла академична програма. Следователно при широко мащабните изследвания популацията е многобройна и учебното съдържание, което се оценява, е с голям обем. Представителни примери за такива изследвания са *национални или международни оценъчно-диагностични изследвания*. При тях обикновено популацията се състои от всички ученици от определена възраст в дадена държава (или във всички участващи държави) и оценяването се извършва по учебното съдържание на даден предмет (или предметна област) за цял етап от обучението.

Една от основните цели на широко мащабните ОДИ е така наречената *отчетност* (accountability), което означава, че анализът на резултатите помага на обществеността и най-вече на образователните политици да видят до колко добре „работи" образователната система; сама за себе си (национални оценявания) или съпоставена с други такива (международни оценявания). Правилното използване на тези анализи е силно оръжие в ръцете на управленските образователни органи за по-адекватно провеждане на целенасочена образователна политика.

Историята на широко мащабните ОДИ е сравнително кратка. Тя започва от средата на 80-те години на XX век (Popham W.J., 2001). По-ранното провеждане на такива изследвания не е било възможно по различни причини, основните от които са следните. Както вече подчертахме, широко мащабните ОДИ работят с многобройни популации и голям обем от учебен материал. Обикновено се изследва извадка от популацията чрез извадки от тестовия материал. Това, от една страна, е свързано с използването на сложни методики, комбиниращи нестандартни математико-статистически методи и подходящи модели. От друга страна, нужна е съвременна компютърна техника, която да обслужва тези методики. Основните теоретични разработки (математически, статистически и др. модели) се появяват след средата на XX век. Те още са обект на развитие с оглед на нуждите на тяхното приложение. Например, специфичните условия на всяка държава понякога изискват модификация на известни методи, за да могат те да бъдат прилагани към съответната реалност. Въпреки сравнително кратката си история, световната практика в провеждане на широко мащабни ОДИ е доста богата.

През последните 15-20 години в страните от Югоизточна Европа се наблюдават сериозни промени в областта на оценяването на ученическите постижения. Много са причините за тези промени, като най-важните от тях са:

а) Вниманието на образователните системи все повече се фокусира върху резултатите от обучението.

б) Необходима е по-добра отчетност и образователна политика, обоснована от конкретни данни.

в) Появява се обществен натиск за подобряване на учебния процес, за което не малка роля имат международни изследвания като TIMSS, PIRLS, PISA и др.

г) Засилване на обществения интерес за „справедливо” разпределяне на местата във ВУ или на работните места.

д) Нуждата от ефективно използване на оскъдните ресурси за образование; и др.

Държавите от Югоизточна Европа реформираха или все още реформират системите си за оценяване на учениците. Някъде това става по-лесно; другаде – по-трудно, мудно и с много усилия. За съжаление, България е на едно от последните места в региона по тези реформи. Системата за национално оценяване още „прохожда”. Няколко пъти се е създавало специализирано национално

звено за оценяване в образованието, което е закривано и преоткривано поради липса на ясно изразена философия и политическа воля за реформа в областта на оценяването. От няколко години такова звено съществува (ЦКОКУО) и са налице конкретни негови дейности. След няколко неуспешни опита за въвеждане на матура, изглежда тя вече се утвърди като практика, макар че все още е под въпрос обективността на оценяването с нея. Нормативно е определено от МОМН да има национални оценявания при завършване на клас, образователен етап или степен, както и за използване на резултатите им при продължаване на образованието (за класиране). Епизодично са правени опити за извадкови национални ОДИ по отделни предмети (вж. напр. Разработване на стратегия... , 2004, Банков, К., 1997, 1998-Б, Георгиева, Н., Банков, К., 1997, Георгиева, Н., Банков, К., Тодорова, П., 1998), инициирани от научни работници, или неправителствени организации, но те не са изиграли ролята си на такива, главно поради липсата на подкрепа от страна на МОМН. България участва в някои международни изследвания (TIMSS, PIRLS, PISA и др.), по инициатива на ентусиасти – научни работници, въпреки слабия интерес от страна на МОМН.

У нас вече има натрупан опит в международен план при оценяване на определени аспекти на образователните системи. Налична е документацията от поне два национални проекта за разработване на методика и стратегия за външно оценяване в образованието. (Авторът на този труд е бил ръководител на два такива проекта: единият е съвместен проект на СУ „Св. Кл. Охридски" и Фондация „Отворено общество", а другият е към Европейския социален фонд и е ръководен от МОМН и ЦКОКУО.) Това е добра основа, за да се разработи и приеме официална научно обоснована стратегия за оценяване на средното образование в България. Изграждането й трябва задължително да се свърже със съществуващите традиции в образователната ни система и да се основава на всички досегашни опити (както положителни, така и несполучливи) в тази насока в България през последните 25-30 години, от които можем да се поучим.

В този труд вниманието е съсредоточено върху *методите и технологиите за подготовка, провеждане и анализ на данните от широко мащабни оценъчно-диагностични изследвания, както и резултати от някои такива.* **Основната цел** на труда е *да представи основните съвременни методи, технологии и модели за провеждане на такива изследвания, тяхната конкретизация и при-*

ложение в контекста на българската образователна система и резултати от някои ОДИ, в които авторът е участвал.

Въпреки, че теоретичните основи на методите и моделите за широко мащабни ОДИ са разработени през втората половина на XX век, те продължават да са обект на внимание както от страна на учените, така и от страна на тези, които ги прилаган на практика. Причината е, че общите методи и модели се нуждаят от доразвиване, усъвършенстване и конкретизация към различните условия, в които те се прилагат. Например, прилагането на общите теоретични постановки за получаване на извадка от ученици за ОДИ зависят от организацията на конкретна образователна система. Това налага съответните методи за получаване на извадка да бъдат доразвити и приложени. След това, обаче трябва да се доразвият и съответните методи/модели за обработка и анализ на данните, за да бъде моделираната ситуация най-близка до реалната. В този смисъл, **основните приноси на този труд** са *усъвършенстване, конкретизация и приложение на методите и моделите за широко мащабни ОДИ в условията на българската образователна система, както и анализ на някои резултати от проведени такива изследвания чрез използване на съответните адекватни методи.*

Трудът се състои от четири части, обединяващи петнадесет глави. Първата част съдържа две глави, в които се прави обобщен обзор на основните характеристики на националните и международните оценявания на ученическите постижения като представителен вид широко мащабни ОДИ. Първата глава на тази част подробно разглежда целите, задачите, начините на провеждане, обектите и други характерни особености на националните оценявания. В нея се прави и кратък преглед на опитите за национални оценявания в България от 1989 година насам. Приносът на автора е в систематизирането на основните научноизследователски въпроси, които се поставят в тези изследвания; описанието на последователността и отговорностите при тяхното провеждане; анализа на типичните грешки, които се допускат при национални ОДИ. Втората глава прави същото за международните оценявания. Описани са някои от големите такива оценявания, в които участва България или в които авторът има приноси, а именно TIMSS, PIRLS, PISA, MT21 и TEDS-M. Приносът на автора е в систематизирането на някои основни методически проблеми на международните ОДИ.

Във втората част се разглеждат методологични и технологични основи на националните и международни оценявания на ученическите постижения. Тя съдържа пет глави, които са съответно трета, четвърта, пета, шеста и седма глава от труда. В третата глава се разглеждат начините за получаване на извадка за широко мащабни ОДИ. Приносът на автора е в разработването на конкретна методика за получаване на извадка за България. В четвъртата глава се дискутират основни проблеми, свързани със скалирането на широко мащабни ОДИ. Накратко са представени основите на вероятностното моделиране (Item Response Theory – IRT). След това е направен анализ на съгласуваността на модела на TIMSS с данните за някои конкретни задачи за България от изследванията през 1995 и 1999 година – това е приносът на автора в тази глава. Петата глава е посветена на тестовите задачи. Приносът на автора е в систематизирането на известните, но разхвърляни познания за тяхната класификация и основните правила за писане. Авторът е извършил тази дейност във връзка с лекционните курсове със студенти, учители и други слушатели в образователни програми. В шеста глава се описва процеса на създаване на измерителен инструментариум (така наречения *тест*). Авторът представя разработки на съдържателна рамка, тестови спецификации, тестов дизайн и др., които той е писал или е бил активен участник в тяхното създаване. В седма глава са представени рамката, тестовите спецификации, дизайна и някои от задачите в изследването TIMSS. Авторът представя приносите си като член на експертния съвет на TIMSS: участие в разработването на рамката и спецификациите на математическия тест на TIMSS-2003 (TIMSS Assessment Frameworks and Specifications 2003), на TIMSS-2007 (TIMSS 2007 Assessment Frameworks) и TIMSS-2011 (TIMSS 2007 Assessment Frameworks). Той е един от авторите на тези рамки, както и на не малка част от задачите, включени в математическия тест на TIMSS. Полезни са описанията на класификациите за всяка представена в главата задача.

Третата част съдържа главите осма, девета, десета единадесета и дванадесета. В тях се разглеждат основни процедури за обработката и анализа на данни от широко мащабни ОДИ. Осма глава описва така наречените *извадкови тегла* – възможност да се вземе пред вид вероятността, с която учениците се избират в извадката. Тук няма съществени авторски приноси, но написаното е важно за разбирането на следващите глави. В девета глава се представя начин да се отчете сложността на извадката при пресмятане на стандартното отклонение.

Приносът на автора е в пресмятането на някои важни конкретни стойности свързани с националната извадка за България в контекста на изследването TIMSS. Десета глава разглежда една възможност за по-адекватно описание на параметрите на популацията, от която е изследвана сложна извадка и е използван непълен свързан тестов дизайн. Пълното описание и разбирането на правдоподобните стойности, на които е посветена десета глава, изискват сложна математическа и статистическа теория, която авторът съзнателно не представя в този труд. Тази теория е разработена от други и той няма приноси към нея. Един от приносите на тази глава е, че сложни за разбиране математически формули са представени на сравнително разбираем език. Друг принос е използването на тази съвременна методика от автора за анализ на данни от изследването TIMSS. В единадесета глава се разглеждат някои йерархични линейни модели за анализ на данни от широко мащабни оценявания ОДИ. Приносът на автора е използването на тези нетрадиционни методи за анализ на данни с цел получаване и представяне на някои резултати от данните на TIMSS за България. Дванадесета глава описва начини за количествено и качествено описание на резултатите на ученици от направено оценяване. Дават се методи за построяване на скали и норми, за определяне на опорни точки и качествено описание на различни нива на постижения за популацията ученици. Систематизиран е опитът на автора в участието му като експерт в тази дейност на TIMSS. Приносът му в този труд е в систематичното описание на начините за конструиране на скали и норми, на метода на опорните точки и неговото използване за получаването на някои изводи от данните на TIMSS за България.

Трите глави в четвъртата част са посветени на международни изследвания, свързани с подготовката на учители по математика. Тринадесета глава обсъжда основни елементи от дейностите на учителя по математика и необходимостта от изследване в тази посока. Приносът на автора за някои мисли относно така наречените „математически знания, необходими за преподаване на математика”. В четиринадесета глава авторът представя собствени приноси от участието си в основните дейности на изследванията MT21 и TEDS-M. Той е един от авторите на рамки те на тези изследвания, както и на част от задачите, някои от които са включени в тази глава, заедно с техните спецификации. Петнадесета глава представя резултати от участието на България в изследването MT21. Приносът на автора е в анализа и тълкуването на тези резултати.

В заключението се обсъжда важността на въпроса за съвместимост на теоретичните модели за обработка и анализ на данните от широко мащабни ОДИ с реалните експериментални данни. Цитират се данни от съвместно изследване на автора с чуждестранен колега относно съгласуваността на българските данни с моделите, използване в TIMSS. Общото заключение е, че методите, технологиите и моделите за измервания и анализ на данни трябва да бъдат внимателно конкретизирани към специфичните условия, за да могат най-адекватно да опишат и обяснят експерименталните данни, събрани в съответната ситуация.

ПЪРВА ЧАСТ
НАЦИОНАЛНИ И МЕЖДУНАРОДНИ ОЦЕНЪЧНО-ДИАГНОСТИЧНИ ПЕДАГОГИЧЕСКИ ИЗСЛЕДВАНИЯ

В тази част се описват основните характеристики на националните и международните ОДИ. През последните няколко десетилетия те станаха широко разпространен начин за оценка на качеството на образованието. От методологична гледна точка тези оценявания имат две важни характеристики. Първо, те променят традиционното използване на тестовете за оценка на *индивидуалните* ученически постижения към прилагането им за получаване на информация за постиженията на *цяла популация* ученици, или за *цялостната образователна система*, или за отделни нейни елементи. Второ, те преместват тежестта на преценката за качеството на образованието от оценка на входа (учебните планове, учителите, училищните условия и др.) към оценка на изхода (знанията и уменията, които учениците са получили в училище) (Kellaghan, T., & Greaney, V., 2001).

Целта на националните ОДИ е да опишат ученическите постижения в определен момент от обучението, и/или промените, които настъпват в тях, за даден период от време. Въз основа на тези оценявания може да се опише постигането на националните образователни стандарти по основните предмети. Също така могат да се сравняват постиженията на основни групи ученици от популацията (момичета и момчета, ученици в големи и малки населени места и др.).

Международните ОДИ дават възможност да се правят сравнителни оценявания на ученическите постижения между участващите образователни системи. Съществена особеност за тях е необходимостта тестовият инструментариум да бъде достъпен за всички участващи държави, което налага ограничение върху рамката и спецификациите на теста. Така тестът може да не е напълно адекватен на спецификата на обучението във всяка участваща държава.

Понякога се провеждат и така наречените „регионални оценявания” на ученическите постижения, в които участват държави от определен географски район (например Балканския полуостров). В големи държави като САЩ, Канада, Бразилия и др. се провеждат „под-национални” оценявания, в които участват само няколко щата (провинции, региона) от държавата.

ПЪРВА ГЛАВА
НАЦИОНАЛНИ ОЦЕНЪЧНО-ДИАГНОСТИЧНИ ИЗСЛЕДВАНИЯ

Целта на националните ОДИ е да се опишат ученическите постижения по даден предмет или предметна област, така че да се получи обобщена информация за образователната система като цяло, или за отделни нейни елементи (за даден регион, училище, клас или възрастова група). Обикновено това се прави чрез тест за постижения по даден предмет или предметна област за определената възрастова група (популация) ученици. Учители, директори на училища, понякога и родители, също дават информация за учебната и семейната среда чрез специално подготвени въпросници от тип анкета.

Националните ОДИ са различни в различните образователни системи, но те имат и редица общи черти. Навсякъде се правят национални оценявания за постижения по език и по математика, които се считат за основни учебни предмети. Много държави правят оценявания по природни науки, обществени науки, втори език и др. Някъде националните оценявания се провеждат всяка година, на други места – през период от няколко години. Всички национални оценявания търсят отговори на следните важни въпроси:

- Как се обучават учениците по даден предмет или предметна област (по отношение на официалните общи очаквания, целите на обучението и др.)?
- Има ли теми от оценявания предмет, по които учениците имат ясно изразен напредък и такива, по които те очевидно изостават?
- Има ли подгрупи ученици, които изостават в обучението по предмета? (Например, има ли статистически значима разлика между постиженията на момичета и момчета; на ученици в големи и малки населени места; на ученици от различни етнически групи и др.)
- Кои фактори (училищни, семейни и др.) влияят на успеваемостта на учениците?
- До колко се постигат националните стандарти както по отношение на ученическите постижения, така и по отношение на ресурсите за обучение?
- Как се променят ученическите постижения с течение на времето? (Този въпрос е много важен, когато се обмислят образователни реформи.)

Организацията на националните оценявания на ученическите постижения се подчинява на следните основни дейности, които обикновено се извършват в указания ред (Kellaghan, T., & Greaney, V. 2001).

- Министерството на образованието (МО) или аналогична на него институция назначава организация за провеждане на оценяването. Това може да е отдел към МО или външна организация (например, научна професионална организация, звено към университет, др.). МО осигурява финансирането.
- МО и членове на организацията, която ще провежда оценяването, определят основните цели и задачи на оценяването, както и важните въпроси, на които ще се търси отговор.
- МО и/или назначен Управителен съвет определя възрастовата група (популацията) на оценяването.
- Определя се областта/предмета на оценяване (математика, език, грамотност, и др.).
- Организацията, която провежда оценяването, подготвя инструментариума (тестове, въпросници, анкети, указания и др.).
- Организацията, която провежда оценяването, прави извадка от училища, ученици, учители и т.н. и с помощта на МО установява връзка с тях.
- Организацията, която провежда оценяването, прави инструктаж за техническото провеждане на теста.
- Провежда се апробация и се оформят крайните варианти на изследователските материали.
- На определена дата (дати) се провежда основното оценяване.
- Всички изследователски материали се връщат в организацията, която провежда оценяването.
- Данните се въвеждат в компютър и се прави статистически анализ.
- Подготвят се примерни отчети, които внимателно се четат и коментират от специалисти на организацията, която провежда оценяването и от МО.
- Подготвят се официалните отчети.
- Резултатите се публикуват и се разпространяват в медиите.

За успеха на едно национално оценяване е важно то да изследва не само важни области/предмети, но и да фокусира на значими знания и/или умения, които се изучават/придобиват по този предмет. Важно е да се направи прецизна извадка. Важно е събраните данни реално да описват обучението по оценяваната област/предмет. Важно е да се направи статистически анализ, който е адекватен на начина на събиране на данни и на тяхната структура. Всички тези дейности изискват значителни ресурси (материални и човешки) и голяма политическа подкрепа (от МО).

Трябва да се подчертае, че националните изпити не могат да играят ролята на национални ОДИ. По редица причини изпитите не дават информацията, която е съществена за оценяванията. Първо, изпитите имат за цел да удостоверят завършване на образователна степен или селекция за постъпване в по-висока степен на образование. Поради това те нямат за цел (и не го правят) да покрият (почти) напълно учебното съдържание по предмета на изпита. Второ, учениците, които се явяват на изпити, се сменят всяка година, което не дава възможност за сравнения от година в година. Трето, резултатите от изпита са от голяма важност за индивидуалния ученик и това в известна степен намалява валидността на изпита като инструмент за ОДИ. Обратно, обикновено при националните оценявания резултатите не са от значение за индивидуалния ученик. Четвърто, националните изпити обикновено се правят в по-високите класове на обучение, докато образователната система се нуждае от информация за състоянието (национално ОДИ) в по-ранен стадий. Пето, информацията за учебната среда, в която става обучението и която е важна за обяснението на резултатите, не се събира при провеждане на изпит. В Таблица 1 са описани основните различия между национално ОДИ и национални изпити.

Таблица 1. Разлика между национални ОДИ и национални изпити

	Национални оценявания	***Национални изпити***
Цел	Да информира хората, вземащи образователни решения	Да дава сертификати или да подбира ученици
Честота	През определено време (напр. веднъж на 4 години) за даден предмет	Ежегодно или по-често
Продължителност	Един или два дни	До няколко седмици
Кои се подлагат на тест?	Обикновено извадка от ученици от популацията	Всички ученици, които искат да се явят на теста (и отговарят на условията)
Формат	Обикновено задачи с избираем или кратък отговор	Обикновено задачи с избираем и разширен свободен отговор

	Национални оценявания	***Национални изпити***
Важност за индивидуалния ученик	Не е важно	Изключително важно
Покриване на учебната програма	Обикновено частично покрива един-два предмета	Покрива много предмети
Ефект върху дейността на учителя	Слаб ефект	Силен – учителят се стреми да преподава онова, което е включено в изпита
Търсят ли учениците допълнителни уроци?	Почти не	Много често
Научават ли учениците резултатите си?	Рядко	Да
Събира ли се допълнителна информация от/за учениците?	Много често чрез въпросници и анкети	Много рядко
Обработка на резултатите	Обикновено са нужни сложни статистически процедури	Обикновено се използва проста схема за оценяване
Влияние върху посещаемостта	Няма влияние	Лошите резултати водят до напускане на училище
Възможност за мониторинг на постиженията с течение на времето	Големи, ако тестът е построен с такава цел	Няма възможност

При планирането на национални ОДИ трябва да се вземат редица важни решения. По-долу ще разгледаме някои основни въпроси, които изискват такива решения.

1. *Кой трябва да има политическо, научно и методическо ръководство на националното оценяване?* Разпространена практика е Министерството на образованието (МО) или аналогичната за страната институция да назначи Национална комисия. В комисията участват представители на МО, на организацията, която провежда оценяването, изтъкнати учени в съответните области, представители на големи етнически, религиозни, лингвистични и др. групи. Тази комисия има особено активна функция в началото на оценяването докато се изяснят и приемат основните научно-методически и политически въпроси, които стоят пред оценяването. Комисията не изпълнява технически функции, а има ролята на консултативен орган, който определя основните решения от политически и научно-методически характер.
2. *Кой провежда националното оценяване?* Това е организация, която е специално назначена от МО за провеждане на ОДИ. Възможностите за избор на организация са най-разнообразни, като се започне от поделения на МО, мине се през научни звена на институти/университети и се стигне

до неправителствени професионални организации. Различните възможности предлагат различни предимства и недостатъци. Важно в случая е да се избере организация, която има отлични професионалисти с нужния опит; организация, на която може да се има доверие; организация, която има добри отношения с МО и лесно може да комуникира с него. Организацията, която провежда националното оценяване, трябва да предложи екип от най-разнообразни специалисти, включително компютърни специалисти и такива, които имат опит в поддръжка и обработка на големи бази данни.

3. *Кой провежда (администрира) теста с учениците?* Известни са различни практики. В някои държави се привличат завършващи студенти (Колумбия), на други места (Замбия) това правят регионалните инспектори. В Аржентина всяка провинция назначава независими квестори. Може да се ползва и вариант учители от дадено училище да провеждат теста в училища от други региони. Не бива да се допуска учители да провеждат теста на собствените си ученици.
4. *По кое време на обучението си в училище (кой клас) да се оценяват учениците?* Отговорът на този въпрос зависи в голяма степен от организацията на учебния процес. Първият въпрос е дали популацията да се дефинира на основата на „клас" или на основата на „възраст". Когато образователната система е така организирана, че в класовете са концентрирани сравнително еднакво възрастни деца, желателно е да се използва дефиницията с „клас". Ако в класовете има ученици с най-разнородна възраст, по-добре е популацията да се дефинира чрез „възраст". Обикновено политиците в образованието се интересуват от резултатите на образователната система в края на определени етапи/степени. В САЩ, например, това става в края на класовете 4, 8 и 12, в Колумбия – в класовете 3 и 5. Малко страни правят национални оценявания преди завършване на началния етап на обучение (класовете 1-4).
5. *Цялата популация ли да се изследва или само (представителна) извадка от нея?* По технологични и финансови причини повечето национални оценявания се правят с (представителна) извадка. Целта на националните оценявания е да се получи обобщена информация за цялата популация, а не за отделните ученици. Затова не се налага да се харчат много пари и

да се правят сложни схеми за организация и провеждане на теста върху всички ученици от популацията. Трябва да се обърне внимание обаче, че получаването на необходимите статистически оценки за популацията от извадката неминуемо изисква използването на сложни съвременни статистически процедури и специализиран софтуер.

6. *Какво да се оценява?* Основно се оценяват знания и умения, които учениците са придобили по важни области/предмети през определен период от престоя си в училище. Всички национални оценявания съдържат предметите език и математика. Включването на други предмети варира от държава в държава и е свързано с особеностите и организацията на конкретната образователна система. Напоследък се забелязва, че в европейските страни ударението се поставя на оценяване на компетентности (езикова грамотност, математическа грамотност и др.), а не на възпроизвеждане на знания (Key competences for lifelong learning in Europe, 2005). Така, вместо оценяване по даден учебен предмет се преминава към оценяване на основни компетентности. Каквото и да се оценява обаче, важното е да има ясна рамка за оценяването, върху която се изработва теста и която служи, за да се анализират получените резултати. Понякога такава рамка може да се съдържа в учебните планове, ако в тях ясно е операционализирано какво и как се оценява като резултат от обучението. Практиката показва обаче, че обикновено учебните планове не съдържат такава част и екипът трябва да разработи рамката за оценяване. При националните оценявания обикновено се събира информация за влиянието на учебната среда чрез въпросници/анкети към учениците, учителите, директорите и (понякога) родителите.
7. *Как да се оценяват ученическите постижения?* При подготовката на инструментариума (теста) обикновено се поставя въпросът дали измерването е нормативно или критериално. На практика при националните оценявания тестът се конструира и по двата принципа. Започва се с описание на рамката на измерването, като се дават спецификации на очакваните знания и умения, които ще се измерват. След това се пишат тестовите задачи, за да се измери до каква степен учениците са постигнали тези знания и умения. През последните 15-20 години се промени технологията на скалиране. Използването на вероятностното моделиране (IRT)

дава възможност различни групи ученици да работят по различни части от теста (виж четвърта глава). Това дава възможност да се увеличи многократно броят на задачите в теста, а оттам – да се разшири тематиката на целия тест.

8. *Колко често да се правят национални оценявания?* Отговорът на този въпрос зависи от целите на оценяването и от традициите на съответната държава. Националното оценяване на ученическите постижения е скъпа, продължителна и трудоемка дейност. Освен това, образованието е „консервативно", в смисъл, че не могат да се очакват големи промени в кратък срок от време. Тези съображения водят до идеята национални оценявания да се правят през определен период от време (например на всеки 2 или 4 години). Възможно е да има редуване, като например всяка четна година да има национално оценяване по математика, а всяка нечетна – по език. В подобна схема могат да се вкарат и други предмети/области. В някои образователни системи (например Англия и Чили) една от целите на националните оценявания е да играят роля на ежегодна отчетност за училищата и учителите. При наличието на такива цели, естествено е национални оценявания да се правят всяка година.
9. *Как да се прави отчет на ученическите постижения?* Практиката показва голямо разнообразие в това отношение – като се започне от отчет за броя (процента) ученици, отговорили правилно на всяка задача (нормативен подход) и се стигне до процента ученици, постигнали определени нива на постижение (например „основно", „средно" и „напреднало") (критериален подход).
10. *Какъв статистически анализ трябва да се направи?* Отговорът зависи от поставените политически и научни въпроси. Почти винаги се прави анализ на постиженията по различни групи (момичета-момчета, големи и малки населени места), и/или по региони (училища). Анализите трябва да отчитат и различията в учебната среда – информация, събрана от въпросниците и анкетите с учители и директори. Често пъти анализът се прави със сложни и съвременни математико-статистически процедури, които отчитат особеностите на извадката и на структурата от данни (виж трета глава).

11. *Как да се разпространяват и използват резултатите от национално оценяване?* (Bankov, K., O'Sullivan B, 2008) Отчетите трябва да се приготвят без излишно забавяне. Разбира се, нужно е внимателно да се прегледат всички изводи и анализи за евентуални грешки в данните и тълкуването им. Но колкото по-бързо стигнат окончателните анализи до потребителя (Министерството на образованието), толкова по-добре. Преди време се считаше, че дългите отчети с множество технически подробности са за предпочитане. Един такъв отчет, разбира се, е необходим, но все повече се ценят по-кратки отчети от типа на резюме, от което заетите с много дела политици лесно могат да получат синтезирана информация. Подготвят се и сравнително кратки отчети за определен кръг читатели (например за групите, работещи по нови учебни програми и др.). На някои места се изготвят отчети за отделните региони от страната.

Провеждането на национални ОДИ е многостранна и сложна дейност. Тя трябва да се извършва от многочленен екип с отличен ръководител. Въпреки това, допускането на различни грешки съпътства всяка практическа дейност. Най-често допусканите грешки, които са систематизирани по-долу.

Грешки при планирането:

- Недостатъчно финансиране, особено на дейностите, свързани с писането на отчетите и разпространението на резултатите.
- Липса на Национална комисия (или неефективна такава), която да бъде главен консултант и да решава най-важните политически, научни и методически въпроси на оценяването.
- Липса на ясно изразено желание и политическа подкрепа от страна на Министерството на образованието или аналогичната му институция.
- Недостатъчна ангажираност на големи маси, като например учители, в организацията на оценяването.
- Пропускане да бъдат обхванати големи групи от популацията (например частните училища).
- Поставяне на нереалистични изисквания към тестовия бал (например повишаване на средния тестов бал с 25% за 4 години).

- Недостатъчно планирано време за подготовката на теста.

Грешки при провеждане на оценяването:

- Подготовката на теста се поверява на екип от хора, които не са добре запознати с нивото на знания и умения на учениците.
- Неадекватно разпределение на задачите според учебния план.
- Не се провежда пилотно тестване (апробация).
- Броят на задачите в крайния вариант на теста е много голям (или много малък).
- Липса на баланс в задачите по отношение на това, което учениците трябва да знаят и могат и онова, което те в действителност знаят и могат (което води до труден или лесен тест).
- Няма ясна дефиниция на конструкта, който се оценява.
- Включването в пилотния тест на недостатъчно задачи от формат, за който се предполага, че учениците не са добре запознати с него.
- Не се обръща сериозно внимание на многократно препрочитане на тестовия материал, въпросниците/анкетите и указанията за провеждане преди те да бъдат официално отпечатани.
- Използване на стари данни за получаване на извадката.
- Използване на неадекватни процедури за получаване на извадката.
- Неадекватна инструкция (или липса на такава) за техническото провеждане на теста с учениците.
- Голяма външна намеса по време на провеждане на теста (например присъствието на учителя на класа).
- Учениците стоят много близко един до друг и си разменят информация.
- Възможност учениците да работят повече от определеното време.
- Възможности за подсказване, преписване и др.

Грешки при анализа на данните:

- Използването на неподходящи статистически методи (включително не използване на извадкови тегла).
- Използване на резултати, получени върху малък брой данни.

- Сравняване **само** на средните стойности на учениците върху два различни предмета и правене на изводи за по-голямата им успеваемост по един от предметите.
- Не се дава (публикува) стандартната грешка за всяка статистика.
- Сравняване на училищата подредени по получената средна стойност на учениците от теста, без да се вземат пред вид особеностите на учебната среда за всяко училище. Още по-голяма грешка е, ако такъв списък се публикува в общественото пространство.
- Използване на необосновани причинно следствени връзки (например връзка между броя на учениците в класа и постижението на класа).
- Сравняването на тестови балове получени през определен период от време (например 2 години) когато не са използвани еквивалентни (паралелни) теста.
- Сравняването на тестови балове получени през определен период от време (например 2 години) получени с еквивалентни (паралелни) теста, но без да се вземе пред вид промяната на учебната среда за този период.

Грешки при писане на отчетите:

- Представяне на отчет, който набляга предимно на техническите и технологичните особености на оценяването.
- Липса на няколко ясно изразени основни заключения.
- Препоръки свързани с конкретна променлива, дори и когато има съмнения във валидността на резултатите по тази променлива.
- Липса на връзка между получените резултати и реалните учебни планове, учебниците, начините за подготовката на учителите и др.
- Не се обяснява кога разликата в средните стойности е (или не е) статистически значима.
- Късно представяне на отчетите (по време, когато вече не могат да повлияят за вземане на решения, за които е направено оценяването).
- Основните резултати не се представят пред съответните заинтересувани групи от хора.

Грешки при използване на резултатите:

- Резултатите не се правят достояние до големи обществени групи, които за заинтересовани от тях (ученици, учители, родители, научна общественост и др.).
- Игнориране на резултатите, когато трябва да се вземат политически решения.
- Големи групи от хора, за които са предназначени резултатите (учители, директори и др.), не се съобразяват с тях и с евентуалните последици, направени в изводите.
- Липса на препоръки от екипа към следващи национални оценявания.

България не е от държавите с трайни традиции в провеждане на национални оценявания на ученическите постижения. Въпреки това са правени опити, макар и епизодични, за провеждане на такива оценявания. Няма да се спирам на периода преди 1989 година, но ще направя коментар на случващото се през последните 20-тина години.

С настъпването на демократичните промени в България (1989) острата необходимост от нова образователна политика излезе на преден план пред Министерството на образованието. С изменението на основния образователен закон (ЗНП, 1991) се поставя началото на изработването и въвеждането в училищното обучение на Държавните образователни изисквания (ДОИ). Единственият подход за реални демократични промени в образованието е ясна стратегия, която може да премине в добре обмислена образователна реформа. Реформа, която трябва да бъде реално осъществима, систематична в целите и намеренията си, реформа на консенсус на всички заинтересовани (учители, ученици, родители, деятели на образованието и др.).

От друга страна, необходимо е да бъде заявена и ясна политическа воля, както и осигуряването на финансово обезпечаване на дейностите свързани с осъществяването на тази реформа.

Основни и важни задачи в този процес са разработването и въвеждането на ДОИ за учебно съдържание (ДОИ за УС), Закона за степента на образование, общообразователния минимум и учебния план (ЗСООМУП) и ДОИ за системата за оценяване (ДОИ за СО). Последното включва както текущото, така и заклю-

чителното оценяване на постиженията на учениците. Поради редица субективни и обективни причини – и най-вече липсата на ясно заявена политическа воля, реформата в българското образование се случва бавно, липсва последователност в действията и много от нещата се правят „на парче".

В периода 1998-2011 могат да бъдат посочени няколко ***проекта на Министерството на образованието*** за проверка в национален мащаб на образователните процеси. По-долу са описани три от тях, в които авторът е взел активно участие.

- *Национално оценяване на постигането на "общообразователния минимум по БЕЛ и по Математика в края на IV кл., VIII кл. и XI кл."* (1998) (Банков, К., 1997, 1998-Б, 1998-В; Георгиева и Банков, 1997; Георгиева, Банков и Тодорова, 1998.).

Оценяването се осъществява по поръчка на МОМН. В този процес са въвлечени специалисти от Научноизследователския институт по образование (НИО), работни екипи от специалисти по математика и български език и литература от СУ, както и практикуващи учители.

Целта на това оценяване е да установи до каква степен се постига общообразователният минимум съгласно учебното съдържание от учениците в различните образователни етапи и степени. Оценяването се осъществява на основата на извадка от училища и ученици. Провежда се от учителите в избраните училища по зададени критерии от съответните работни екипи. Технологията на работа в известна степен следва теоретично известните статистически модели. Започва се с изясняване на целите на оценяването, създава се спецификация като гаранция за валидността на резултатите, и въз основа на нея се съставят задачите. Изводите се правят чрез статистическа оценка на неизвестни параметри (предимно средна стойност) за цялата популация. Целият проект е осъществен за половин година. Това време е по-малко от необходимия минимум за провеждане на такъв род оценяване. Поради това не е направено пилотно оценяване и липсва информация за надеждността на измерителния инструментариум. В резултат на направените анализи МОМН започна разработването на Закона за степента на образование, общообразователния минимум и учебния план (ЗСООМУП) както и Държавните образователни изисквания за учебно съдържание (ДОИ за УС).

- *Изработване и въвеждане на тестова система за прием на учениците след завършен VII клас в Профилираните гимназии* (1998-2001) (Банков, 1999; Банков и Витанов, 1999, 2000, 2001.).

Целта на това оценяване е да направи класиране на кандидатите за профилирани гимназии като измери с тест от задачи с избираем отговор знанията им в четири области: български език и литература, математика, природни науки и обществени науки. Тази цел поставя проекта по-близо до подготовка за провеждане на национален изпит. Причината да го цитирам тук като вид национално оценяване е, че през голяма част от подготвителната фаза този проект играеше ролята на пилотно ОДИ по казаните по-горе предметни области. Използваната теоретична основа е класическата теория на тестовете. През първата година са уточнени спецификациите на теста, създадена е банка от тестови задачи, проведени са няколко пилотни оценявания и са установени качествата на инструментариума. Важно е да се подчертае публичността на подготовката и реализирането на проекта.

През този период в проекта са въвлечени много учители, ученици и техните родители. Тази публичност до голяма степен, помогна на всички заинтересовани лица да се запознаят с новата система за оценяване. Това включва: открита дискусия относно изработването на задачите в различните културно образователни области, както и критериите, по които ще се оценяват тестовете, а така също и процедурите за провеждането на приемния изпит в реални условия. Проведени бяха семинари, работни срещи с учители, експерти и преподаватели от СУ.

Много от готовите примерни тестове бяха публикувани в пресата, издаден бе справочник, които включва не само тестови примери, но и указание за тяхното решаване и оценяване. Това допринесе, до известна степен, както за самоподготовка на учениците, така и за подготовката на учителите.

Следващите две години (2000 и 2001 година), заедно с попълване на банката от задачи и тяхното апробиране, се провеждат и два (всяка година по един) реален изпит за прием в профилирани гимназии. По време на този проект изцяло се прилагат на практика методи и технологии за измерване съобразени с класическата теория на тестовете. Натрупан е полезен опит в теоретично и практическо отношение. Разбира се, не всичко е правено идеално от самото начало, но с течение на проекта се забелязваше чувствително подобрение на ка-

чеството на работа и съгласуваността между отделните групи в екипа. След всяка приемна кампания е изготвен анализ на оценяването.

- *Национално стандартизирано оценяване в IV, V, VI и VII клас, провеждано от 2007 г.* (Банков и Витанов, 2010.).

В изпълнение на Националната програма за развитие на училищното образование и предучилищното възпитание и подготовка (2007-2015) през 2008 Министерството на образованието, младежта и науката (МОМН) започна въвеждане на система за национално стандартизирано оценяване. Вече е нормативно определено от МОМН да има ежегодни национални оценявания при завършване на клас, образователен етап или степен, както и за използване на резултатите им при продължаване на образованието (за класиране). Основна цел на националното стандартизирано оценяване е задължителното оценяване на знанията и уменията на учениците в края на всеки образователен етап или степен (IV, VII, XII клас) чрез национални стандартизирани изпити.

До 2010 година тези оценявания се провеждаха в IV, V, VI и VII клас, а от 2011 година – само в IV и VII клас (защото това са класовете, в които има завършване на образователен етап или степен). Темите (тестовете) се подготвят в Центъра за контрол и оценка на качеството на училищното образование (ЦКОКУО), като за подготовката им се привличат външни специалисти. Авторът на този труд е ръководил подготовката на темите по всички предмети за IV клас през 2008, 2009, 2010 и 2011 година и е участвал в екипа по за оценяването по математика в VII клас през 2010 и 2011 г.

ВТОРА ГЛАВА
МЕЖДУНАРОДНИ ОЦЕНЪЧНО-ДИАГНОСТИЧНИ ИЗСЛЕДВАНИЯ

Международните ОДИ имат много общи черти с националните такива. Двете използват едни и същи процедури за извадка, за подготовка на инструментариум и събиране на данни, за анализ и обработка на данни. Те имат също много общи неща в техните цели: да изследват как и какво научават учениците в училище; да установят силните и слабите страни в знанията и уменията на учениците; да сравнят отделни големи групи ученици помежду им по отношение на знанията и уменията по оценявания предмет; да потърсят връзка между постиженията на учениците и условията на учебната, домашната и обществената среда; да изследват промяната на ученическите постижения с течение на времето.

Основната разлика между международните и националните ОДИ е в това, че в международните участват няколко образователни системи. Така тези оценявания имат за цел да дадат сравнителна информация за постиженията на учениците за всяка страна в сравнение с другите участващи в оценяването. Международните ОДИ са основани на идеята, че сравнителната информация ще доведе до по-дълбоко разбиране на факторите (различни за различните образователни системи), които определят националните различия в ученическите постижения.

За интереса към международните ОДИ допринася още и увереността, че повишаването на знанията на човечеството (особено тези по математика и природни науки) играе важна роля за икономическия растеж. Затова целта е да се повиши средното световно постижение на учениците като предпоставка за икономически растеж (Raminez, F.O., Luo, X., Schofer, E., & Meyer, J.W. , 2006).

Заслужават да се отбележат най-известните международни оценявания, които се провеждат в момента: TIMSS, PIRLS, PISA, TEDS-M.

Нараства броят на образователните системи, които участват в международните оценявания. През 80-те години около 20 държави участваха в едно международно оценяване. През 1991 г. 32 държави участват в IEA Reading Literacy Study. През 2003 г. 52 държави участват в TIMSS 2003 и 41 държави участват в PISA. Почти всички международни оценявания имат за цел да изследват „тенденции”, т.е. промяна на ученическите постижения за определен период от време (3 години за PISA, 4 години за TIMSS и 5 години за PIRLS).

По-долу са изброени някои от „ползите” при участие в международни ОДИ.

1. Международните оценявания дават сравнителен анализ на обема знания, които учениците изучават в участващите държави, както и други ресурси, свързани с обучението. Това често пъти се оказва сериозен стимул за включване (или изключване) на учебен материал от учебните планове и/или за повишаване на учебните ресурси. Например, в резултат на слабото представяне на Южна Африка в TIMSS, беше взето решение на национално ниво да се увеличат ресурсите за обучение по математика и природни науки (Reddy, V. 2005).
2. Сравняването на резултати от различни държави показва какво е постижимо от учениците в международен мащаб и къде се намира съответната държава в това отношение. Например, „шокиращо” слабите резултати на Германия в TIMSS 1995 и PISA 2000 доведоха до много сериозни дебати за качеството на образователната система в Германия и се набелязаха съответно промени, както за средното училище, така и за подготовката на учители (Blomeke, S. 2007).
3. Международните оценявания са изключителна лаборатория за педагогически изследвания. Богатата информация за организацията на учебната среда и условията за учене в различните държави раждат много хипотези относно факторите, които обуславят и влияят на ученическите постижения. Всяка от тези хипотези се нуждае от своето научно оценяване.
4. Международните оценявания водят до научно осмисляна не редица понятия от педагогиката (например какво е математическа грамотност) в международен контекст (Банков, 2006).
5. Участието в международни оценявания води до увеличаване броя на местните специалисти с богати знания и опит в провеждането на сравнителни ОДИ на всякакво ниво. Тези оценявания са своеобразен университет за повишаване на нивото на кадрите, работещи в тази област.

Въпреки тези очевидни „ползи” от участието в международни оценявания, има някои проблеми, които трябва да се имат пред вид преди да се вземе окончателно решение за участие в международно оценяване.

1. Въпреки че учебните планове на различните държави по основните предмети имат много общи неща (особено в началния етап на обучение), съществуват редица различия, които затрудняват построяването на рамката на измерването. В крайна сметка рамката е „разумен компромис" от онова, което изучават учениците в участващите образователни системи. Така, измервателният инструмент (тестът) се оказва по-адекватен за обучението в някои държави и по-неадекватен за други. Например, за математическия тест в TIMSS 2003 процентът на задачите, които се изучават от учениците в участващите държави варира от 33% (Южна Африка и Мароко) до 98% (САЩ). За България – 78% от задачите в математическия тест на TIMSS 2003 са от задължителната учебна програма до 8 клас (Mullis, I., at al., 2004).
2. В международните оценявания чрез въпросници (анкети) за ученици, учители, директори, понякога и родители се събират данни за организацията на учебния процес и особеностите на учебната среда. Поради големите различия в участващите държави, получените от тези въпросници резултати нямат силата на причинно-следствени връзки по отношение на ученическите постижения. Корелацията между променливите, описващи учебната среда, и ученическите постижения варира много от държава в държава. Информацията, която тези оценявания дават в това отношение, може да се интерпретира в контекста на съответната образователна система и трудно може да се обобщи за други такива.
3. Често възникват проблеми свързани с необходимостта тестовият материал да се превежда на различни езици. Когато трябва да се прави сравнение между постиженията на държави по една и съща тестова задача, трябва да се има пред вид, че особеностите при превода могат също да са причина за получените различия в постиженията.
4. Популацията от ученици и извадките на различните държави може да не са сравними. Повечето международни оценявания допускат от популацията да се изключат известен процент ученици (например, за TIMSS е допустимо да се изключат 5% от учениците в популацията). Това обикновено са ученици с увреждания или ученици в отдалечени и трудно достъпни райони и др. Изключването на част от учениците, макар и в малък

процент, може да доведе до проблеми при сравняване на популациите и извадките.

5. В международните оценявания участват държави с голям диапазон на ученическите постижения. Това трудно може да се вземе пред вид при конструкцията на теста. Не малка част от задачите се оказват много трудни за учениците в някои държави и много лесни за учениците в други държави. Това води до ограничаване на дисперсията на баловете.
6. Резултатите от международни сравнителни оценявания винаги съдържат подредба на участващите държави по средния бал на учениците в тях. На тези данни не трябва да се набляга. Подредбата на държавите говори малко на читателя, ако не са обяснени факторите, които имат влияние върху ученическите постижения в съответната държава. Освен това, мястото на дадена държава в този списък зависи много от другите участващи държави. Ако в оценяването има повече държави, чиито ученици получават високи балове, мястото на дадена държава в списъка ще бъде по-назад, без причината за това да е спад в ученическите постижения в самата държава.
7. Свързан с предишната точка е и проблемът с политическия риск, който се появява от „лошото" представяне нададена държава. Това е причина някои държави да отказват да бъдат включени в списъка на държавите, подредени според ученическите постижения.
8. За държави, които нямат голям опит в национални оценявания на ученическите постижения и/или нямат нужния потенциал от хора и администрация, е трудно да се впишат във времето, дадено от международния център за извършване на дейностите по оценяването. Често участието в международни оценявания е много скъпо и трябва внимателно да е прецени дали съответната държава може да си позволи това участие.

По-долу са описани няколко от големите международни сравнителни ОДИ, които се провеждат в момента или са приключили преди 2-3 години. Дадена е информация за участието на България, както и за дейността на автора в някои от тях.

TIMSS (http://isc.bc.edu/)**.** Съкращението TIMSS е съкращение на името Trends in International Mathematics and Science Study, което означава Световни

тенденции в обучението по математика и природни науки. Оценяването е с четиригодишен цикъл. До сега са проведени пет оценявания: TIMSS-1995, TIMSS-1999, TIMSS-2003, TIMSS-2007 и TIMSS-2011, като България е участвала в първите четири от тях.

TIMSS се организира и провежда от Международната асоциация за оценяване на постиженията в образованието (*IEA – International Association for the Evaluation of Educational Achievement*). Вече половин век IEA провежда международни сравнителни оценявания по различни учебни предмети. Всяко оценяване става едновременно в няколко държави, като се използват единни методики и анализ на данните както на национално, така и на международно равнище. България членува в IEA от 1991 г.

Както всички оценявания на IEA, така и TIMSS се провеждат под строг контрол, за да се гарантира висока валидност, надеждност и сравнимост на резултатите. Извадката се наблюдава от специален център. Преводът на материалите се контролира на няколко нива. Стриктно се следи и самото събиране на данните. Специално обучени хора наблюдават как се провежда цялото оценяване.

Страни-членки на IEA използват по различен начин резултатите от отделните оценявания. Самият факт обаче, че в тях винаги участват както високо развитите индустриални, така и други държави, доказва голямата заинтересованост от получаване на сравнителни данни за ефективността на отделните образователни системи, съизмерени с постиженията и на други страни. За образователната политика на всяка отделна държава такива данни са много важни. Резултатите от оценяванията на IEA са обобщени в над 100 научни публикации.

Ето защо TIMSS си извоюва престижа на международно признато надеждно ОДИ, което дава обективна обратна информация не само за обучението по математика и природни науки, а и за функционалността на оценяваните образователни системи. Въз основа на него сериозни световни организации като UNESCO (United Nations Educational, Scientific and Cultural Organization), OECD (Organisation for Economic Co-operation and Development), ETS (Educational testing Service), CITO-Group, NCF (National Science Foundation), NAEP (National Assessment of Educational Progress) и др. правят своите прогнози. Когато се правеха европейските индикатори за качество на образованието (1999-2001 г.), в частта си по математика и природни науки те бяха базирани именно на това

оценяване.

Редно е да се отбележи, че TIMSS не е състезание. Това е научно ОДИ, чийто основен интерес е общото състояние на обучението по математика и природни науки в участващите държави. Работи се с математико-статистически методи въз основа на представителна за страната извадка.

Целта на TIMSS е да изследва не само постиженията на учениците по математика и природни науки, но също така и националните програми, училищната и социална среда на учениците в участващите държави. Тестовете за постижения са конструирани така, че да дават информация за степента на постигане на поставените цели в националните учебните програми. Чрез въпросниците се изследват факторите, които влияят върху обучението по математика и природни науки. Основен въпрос е сравняването по различни показатели между участващите образователни системи. За страните, които са участвали в предишни фази на TIMSS, се прави анализ на тенденциите на развитие от една фаза до друга.

PIRLS (http://isc.bc.edu/)**.** Съкращението PIRLS произлиза от Progress in International Reading Literacy Study. Това оценяване на IEA се провежда на всеки 5 години. До сега са проведени три оценявания на PIRLS, съответно през годините през 2001, 2006 и 2011, като България участва и в трите. Идеята на PIRLS е да се изследват международните тенденции в грамотността на учениците в края на началното училище. България участва в PIRLS от самото начало.

Основната цел на PIRLS е да осигури информация за степента на развитие на четивните умения на учениците в края на началната училищна степен (9-10 годишни четвъртокласници), както и да идентифицира факторите, които им влияят. Структурната рамка на PIRLS разглежда *четенето с разбиране* като сложен интерактивен процес. При него на преден план се извежда *възможността да се разбират и използват тези форми на писмения език, които са наложени от обществото и/или са оценени като важни от индивида. Младите читатели трябва да могат да извличат значението от най-разнообразни текстове. Те трябва да могат да четат, за да научат нещо, за да участват в читателската общност, както и за собствено удоволствие* (Campbell, Kelly, Mullis, Martin, & Sainsbury, 2001).

В оценяването са използвани два вида основни измерителни инструменти – дидактически тестове (за диагностика на разбирането при четене) и ан-

кети (предназначени за събиране на данни относно процеса на формиране на читателските навици, отношението към четенето, както и факторите, които им влияят).

Разбирането при четене е многомерен и сложен психологически конструкт, чието измерване е изключително трудно. Една от основните причини за това е, че четенето е процес на взаимодействие между индивида и текста. Ето защо крайният резултат от четенето – степента на разбиране на смисъла и значението на определен текст, зависи не толкова от отделните характеристики на всеки от двата компонента (индивид и текст), колкото от тяхното взаимодействие.

Разработката на тестовите задачи за оценяването е съобразено с четири основни аспекта на разбирането при четене, които са възприети от екипа на PIRLS, а именно:

- Търсене и извличане на конкретна информация;
- Правене на директни умозаключения;
- Интерпретация и обобщение на факти и идеи;
- Анализ и оценка на съдържанието, езика и структурата на текста.

Текстовете, включени в тестовете за разбиране при четене, се разделят условно на два основни типа – художествени и научнопопулярни. Изборът на тези два типа текстове е свързан пряко с двете основни функции на четенето, а именно – за удоволствие и за получаване на информация.

PIRLS, както и TIMSS, използва така наречения непълен свързан дизайн (incomplete linked design) или матрична извадка (matrix sampling). Този подход позволява всеки ученик да работи само върху част от текстовете и да отговори на свързаните с тях тестови задачи. Едновременно с това за всеки текст и задачите към него трябва да се осигури представителна извадка от ученици, които са работили по него.

Използвани са два основни тестови формати – задачи с избираем отговор (1 от 4 възможни) и задачи със свободен отговор, оценявани по скали с две, три или четири степени, респ. балове. Анкетите са вторият основен метод на оценяването и са насочени към събиране на данни относно читателските навици на учениците, отношението им към четенето, най-често използваните от учителя дидактически методи, средства и форми при обучението по четене, материално-

техническата база в училищата и микроклимата в тях, квалификацията на учителите, семейната среда и социалното обкръжение на учениците.

PISA (www.pisa.oecd.org**).** През 2000 година Организацията за икономическо сътрудничество и развитие (OESD – Organization for Economic Cooperation and Development) започна едно от най-амбициозните си международни оценявания под названието PISA (Progress in International Student Achievement). То обхваща 15 годишните ученици и се провежда в три части: Езикова грамотност (2000 г.), Математическа грамотност (2003 г.) и Природонаучна грамотност (2006 г.). Тези части се редуват през 3 години, като през 2009 беше изследвана пак езиковата грамотност.

Оценяването започна само с държави членки на OECD, но още на следващата година (2001 г.) бяха поканени и страни, които не са членки на Организацията, между които и България. Така в PISA участват около 35 държави.

Методологията на PISA е близка до тази на IEA оценяванията.

PISA не изследва специфичните за всяка наука и изучавани в училище когнитивни знания, а главно и преди всичко *разбирането, рефлексията, способностите за решаване на проблеми и за приложение на знанията* като основа за перманентното учене.

Получава се информация за условията за учене, което става с помощта на специално разработени за целта въпросници. По този начин се обхващат не само усвоените знания и свързаните с тях училищни и извънучилищни фактори, но и индивидуалните *стратегии на учене, както и мотивацията на учениците*.

Правят се сравнения между резултатите от обучението и различните образователни системи, финансовите средства, както и методите на обучение. Търси се отговор на въпроса в каква степен училищните системи и учениците са способни да „компенсират" различията в социално-икономическия статус на учениците, кои са другите фактори, които обуславят различията. PISA не търси "световен майстор" например по четене, а търси отговор на въпроса в каква степен се отдава на учениците в отделните страни да усвоят онези незаменими за четенето и за живота основни компетенции, колко големи са различията между учениците вътре във всяка държава и на какво могат да се дължат те, както и каква е частта на онези изоставащи ученици, които в училище не са подготвени за своя бъдещ живот.

Организаторите на проекта са ангажирали най-добрите специалисти от участващите страни, с помощта на които е подготвена и публикувана за обсъждане „рамкова концепция“ (Frameworks). В нея са предложени съответни дефиниции на компетентността, целите, компонентите на учебното съдържание, чието усвояването ще се оценява. Посочени са също така измерителните инструменти и критериите за оценка, както и примерни задачи. Важно е да се посочи, че всички материали и документи за оценяването се обсъждат и приемат предварително с консенсус от съвета на участващите държави. Това е направено по този начин, тъй като от много други оценявания е известно, че не може да се конструира „световен курикулум”, който да е еднакво валиден за всички държави, но е възможно да се постигне единство, консенсус между специалистите по отношение на най-важните според тях компетенции, които са важни за техните ученици, за да могат те да се справят успешно с изискванията и предизвикателствата на техния бъдещ живот. Всички материали, мнения и предложения се включват във Web-страницата на оценяването. Така авторите и организаторите се надяват, че ще постигнат една жива и реална комуникация с всички заинтересовани и участници в това оценяване.

Официален представител за България в PISA е ЦКОКУО.

TEDS-M (http://teds.educ.msu.edu/default.asp) **и**

MT21 (http://usteds.msu.edu/related_research.asp).

През 2002 година Международната асоциация за оценяване на постиженията в образованието (*IEA – International Association for the Evaluation of Educational Achievement*) реши да започне оценяване на обучението на учители (TEDS – Teacher Education and Development Study), фокусирайки върху учителите по математика. През втората половина на 2003 година експертна група (член на която е и авторът на този труд) разработи концептуалната рамка на оценяването. Това е първи опит за оценяване на обучението на учители от международен мащаб. Координационни центрове на TEDS са Michigan State University (MSU) и Australian Council for Educational Research (ACER). Поради финансови затруднения обаче, TEDS стартира по същество през 2006 г.

От края на 2004 година MSU започна предварително оценяване на обучението на учители по математика MT21 (Mathematics Teaching in the 21st Century) „Учителите по математика в 21-ви век”, което беше финансирано от National Science Foundation. Целта на MT21 е да изпробва някои методики, тестови за-

дачи, въпросници и др., които по-късно да са база за TEDS-M. Нуждата от такава предварителна (пилотна) дейност се обосновава с факта, че няма разработени методики за подобен род международно оценяване. В MT21 участват 6 държави, България (с координатор авторът на този труд), Германия, Корея, Мексико, Тайван и САЩ. Някои от резултатите на MT21 са представени в петнадесета глава.

В началото на 2006 година започна IEA оценяването TEDS-M (буквата M е поставена, за да се подчертае, че се изследва подготовката на учители само по математика). Участват 18 държави. За съжаление България не е сред участващите, поради липсата на интерес от страна на МОМН (и съответно поради липса на национално финансиране). Авторът на този труд е привлечен към Международния координационния център в MSU като консултант, а от юни 2007 до края на TEDS-M през 2010 работи като координатор по математика в този координационен център.

Изследванията MT21 и TEDS-M са широко мащабни международни ОДИ. Те използват основните методики и технологии от изследванията TIMSS, PIRLS и PISA. Основна разлика е, че обект на изследването за MT21 и TEDS-M са *студенти*, обучавани за учители по математика. За разлика от учениците, нужни са други средства за мотивация за участието им в изследването. Един от подходите, използван в MT21 и TEDS-M, е в теста да се задават математически задачи в педагогически контекст. Предположението е, че такива задачите са „по-близки” до бъдещата професия на студентите и така те са „по-интересни” за тях. Примери за такива задачи са могат да се намерят в петнадесета глава.

ВТОРА ЧАСТ
ПЛАНИРАНЕ И ПОДГОТОВКА НА ШИРОКО МАЩАБНИ ОЦЕНЪЧНО-ДИАГНОСТИЧНИ ИЗСЛЕДВАНИЯ

В тази част се разглеждат някои основни методи и технологии, използвани в национални и международни ОДИ. Планирането на всяко ОДИ започва с определяне на методите и технологиите, които ще се използват в него. Цялостната организация на изследването в голяма степен зависи от тях. В този смисъл, разглежданите тук методи и технологии са основополагащи за всяко такова изследване.

Националните и международните ОДИ твърде често се правят с използване на извадка. След това, като се използват подходящи статистически процедури, от статистиките на извадката се получават съответните параметри за цялата популация. Методите за получаване на извадката са важни не само за планирането на тези процедури, но и за оформяне на цялостния дизайн на оценяването. Основните принципи и методи за получаване на извадка за ОДИ са разгледани в трета глава.

Моделирането на ученическите постижения също трябва да се обмисли в началото. Видът на скалите и теоретичната основа, на които те са изградени, сериозно влияят на създаването на рамката и спецификациите на теста. Влиянието им се разпростира и до начина на създаване на тестовите задачи, както и на техния брой. Основни въпроси от скалирането са обект на разглеждане в четвърта глава.

Тестовите задачи са градивните елементи на всеки тест. Колкото „по-хубави” тестови задачи са написани, толкова „по-хубав” тест може да се направи. Опитът на автора в създаване на тестови задачи е систематизиран в пета глава.

Към тази част спадат също методите и технологиите за създаване на рамката и спецификациите на теста, както и за използваните процедури при цялостното му създаване. Тези въпроси се разглеждат в шеста глава.

Седма глава е посветена на изследването TIMSS. Представени са рамката, тестовите спецификации, дизайна и някои от задачите в изследването. Дадени са описанията на класификациите за всяка представена в главата задача.

ТРЕТА ГЛАВА
МЕТОДИ ЗА ФОРМИРАНЕ НА ИЗВАДКА ОТ УЧЕНИЦИ

При национални и международни ОДИ обекти на изследване са големи популации от ученици. Целта на такива оценявания е да се представи обобщена картина за постиженията на учениците за цяла страна, регион на света или целия свят. Това означава, че данните за отделни ученици, паралелки или училища се прави извод за цялата популация. Поради многобройните популации не е нужно да се тества всеки един ученик, което е изключително трудна, обемиста и скъпо струваща дейност. Мащабните ОДИ обикновено използват статистически методи и се работи с така наречената представителна извадка. Тази глава разглежда въпросите за формиране на такава извадка. Използвани са методите за получаване на извадка от международните изследвания TIMSS, PIRLS, PISA и др. Тези методи са международно признати. Тук те са конкретизирани за участие на България в международни изследвания или за нуждите на национални изследвания, т.е. те са съобразени и адаптирани към българската действителност.

1. Особености на извадките за широко мащабни ОДИ

Една от основните задачи на ОДИ е да се направи статистическа оценка на изследваните параметри за съответните популации, като се използват получените статистики на извадките. Формирането на валидна и ефективна извадка за всяка участваща популация е важно условие за качеството на оценяването. Точността на получените резултати зависи от прецизността, с която е планирана и направена извадката. Понякога лошо направени извадки водят до разточително тестване на голям брой ученици или до неефективно използване на училищата, в които вече са избрани ученици за тестване. Това увеличава бюджета на измерването. Другата крайност е, че се тестват недостатъчно на брой ученици или училища, което води до незадоволителна точност на получените резултати.

Популациите за национални или международни ОДИ са големи (например, за TIMSS – всички ученици от VIII клас в България), което налага използване на представителна извадка. Съответните статистически методи са добре развити за случаите, когато се използва така наречената проста случайна из-

вадка. От практически съображения, обаче тя не е удобна в разглежданите случаи. Най-напред, трудно е да се направи пълен списък на учениците от цялата популация, което е изискване за простата случайна извадка. Освен това, учениците са групирани в училища и паралелки. От тази гледна точка е много по-удобно (по-евтино, по-малко трудоемко и т.н.) да се изследват цели паралелки, отколкото да се изследват по 2-3 ученика от различни паралелки и/или училища. Избирането на малък брой ученици от дадено училище не е ефективно, защото подготовката за провеждането на измерването изисква много време и усилия. Те са скъпо струващи и трудоемки дейности и е по-целесъобразно да се събират сведения за повече ученици от училище, което е попаднало в извадката.

На практика се прави така наречената много-стъпкова стратифицирана клъстерна извадка. Това означава следното: *много-стъпкова*, защото става с последователно избиране: най-напред на училища и след това на паралелки в тези училища; *стратифицирана*, защото предлага възможност за разделяне на училищата (съответно учениците) на отделни групи в зависимост от някакъв съществен признак за измерването; *клъстерна*, защото се избират цели клъстери от ученици, в случая – паралелки. По-подробно описание на отделните стъпки е дадено по-долу.

Стъпка 1: Избиране на училища. Това става с метода PPS (Probabilities proportional to their size), което означава, че вероятността дадено училище да попадне в извадката е пропорционална на „големината на училището”, т.е. на броя на учениците от разглежданата популация, които учат в това училище.

Стъпка 2: Избиране на клъстер(и) (паралелка(и)) от избраните вече училища. Това става чрез проста случайна извадка от списъка на всички паралелки в даденото училище. Ето защо във втората стъпка паралелката, а не отделният ученик, е самостоятелна единица за извадката. От избраната паралелка в изследването участват всички ученици.

Ако се направи стратификация (разделяне на училищата в групи според признак, който е важен за измерването) на училищата, трябва да се приложат описаните две стъпки последователно за всяка една от получените страти. В резултат, извадката от училища се състои от училищата във всичките страти, а извадката от ученици (паралелки) е обединението от всички ученици (паралелки) от всичките страти.

Разглежданата тук много-стъпкова стратифицирана клъстерна извадка е практически лесна за изпълнение, осигурява нужната точност и надеждност (ако е направена добре), макар и с по-голям брой ученици. Тя е структурно различна от простата случайна извадка. Това означава, че статистическата теория, която се базира на проста случайна извадка, трябва да бъде „коригирана" по подходящ начин, за да бъде приложена към този вид извадка. Тази „корекция" води до избиране на повече ученици, отколкото предписва простата случайна извадка, както и до използване на специфични методики за анализ на данните. Тези въпроси са разгледани в следващите глави.

2. Коефициент на паралелкова корелация

Педагогическата практика показва, че ученици, които учат в една и съща паралелка, показват по-близки постижения, отколкото, ако същите ученици биха учили в различни паралелки. Обяснението е, че учениците от една паралелка учат и живеят заедно през голяма част от активното за тях време. Хомогенността на паралелките (в казания смисъл) за дадена популация се измерва с *коефициента на паралелкова корелация* ρ. На практика той се получава от данните от предишни оценявания, като се пресметне отношението k на дисперсията на вариационния ред от данните на учениците в паралелките към дисперсията на вариационния ред от данните на всички ученици. Тогава $\rho = 1 - k$. Това означава, че $\rho \in [0; 1]$, като близките до 0 стойности на ρ означават, че популацията се състои от относително хетерогенни паралелки, а близките до 1 – че тя се състои от хомогенни паралелки. Коефициентът ρ има значение за пресмятане на големината на извадката.

Преди да се пресметне големината на извадката, трябва да се определи стандарт за точност на статистическите оценки, получени чрез извадката. Тук приемаме международните стандарти на TIMSS, PIRLS и PISA, а именно 95% доверителен интервал за статистическа оценка на средна стойност, процент и коефициент на корелация, както следва:

- *Средна стойност:* $m \pm 0{,}1s$, където m е статистическата оценка на средната стойност и s е статистическата оценка на стандартното отклонение;

- *Процент:* $p \pm 5\%$, където p е статистическата оценка на процента на ниво ученик;

1. *Корелация:* $r \pm 0{,}1$, където r е статистическата оценка на корелационния коефициент на ниво ученик.

Когато се работи с проста случайна извадка, статистиката ни дава алгоритъм, с който може да се намери точният брой E на елементите от популацията, които трябва да попаднат в извадката, за да се получат зададените стандарти. Това не е така в случаите на много-стъпкова стратифицирана клъстерна извадка. Тук броят на учениците, които трябва да попаднат в извадката, съществено зависи от коефициента ρ.

Да предположим, че от всяко училище избираме клъстер, състоящ се от n ученика. Ако коефициентът на паралелкова корелация за популацията е ρ, то за да получим дадени стандарти трябва да имаме извадка от $T = \left(1 + \left(n-1\right)\rho\right)E$ на брой ученика, където E е броят на учениците за същите стандарти при проста случайна извадка (TIMSS 2003 School Sampling Manual, 2001). Ясно е, че колкото по-хомогенни са паралелките на генералната съвкупност, толкова повече ученици трябва да съдържа извадката.

За формулата по-горе предположихме, че от всяко училище избираме клъстер с еднакъв брой ученици. На практика това предположение не се реализира, понеже клъстерите са паралелки, а паралелките са различно големи. За практически цели числото n се избира да е равно на средния брой ученици в паралелка за популацията. След това от всяко училище се избира по една паралелка.

Описаният модел е операционализиран по-долу за получаване на извадка при специфичните за България условия.

3. Генерална съвкупност (популация) за измерването

В рамките на тази глава изразите "генерална съвкупност" и "популация" означават едно и също нещо. Множеството от всички ученици, които са обект на дадено изследване (измерване), се нарича *желателна генерална съвкупност (желателна популация)*. Това множество трябва да бъде строго дефинирано в самото начало на изследването.

Например: Всички ученици от 8 клас в България за времето на изследването (например, дадена учебна година).

Очаква се, че всички ученици от желателна генерална съвкупност трябва да бъдат в рамката на извадката. Понякога обаче, по една или друга причина (спецификата на някои видове училища, отдалечени и трудно достъпни райони и др.) се налага редуциране на тази съвкупност. Така се достига до *ефективна генерална съвкупност*, която или съвпада, или е малко по-малка от желателната генерална съвкупност. Ефективната генерална съвкупност е рамката на извадката. Международно приетите норми (например тези на IEA изследванията) са, че за получаването на ефективната генерална съвкупност е допустимо да се редуцират до 5% от учениците в желателна генерална съвкупност.

В Българските условия ефективната генерална съвкупност е удачно да се получи като от желателната генерална съвкупност се извадят учениците в училищата за умствено изостанали деца, тези с криминално поведение и много малки училища от статистическа гледна точка. За да дефинираме "много малките училища" изхождаме от конкретните данни за популацията. Например за изследването PIRLS това са училищата, в които има по-малко от половината от средния брой на ученици в паралелка за популацията. За TIMSS, това са училищата, в които има по-малко от една четвърт от средния брой на ученици в паралелка за популацията. Не трябва да се забравя стандартът, че редуцираните ученици не трябва да надхвърлят 5% от учениците в желателната популация.

4. Определяне на големината (обема) на извадката

Ако работим с проста случайна извадка, приетият стандарт за точност на статистическите оценки се постига при обем на извадката от около 400 ученика.

За целите на двустъпкова стратифицирана клъстерна извадка, клъстерите са паралелки, т.е. от всяко избрано училище избираме по една паралелка. Този начин за получаване на извадка е използван в международните изследвания TIMSS, PIRLS и PISA, което означава, че той е международно признат за условията на България.

За получаване на големината на извадката използваме формулата $T = \left(1 + \left(n - 1\right)\rho\right)E$, в която $E = 400$, ρ е коефициентът на паралелкова корелация и n е големината на клъстерите. За практически цели числото n се избира равно

на средния брой ученици в паралелка за популацията. Това число се пресмята за всяка популация отделно, когато са налице всички данни за училищата, паралелките и учениците. За коефициента на паралелкова корелация разполагаме със следните данни:

1. От изследването TIMSS 1999, за математика VIII клас $\rho = 0,5228$.
2. От изследването PIRLS 2001, за четене, IV клас $\rho = 0,3447$.

Приетите международни стандарти казват, че при липса на данни може да се приеме, че стойностите на ρ са в интервала [0,3; 0,4].

В широко мащабни изследвания, за които става дума тук, се използват вероятностните модели за измерване постиженията на учениците (виж четвърта глава). При тях всеки ученик работи само по част от тестовите задачи. Използването на тези модели изисква големината на извадката да се увеличи още с около 13%. По-точно, формулата се коригира по следния начин $T = \left(1 + \left(n-1\right)\rho\right)E + \frac{1,25E}{n}$, от където лесно се получава точната стойност на *T*.

Казаното до тук дава възможност да определим броя *T* на всички ученици в извадката. За първата стъпка на извадката обаче, е нужно да се определи броят на паралелките (съответно училищата), който е $c = \frac{T}{n}$.

Ето един примерен план за извадка от популацията „Всички ученици от VIII клас в България" при изследване на постижения по математика. По данни от 2002 г. средният брой ученици в паралелка от VIII клас за страната е 19. Приемаме $E = 400$ и $\rho = 0,5$. Така получаваме $T = 4026$ ученици. Броят на паралелките, които трябва да попаднат в извадката, е приблизително 212. Един начин да направим извадката е да подберем по описания по-долу начин 212 училища и във всяко от тях да изберем по една паралелка. Ако по технически причини считаме, че броят на училищата е много голям, може да подберем по-малко училища, например 152; в най-големите 40 от тях да изберем по две паралелки, а в останалите – по една паралелка. По този начин беше осъществена националната извадка за TIMSS-2003.

5. Стратификация

Преди да бъдат избрани, училищата могат да бъдат стратифицирани. Това означава те да се групират по определен признак (вид училище, брой ученици от популацията в училище, различните програми на училищата, урбанизацията на региона и др. признаци, които са важни за изследването) като всяко училище принадлежи на точно една от тези групи. Стратификацията се прави, за да се повиши надеждността на статистическите оценки от изследването, да се приложи различен дизайн на извадката за различните страти, да се осигури пропорционално представяне на определени групи от популацията. За целите на извадката стратификацията е два вида:

Пряка стратификация: Тя се прави с цел да се приложи отделен дизайн за различните страти. Това означава, че за всяка страта се прави отделен списък на училищата и е възможно училищата да се избират по различен начин за различните страти.

За изследването TIMSS, например, е направена пряка стратификация с две страти: страта 1 („големи" училища) – училища с брой ученици от популацията не по-малък от средния брой ученици в паралелката и страта 2 („малки" училища) – всички останали училища. В страта 1 училищата се избират по метода PPS, а в страта 2 – по проста случайна извадка. (Причината е, че училищата в страта 2 не се различават съществено по броя ученици от популацията).

Да отбележим още, че когато се ползва пряка стратификация, стандартната грешка на цялата извадка за величината *y* се пресмята по формулата

$$se(y) = \frac{\sqrt{N_1^2 \operatorname{var}(y_1) + \ldots + N_k^2 \operatorname{var}(y_k)}}{N},$$

където $se(y)$ е стандартната грешка на цялата извадка за *y*, $\operatorname{var}(y_i)$ е дисперсията на *y* за стратата с номер *i*, N_i е броят на учениците в стратата с номер *i*, *N* е броят на учениците в цялата популация.

Непряка стратификация: Тя се прави, за да се осигури пропорционално представяне на определени групи от популацията. За нея не е нужно да има отделен списък от училищата. В общия списък училищата се сортират по страти и се определя съответният процент за избор на училища във всяка страта. Да отбележим, че в общия списък първата непряка страта се сортира по намаляващ

ред на "големина на училището", следващата страта – по растящ ред на същата променлива, след това – пак в намаляващ, после пак растящ и т.н.

Определянето на непреките страти става въз основа на признак, който е съществен за измерването. Например, в изследването TIMSS е направена непряка стратификация на учениците от не профилираните от профилираните по математика училища, тъй като тези две групи ученици учат по различен учебен план по математика, а това е съществен параметър за изследването.

6. Необходими начални данни за получаване на извадката

За да се направи извадката на училища е необходимо да разполагаме с данни за *всички* училища, в които има ученици от интересуващата ни популация. Тези данни са името на училището, адрес за кореспонденция, телефони (електронна поща) за връзка, брой паралелки от интересуващата ни популация и брой ученици във всяка от тези паралелки. За прегледност в таблица 2 е даден пример за събиране на данните от всяко училище.

Таблица 2. Необходими данни за училищата за получаване на извадка.

Име на училището	**Адрес за кореспонденция**	**Телефон (e-mail)**	**Паралелки VIII клас**
СОУ "Ив. Вазов"	ул."Асен Златаров" 35, 4000 Пловдив	(032) 45-67-23	VIII а – 27 уч. VIII б – 25 уч. **Общо: 52 уч.**

Получените данни се въвеждат в електронна таблица (Excel) и училищата се сортират в намаляваща редица по "големина на училището". Така лесно може да се пресметне средният брой ученици в паралелка за популация за страната, както и да се направят съответните изключвания на училища, за да получим училищата от ефективната генерална съвкупност.

7. Технология за избиране на училища

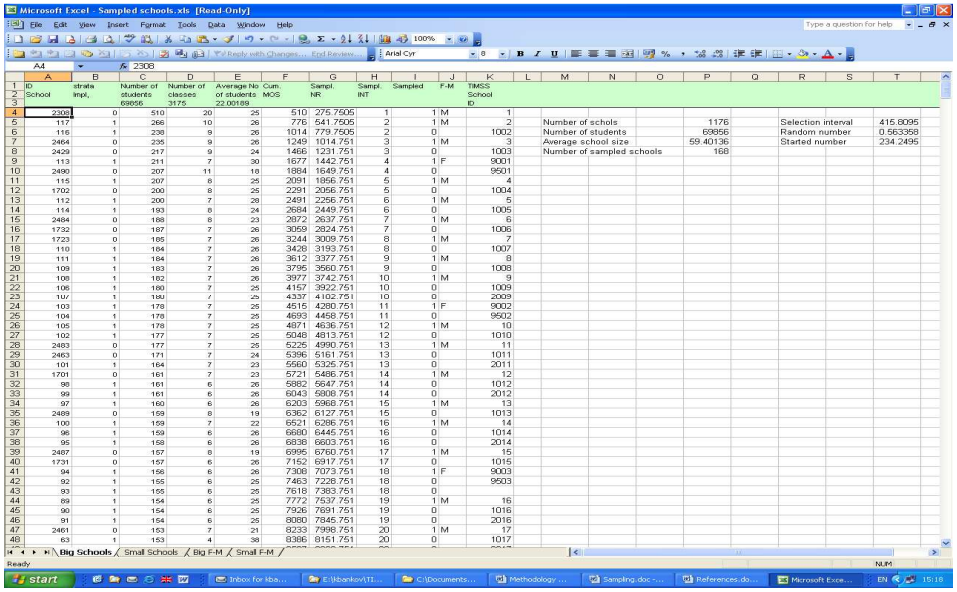
Ще опишем технология, осъществяваща PPS метод на извадка (вероятността дадено училище да попадне в извадката е пропорционална на "големината на това училище") за една пряка страта (Банков, 1997, 1998-А). Ако има повече преки страти, технологията се повтаря за всяка от тях.

Предполагаме, че всички непреки страти са сортирани, както е казано по-горе (края на параграфа "Стратификация"). Освен това те са номерирани последователно с числата 1, 2, 3 ...., m.

1. Пресмятане на извадков интервал I. Той е равен на сбора от "големина на училището" (брой ученици в стратата в това училище) за всички училища в стратата, разделен на броя на училищата, които трябва да се изберат от тази страта.
2. Избиране на случайно число R. Това е число в интервала (0; 1), което лесно може да се избере с компютърен софтуер за случайни числа.
3. Пресмятане на изборни числа. Тези числа $S_1, S_2, ..., S_m$ се пресмятат за всяко училище отделно по следния рекурентен начин: $S_1 = R.I$, $S_p = S_{p-1} + I$ за всяко $p = 2,3,...,m$. Те могат да се получат и по формулата $S_i = R.I + (i-1)I$ за всяко $i = 1,2,...,m$.
4. Пресмятане на изборни индекси. Тези числа $T_1, T_2, ..., T_m$ се пресмятат за всяко училище по формулата $T_i = \left[\frac{S_i}{I}\right] + 1$, $i = 1,2,...,m$, където средните скоби означават цяла част на число.
5. Определяне на избраните училища. За всяко училище пресмятаме символите $V_1, V_2, ..., V_m$ по следния начин: $V_1 = \begin{cases} *, \text{ ако } T_i = 1 \\ -, \text{ ако } T_i \neq 1 \end{cases}$ и $V_p = \begin{cases} *, \text{ ако } T_p - T_{p-1} = 1 \\ -, \text{ ако } T_p - T_{p-1} \neq 1 \end{cases}$ за всяко $p = 2,3,...,m$. Училищата от извадката са онези училища, за които съответният символ V е *.

За всяко избрано училище, училището, което е след него в списъка (със следващ номер), се счита, че е негов заместник. Ако първоначално избраното училище по една или друга причина откаже участие в изследването, привлича се училището-заместник. В някои изследвания (например TIMSS, PIRLS) всяко училище има по два заместника.

По-долу е даден част от електронен файл, който реализира описаната технология. Файлът осъществява извадка на български училища за изследването TIMSS 2003. Едновременно се избират училища за пилотното тестване (F) и за основното тестване (M).



ID School	strata Impl.	Number of students 69856	Number of classes 3175	Average No of students 22.00189	Cum. MOS	Sampl. NR	Sampl. INT	Sampled	F-M	TIMSS School ID
2308	0	510	20	25	510	275.7505	1	1	M	1
117	1	266	10	26	776	541.7505	2	1	M	2
116	1	238	9	26	1014	779.7505	2	0		1002
2464	0	235	9	26	1249	1014.751	3	1	M	3
2429	0	217	9	24	1466	1231.751	3	0		1003
113	1	211	7	30	1677	1442.751	4	1	F	9001
2490	0	207	11	18	1884	1649.751	4	0		9501
115	1	207	8	25	2091	1856.751	5	1	M	4
1702	0	200	8	25	2291	2056.751	5	0		1004
112	1	200	7	28	2491	2256.751	6	1	M	5
114	1	193	8	24	2684	2449.751	6	0		1005
2484	0	188	8	23	2872	2637.751	7	1	M	6
1732	0	187	7	26	3059	2824.751	7	0		1006
1723	0	185	7	26	3244	3009.751	8	1	M	7
110	1	184	7	26	3428	3193.751	8	0		1007
111	1	184	7	26	3612	3377.751	9	1	M	8
109	1	183	7	26	3795	3560.751	9	0		1008
108	1	182	7	26	3977	3742.751	10	1	M	9
106	1	180	7	25	4157	3922.751	10	0		1009
107	1	180	7	25	4337	4102.751	10	0		2009
103	1	178	7	25	4515	4280.751	11	1	F	9002
104	1	178	7	25	4693	4458.751	11	0		9502
105	1	178	7	25	4871	4636.751	12	1	M	10
102	1	177	7	25	5048	4813.751	12	0		1010
2463	0	177	7	25	5225	4990.751	13	1	M	11
2463	0	171	7	24	5396	5161.751	13	0		1011
101	1	164	7	23	5560	5325.751	13	0		2011
1701	0	161	7	23	5721	5486.751	14	1	M	12
98	1	161	6	26	5882	5647.751	14	0		1012
99	1	161	6	26	6043	5808.751	14	0		2012
97	1	160	6	26	6203	5968.751	15	1	M	13
2489	0	159	8	19	6362	6127.751	15	0		1013
100	1	159	7	22	6521	6286.751	16	1	M	14
96	1	159	6	26	6680	6445.751	16	0		1014
95	1	158	6	26	6838	6603.751	16	0		2014
2487	0	157	8	19	6995	6760.751	17	1	M	15
1731	0	157	6	26	7152	6917.751	17	0		1015
94	1	156	6	26	7308	7073.751	18	1	F	9003
92	1	155	6	25	7463	7228.751	18	0		9503
93	1	155	6	25	7618	7383.751	18	0		
89	1	154	6	25	7772	7537.751	19	1	M	16
90	1	154	6	25	7926	7691.751	19	0		1016
91	1	154	6	25	8080	7845.751	19	0		2016
2461	0	153	7	21	8233	7998.751	20	1	M	17
63	1	153	4	38	8386	8151.751	20	0		1017

Number of schols	1176
Number of students	69856
Average school size	59.40136
Number of sampled schools	168

Selection interval	415.8095	
Random number	0.563358	
Started number	234.2495	

8. Технология за избиране на паралелки

От всяко от избраните училища избираме по една паралелка по метода на проста случайна извадка. Този метод лесно се реализира на компютър. Една негова реализация може да получим като приложим описаната технология за избор на училища, като “големина на училището” на всякъде има една и съща стойност (различна от 0), например 1. Разбира се, това се прави за всяко училище поотделно, като се направи списък на всички паралелки в съответното училище.

ЧЕТВЪРТА ГЛАВА
СКАЛИРАНЕ ЧРЕЗ ВЕРОЯТНОСТНО МОДЕЛИРАНЕ

В педагогическите изследвания често се налага да се измерват абстрактни величини като: математически способности, социално икономически статус, отношение към математиката и др. Такива величини не могат да се измерват директно (например, както се измерва дължина на пръчка). Те се наричат ***конструкти*** и за измерването им се въвеждат латентни (скрити) променливи. На практика, измерването се извършва с тестове или въпросници (анкети), чрез които се наблюдават отговорите на учениците. Така разполагаме с явни променливи (ученическите отговори), за които „вярваме", че зависят от „стойностите" на някои от латентните променливи. Например, вярваме, че ученик с положително отношение към математиката би отговорил на въпроса „Важен предмет ли е математиката в училище?" по различен начин от ученик с отрицателно отношение към нея.

Тъй като никога не може да получим действителен бал за латентните променливи, опитваме се да ги изучаваме като обединяваме по подходящ начин баловете от наблюдаемите променливи, които считаме, че представят съответния конструкт. Терминът ***скалиране*** се използва за описание на начина, по който обединяваме баловете от множество тестови задачи или въпроси от анкета, за да получим един общ бал.

Съществуват различни начини за скалиране. Преди всичко, обаче трябва да определим кои от променливите, с които работим, са дискретни и кои са непрекъснати. Например, пола на ученика, отговора на дадена задача (кодиран с „вярно-невярно") са примери за дискретни променливи. От друга страна математическите способности, социално икономическия статус обикновено се приемат за непрекъснати променливи.

Когато латентната променлива е непрекъсната и наблюдаемите променливи също са непрекъснати, най-често използваният метод за скалиране е факторният анализ. Той се използва, за да се определи броят на латентните променливи, които се описват с множеството от наблюдавани балове, както и за да се построи скала за наблюдаваните отговори.

Когато латентната променлива е непрекъсната и наблюдаемите променливи са дискретни, се използват методите на вероятностното моделиране. Основната идея е, че всеки ученик има някакво (неизвестно) „място" в скалата на

латентната променлива, което искаме да установим. В зависимост от това положение той дава определен набор от отговори (стойности) за дискретните наблюдаеми променливи. По този известен набор може да се установи (с някаква точност) мястото на ученика в скалата на латентната променлива. Има различни вероятностни модели, които по различни методи определят мястото на ученика върху скалата на латентната променлива.

Когато латентната променлива е дискретна и наблюдаемите променливи непрекъснати, се използва клъстърен анализ. С него учениците се класифицират в малък брой клъстери въз основа на стойностите (отговорите) си върху множество от непрекъснати променливи.

Когато латентната променлива е дискретна и наблюдаемите променливи също са дискретни, се използва латентен класифициращ анализ. Подобно на вероятностното моделиране, класификацията на учениците става въз основа на цялостния набор от балове на непрекъснатите наблюдавани променливи.

На практика голяма част от учените работят със сравнително прости методи за скалиране. Например, ако имаме 6 въпроса в 4-степенна скала с възможни отговори 0, 1, 2 и 3, може или да работим с общия брой точки (вариращ от 0 до 18), или да ползваме средната стойност от отговорите (вариираща от 0 до 3). Както единият, така и другият метод дават лесно разбираема скала, с която може да се работи, поне в началния етап от анализа на данни. Ето как тази ситуация може да се моделира и по други начини с подходящо скалиране.

Идеята на Мостелър и Тъкий (Mosteller, F., & Tukey, J. W., 1977) е да се предположи, че латентната променлива има така нареченото „логистично" разпределение (то е подобно на нормалното разпределение, но с „по-дебели" опашки). След това да се разгледат наблюдаваните балове (например тези от 0 до 18) като ординални категории. Да се прекодират тези категории в нова скала като се вземат пред вид съответните им честоти в логистичното разпределение. Наредбата на баловете в новата скала се запазва, но се променят разстоянията между тях, в зависимост от първоначалното им разпределение. Има и други варианти на този метод, които се различават по предполагаемото разпределение на латентната променлива.

Друга възможност е да се направи факторен анализ на наблюдаваните данни. Трябва да се предположи обаче, че скалата, в която са измерени отговорите (0, 1, 2 и 3), е непрекъсната. Разбира се, това е много силно предположе-

ние, но въпреки това чрез факторния анализ ще получим информация дали измерваме един или повече конструкта. Най-простият факторен анализ (анализ на основните компоненти) дава „факторен бал”, който може да се използва като скала за представяне на латентната променлива. Такъв метод е използван в изследването PISA, за да се измерва конструкта „социално икономически статус”.

Трета възможност е да се използват методите на вероятностното моделиране към ситуацията на 6 въпроса с по 4 отговора 0, 1, 2 и 3. Това е дълга процедура, която изисква доста време и специализиран софтуер. Предимството е, че получените скали са в някакъв смисъл „универсални”. Към тях лесно се прилагат предположения за различна форма на началното разпределение на латентната променлива (виж десета глава). Практиката показва, че скалите, получени с вероятностното моделиране са с висока степен на корелация с всяка друга смислена скала. Затова по-долу са разгледани методите на вероятностното моделиране.

1. Измерване на ученическите постижения

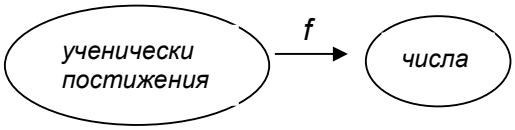
Измерването е една от най-често извършваните дейности. Ежедневно срещаме измерване на дължини, тегла, време, температура и др. *Измерването* процес, при който на всеки един от дадено множество обекти или явления се съпоставя число по дадено правило. Всъщност, измерването се прилага към определени свойства на обектите или явленията, а не към самите тях. Например, като измерваме даден човек, трябва да уточним от кое негово свойство се интересуваме: ръст, тегло, гръдна обиколка и др.

Свойствата, за които стана дума в предишния параграф, са *физически* свойства. Тяхното измерване става директно с помощта на създадени измерителни инструменти (уреди), като линийка (разграфена на сантиметри), кантар, часовник, термометър и др. Обект на измерване в даден човек може да са свойства (качества), които не са физически, т.е. не могат да се измерят директно с уред. Например, възможностите му за критично мислене, уменията му за използване на даден чужд език, знанието му в определена област и др. Това са психологически свойства (качества), които се наричат *конструкти*.

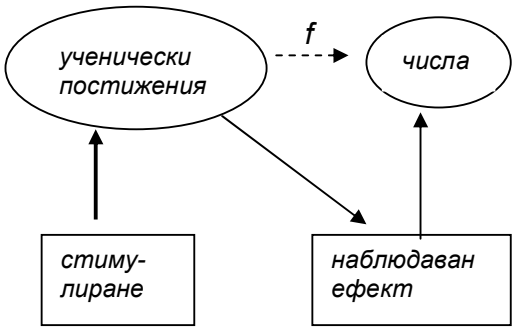
Конструкт е психологическо свойство (качество) на човека, което не може да бъде директно измерено с инструмент (уред). То е хипотетично поня-

тие, плод на мисълта на учените, които го използват, за да обяснят някои страни от поведението на човека. Степента (големината) на определен конструкт за даден човек може да бъда определена (измерена) само като се наблюдава поведението на човека. Следователно, за да се измери конструкт, трябва да се стимулира проявлението му и да се наблюдава полученото поведение (резултатите).

Основен конструкт в тази глава е *ученически постижения*. Така условно наричаме знанията и уменията на даден ученик (или множество ученици) в дадена област. За да се измерят ученическите постижения на дадено множество ученици, трябва на всеки ученик от множеството да се съпостави някакво число, което количествено описва *големината* на неговите постижения. С други думи, трябва да се построи числова функция f (правило), която на постиженията на всеки ученик да съпоставя число (фиг. 1). Като при измерването на всеки конструкт, построяването на такава функция става индиректно. За целта се стимулира проява на ученическите постижения и се измерва наблюдавания ефект при тази проява (фиг. 2). Инструментът, с който става стимулирането, се нарича *тест*, а процесът на стимулиране – *тестиране*. Наблюдаваният ефект е писмената работа на ученика, на която по дадени правила се съпоставят числа във вид на точки или оценка.



Фиг. 1. Измерване на ученици



Фиг. 2. Измерване на конструкт

Така в процеса на измерване на ученическите постижения се оформят два типа обекти: тези, които измерваме (ученици), и тези, с които правим измерването (тестови задачи). Начините, по които могат да се свържат тези два типа обекти, както и редица проблеми, свързани с описаната схема на измерване, са обект на изучаване от *Теорията на тестовете* (Crocker, L. and J. Algina, 1986). Тя е множество математически модели, които предлагат различни решения на проблемите. Исторически първият систематично изграден модел е така нарече-

ната *Класическа теория на тестовете* (Стоименова, 2000). Освен че е в основата на почти всички останали модели, тя се използва и самостоятелно. (Например при националния изпит-тест за прием след завършен 7 клас през юни 2000 и 2001 г. виж Банков, Витанов, 2001.) Тук няма да се спираме на теоретичните основи на класическата теория на тестовете, тъй като те са добре разработени и известни. Класическата теория, обаче, при някои условия на измерване проявява недостатъци. Част от тях се преодоляват с използването на така нареченото *вероятностно моделиране* (Банков, 2002-Б) в теорията на тестовете.

2. Случаи, в които използването на класическата теория не е удобно.

При провеждане на широко мащабни ОДИ основните проблеми са свързани с двата типа обекти. Първо, генералната съвкупност (обектите на измерване) е много голяма. (Например, в TIMSS участват около 50 държави във всяка от които популацията е „всички ученици от осми клас".) Този проблем се решава с подходяща методика за правене на извадка и съответните статистически анализи от нея (вж. трета глава). Второ, тематиката на изследването е обширна, което налага използването на твърде много тестови задачи за стимулиране на ученическите постижения. Необходимите тестови задачи са толкова много, че в рамките на разумното време не е по силите на един ученик да работи по всички тях. Това е сериозен измерителен проблем, тъй като според класическата теория на тестовете измерването на даден ученик става само въз основа на задачите, върху които той е работил.

Един възможен начин да се преодолее проблемът е задачите да се организират в блокове и различни тестови книжки, като има застъпване на блоковете в книжките. Ето един прост пример. Нека всички задачи от теста са организирани в пет блока *A*, *B*, *C*, *D*, и *E*. От тези блокове са съставени пет различни тестови книжки 1, 2, 3, 4 и 5, като книжка 1 се състои от блоковете *A* и *B*, книжка 2 – от блоковете *B* и *C* и книжка 3 – от блоковете *C* и *D*, книжка 4 – от *D* и *E* и книжка 5 – от *E* и *A* (виж Таблица 3).

Всеки от учениците работи само по една тестова книжка, а не по всички задачи от теста, като броят на учениците, работещи по всяка книжка, е приблизително еднакъв. Така имаме връзка между всеки два блока. (Например връзката между блоковете *A* и *D* се осъществява чрез книжки 4 и 5.) Но всеки ученик има резултати само по книжката, по която е работил. Ако сега приложим класи-

ческата теория, ще получим 5 различни криви (или хистограми) на разпределение на резултатите, съответстващи на групите ученици, работещи върху всяка от петте тестови книжки. Нямаме възможност да получим единна крива на разпределение на резултатите на учениците по всички задачи от теста.

Таблица 3. Примерно разпределение на блокове по книжки.

Номер на книжка	**Блокове в книжката**
1	A, B
2	B, C
3	C, D
4	D, E
5	E, A

Ето още едно ограничение, което идва от използване на класическата теория на тестовете. В тази теория постиженията на отделни групи ученици (например отделните държави) се сравняват лесно като се даде информация за средния процент ученици от всяка група, решили вярно задачите от теста. Например, нека имаме тест с три задачи и процентът вярно отговорили ученици по отделните задачи, както и средното за три държави върху него, са представени в таблица 4.

Таблица 4. Примерно представяне на три държави в тест от три задачи.

	Държава А	Държава Б	Държава В
Задача 1	25%	100%	50%
Задача 2	50%	50%	50%
Задача 3	75%	0%	50%
Средно	**50%**	**50%**	**50%**

Въз основа на тези данни правим заключение, че трите държави са се представили еднакво върху теста. Както се вижда, обаче, загубили сме информация за относителната трудност на задачите за отделните държави. Като се има пред вид съдържателната страна за всяка задача, загубената информация е важна за качественото описание на различията в отделните държави.

Ограниченията на класическата теория в този случай могат да бъдат преодолени, ако се намери скала за оценяване (различна от скалите, които ни предлага класическата теория), която има следните свойства:

- По решенията на *част от задачите* да може да се определи мястото на всеки ученик в скалата, както и разпределението на цялата генерална съвкупност по *задачите на целия тест*.
- Да може да се представят в една и съща скала (мерна единица) както тестовите задачи, така и постиженията на учениците.

3. Вероятностно моделиране.

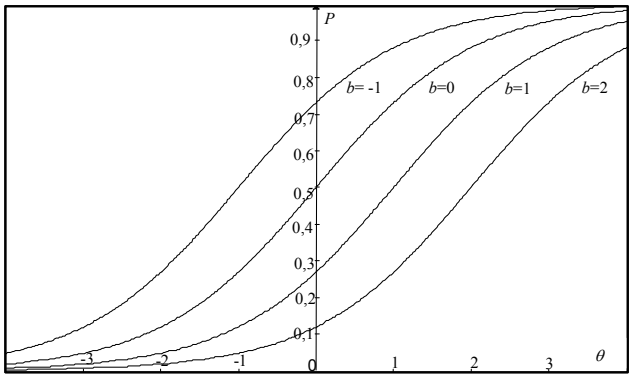
В основата на вероятностното моделиране (Item Response Theory – IRT) стои идеята, че възможността произволно избран ученик да реши дадена тестова задача може да бъде предсказана (или обяснена) чрез изменението на една абстрактна величина, условно наречена *способност*. Математически това се описва с една растяща функция, изразяваща вероятността произволно избран ученик да реши правилно дадена тестова задача в зависимост от *способността*. Тази функция се нарича *характеристична* за конкретната задача (или *характеристична крива*). Различните задачи имат различни характеристични функции. В най-разпространените модели тези разлики се описват чрез един, два, или три параметъра. Съответните модели се наричат едно- дву- и три- параметрични (Hambelton, R., H. Swaminathan, H. Rogers, 1991).

За да опишем по-подробно вероятностните модели, нека най-напред разгледаме множество от n задачи с избираем отговор. Отговорът на всеки ученик на всяка такава задача се кодира или с 1, ако ученикът е посочил правилния отговор (решил е задачата), или с 0, в противен случай. При едно-параметричния модел характеристичната функция на която и да е задача се задава с формулата

$$P_i(\theta) = \frac{e^{\theta - b_i}}{1 + e^{\theta - b_i}} \quad , \qquad i = 1,2,...n \, ,$$

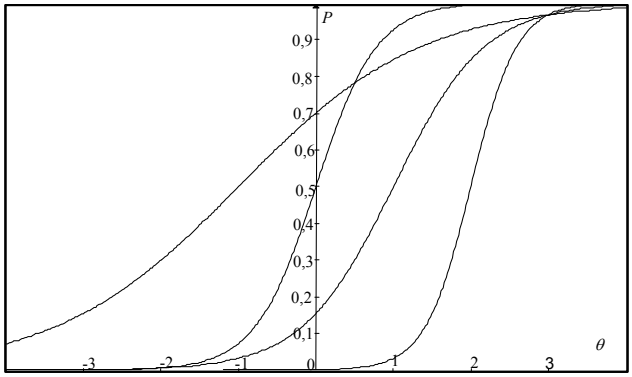
където i е номерът на задачата, $P_i(\theta)$ е вероятността произволно избран ученик да посочи правилния отговор на i-тата задача (да реши задачата) в зависимост от *способността* θ, а b_i е параметър, наречен *трудност*, на i-тата задача. Параметрите b_i за различните задачи може да са различни.

Параметърът b_i е онази стойност на *способността*, при която вероятността да се посочи правилният отговор на *i*-тата задача е 0,5. Следователно параметърът b_i определя положението на характеристичната крива по отношение на скалата *способност*. Колкото по-голяма е стойността на b_i, толкова по-голяма *способност* е нужна за да се получи 50% вероятност за посочване на правилния отговор. Затова, задачите с по-голям параметър b_i са в определен смисъл по-трудни от задачите с по-малък b_i.



Фиг. 3. Характеристични криви – 1 параметър

Скалата *способност* обикновено се "оразмерява" така, че средната стойност на b_i за всички задачи да е 0 и тяхната стандартна грешка да е 1. Тогава „смислените" стойности на b_i обикновено са между –2 и 2. Това означава, че стойност на b_i близка до (–2) съответства на много лесна задача, а стойност на b_i близка до 2 съответства на много трудна задача.



Фиг. 4. Характеристични криви –2 параметъра

На фиг. 3 са представени четири характеристични криви с параметър на трудност съответно –1, 0, 1 и 2.

При дву-параметричния модел характеристичната функция на която и да е задача е

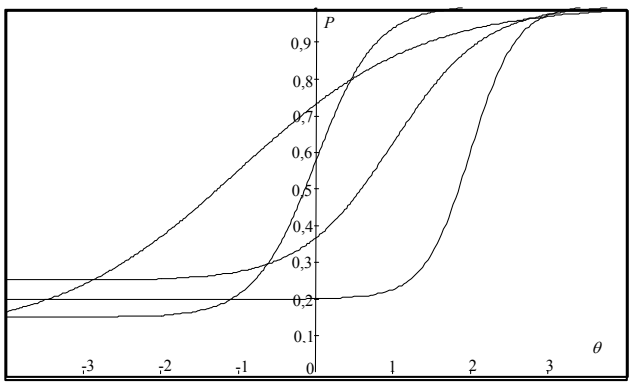
$$P_i(\theta) = \frac{e^{1,7a_i(\theta - b_i)}}{1 + e^{1,7a_i(\theta - b_i)}}\ ,\ i = 1,2,...n\ .$$

Освен известните от едно-параметричния модел означения, тук се появява втори параметър a_i. Той се нарича *дискриминация* на задачата. Този параметър е про-

порционален на наклона на допирателната в точката $\theta = b_i$. Задачи с по-голяма стойност на a_i имат „по-стръмна” графика и следователно по-добре разграничават постиженията на учениците в област от *способности* близки до b_i. Теоретично стойностите на a_i са всички реални числа, но на практика най-често този параметър е в интервала (0, 2).

На фиг. 4 са представени четири дву-параметрични характеристични криви с параметри $(b;a)$ съответно (-1; 0,5), (0; 1,5), (1;1) и (2;2).

Три-параметричният модел добавя още един параметър c_i, наречен *налучкване*. Той взема предвид възможността да се налучква правилният отговор. Графично този параметър променя височината на долната асимптота на характеристичната крива, която е равна на параметъра на налучкване. Характеристичното уравнение на която и да е задача в три-параметричния модел има вида



Фиг. 5. Характеристични криви –3 параметъра

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{1,7a_i(\theta - b_i)}}{1 + e^{1,7a_i(\theta - b_i)}}, \qquad i = 1,2,...n.$$

На фиг. 5 са представени кривите от фиг. 4 с добавен параметър *налучкване* съответно равен на 0,1; 0,15; 0,25 и 0,2.

Вероятностното моделиране за измерване на ученическите постижения е разработено през втората половина на 20-ти век. Като основополагащ труд се счита книгата на датския математик Раш (Rasch, 1960). В нея подробно се разглежда едно-параметричния модел. Поради това този модел често се нарича *модел на Раш*. За техническо удобство, по-нататък ще разглеждаме само модела на Раш, въпреки че всичко, за което ще става дума, е приложимо и в трите модела.

4. “Универсалност” на скалата *способност*.

Когато по-горе описахме някои ограничения на класическата теория на тестовете, се постави въпросът за търсене на скала с две изисквания. Сега ще обясним защо те са на лице при вероятностното моделиране. Ще изпреварим

малко събитията, като предположим, че знаем как да оценим мястото на всеки ученик в скалата *способност*. Това ще бъде направено по-нататък (вж. т. 6). Поточно, ще бъде обяснено как по резултатите на даден ученик върху едно множество от тестови задачи, на този ученик се съпоставя точно определено число от скалата *способност*. Да наречем това число *постижение* на ученика.

Да предположим, че разполагаме с резултатите на множество ученици върху група тестови задачи. С помощта на математико-статистически методи от тези резултати може да се направи оценка на параметрите на всяка една задача. Така разполагаме с характеристичните функции (криви) за всички тестови задачи. Да обърнем внимание на факта, че за оценката на тези параметри не е задължително задачите да бъдат решавани от всички ученици. Отделни групи ученици може да са решавали различни множества задачи. При положение, че знаем числото *постижение* (в скалата *способност*) на всеки ученик, лесно е да пресметнем каква е вероятността този ученик да даде правилен отговор на която и да е от тестовите задачи, дори и за онези, които ученикът не е решавал. Например, ако имаме задача с трудност 0 (b_i=0), то трима ученици с постижения θ_1= 1, θ_2 = 0 и θ_3 = –1 имат вероятност да дадат правилен отговор на задачата съответно

```
 3              |
                |
              x |

              x |
             xx |
 2           xx | 9 22
            xxx |
            xxx | 6 16
           xxxx | 8 11 27 29
 1        xxxxx |
        xxxxxxx | 31
        xxxxxxx | 2 30
      xxxxxxxxx | 13
     xxxxxxxxxx | 19
 0      xxxxxxx | 5 32
       xxxxxxxx | 7 15 28
        xxxxxxx | 4 14 21
       xxxxxxxx | 3 17 20 23
      xxxxxxxxx | 10 18 24
-1       xxxxxx |
           xxxx | 1
           xxxx |
             xx | 12 26
-2          xxx | 25
             xx |
              x |
              x |
              x |
-3            x |
```

Фиг. 6. Единна скала за ученици и задачи

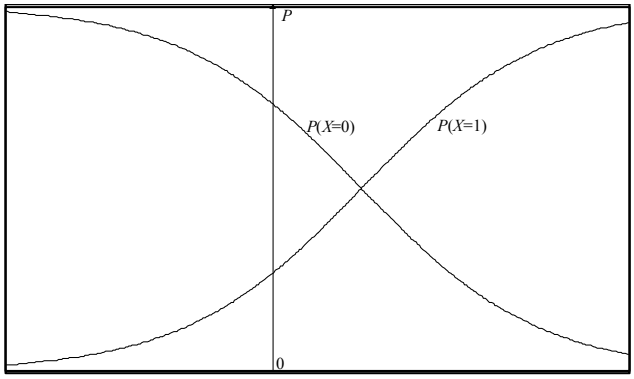
$P(1) = \dfrac{e}{1+e} \approx 0{,}73$, $P(0) = \dfrac{1}{1+1} = 0{,}5$ и $P(1) = \dfrac{e^{-1}}{1+e^{-1}} \approx 0{,}27$ (разглеждаме едно-параметричния модел).

Скалата *способност* служи както за измерване на учениците чрез числото *постижение* за всеки от тях, така и за измерване на задачите чрез параме-

търа трудност за всяка от тях. На фиг. 6 е дадено общо представяне (получено чрез компютърна програма) на учениците и задачите по един тест в скалата *способност*. Скалата е начертана по средата и е насочена нагоре. От ляво са наредени учениците съгласно техните постижения (всяко X е един ученик). От дясно са наредени задачите по номера съгласно тяхната трудност. Учениците, които са "на една височина" с дадена задача, имат 50% вероятност да решат правилно тази задача. Онези, които са „над" дадена задача, имат вероятност повече от 50% да дадат правилния отговор, а тези, които са „под" дадена задача, имат вероятност по-малка от 50% да я решат правилно. Точната им вероятност може да се пресметне по съответната формула.

5. Обобщения на вероятностните модели.

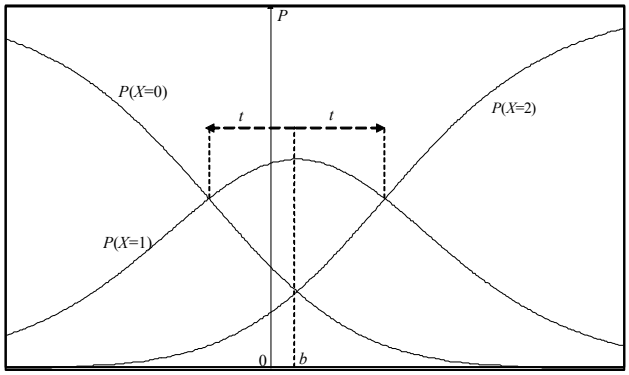
В литературата са описани различни обобщения на разгледаните вероятностни модели, правени за по-общи или по-конкретни цели. Тук ще обясним как вероятностното моделиране е приложимо не само за задачи с избираем отговор, но за други формати на задачите, включително и задачи със свободен отговор.



Фиг. 7. Две характеристични криви на задача

До сега разглеждахме задачи, които се кодират с две числа 1 и 0. Това означава следното. Ако X_i е кодът на отговора на даден ученик на i-тата задача, то X_i може да приема стойности 1 при правилен отговор (или решена задача) или 0 в противен случай. Видяхме, че вероятностното моделиране описва ситуацията с функцията $P_i(\theta)$. В същност, в случая са налице две функции. Едната от тях $P(X_i = 1/\theta) = P_i(\theta)$ е добре познатата ни вероятност произволно избран ученик да даде правилен отговор на задачата. Втората функция $P(X_i = 0/\theta) = 1 - P_i(\theta)$ е вероятността той да даде грешен отговор (да не реши задачата) (фиг. 7). Абсцисата на пресечната точка на двете функции е трудността на задачата.

Нека сега *i*-тата задача се кодира с повече от две числа, например с три числа – 0, 1 и 2. Това означава, че X_i е равно на 2, ако ученикът правилно е решил задачата, X_i е равно на 1, ако той частично е решил задачата и X_i е равно на 0 в останалите случаи. (Смисълът на частично решена задача трябва да бъде точно описан в зависимост от конкретното съдържание на задачата, за да може проверяващият недвусмислено да прецени поставянето на код 1.) Тази ситуация в едно-параметричния модел се описва чрез следните три функции:

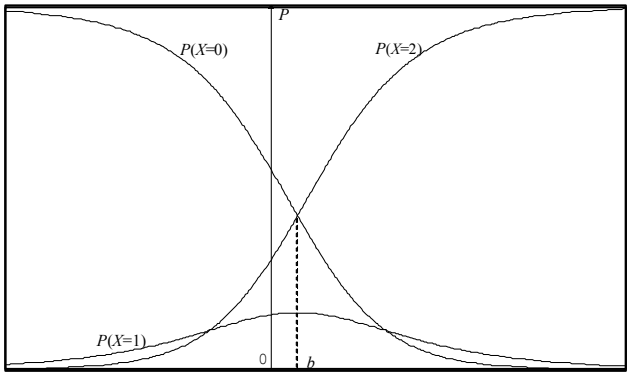


Фиг. 8. Три характеристични криви на задача

$$P(X_i = 2/\theta) = \frac{e^{2\theta - 2b_i - t_{i1} - t_{i2}}}{1 + e^{\theta - b_i - t_{i1}} + e^{2\theta - 2b_i - t_{i1} - t_{i2}}}, \quad P(X_i = 1/\theta) = \frac{e^{\theta - b_i - t_{i1}}}{1 + e^{\theta - b_i - t_{i1}} + e^{2\theta - 2b_i - t_{i1} - t_{i2}}} \quad \text{и}$$

$$P(X_i = 0/\theta) = \frac{1}{1 + e^{\theta - b_i - t_{i1}} + e^{2\theta - 2b_i - t_{i1} - t_{i2}}},$$ където $t_{i2} = -t_{i1}$. Те изразяват вероятностите произволно избран ученик да получи код съответно 2, 1 и 0 като функция на *способността θ*. За определянето на тези функции са достатъчни двата параметъра *b* и *t*, чийто графичен смисъл е ясен от фиг. 8. Параметърът *b* е *трудността* на задачата. По подобен начин се описват и задачи, кодирани с повече кодове, като функциите от типа "камбана" са повече.

От получените графики могат да се правят различни изводи. На фиг. 8 например, учениците с постижения по-малки от *b* – *t* имат най-голяма вероятност да получат код 0 на задачата. Тези с постижения между *b* – *t* и *b* + *t* имат най-голяма вероятност да получат код 1, а с постижения по-големи от *b* + *t* имат най-голяма вероятност да получат код 2.



Фиг. 9. Няма нужда от код „1"

Не винаги, обаче, предварително замисленото кодиране на задачите се оказва смислено. Графиката на фиг. 9 е на задача, която

е кодирана с числата 0, 1 и 2. При нея, обаче, няма интервал от *способностите*, в който вероятността за получаване на код 1 е най-голяма. Това означава, че тази задача би трябвало да се кодира само с два кода 0 и 2 (или по-точно с 0 и 1, съответстващи на „невярно" и „вярно"). Кодът за „частично решение" в случая и излишен, защото няма интервал от скалата на способностите, в който този код да е „преобладаващ".

6. Оценяване на постиженията на учениците.

Сега ще разгледаме как се намира онова число от скалата *способности*, което най-добре характеризира мястото на всеки ученик върху тази скала. Използват се три метода: „оценка на максималната вероятност", „оценка чрез емпирично разпределение" и „оценка чрез правдоподобни стойности" (Mislevy, R., A. Beaton, B. Kaplan & K. Sheehan, 1992.). Всеки от тях има предимства и недостатъци. Изборът се прави в зависимост от поставените цели. Тук ще опишем само един от най-използваните – „оценка на максималната вероятност". Ще се спрем на случая, когато задачите са кодирани само с 0 и 1.

Едно от основните предположения във вероятностното моделиране е, че решението на която и да е задача не зависи от решението на никоя друга задача от теста (предположение за независимост). Да разгледаме един ученик, чиито отговори на тест с n задачи са кодирани с вектора $X = (X_1, X_2, ...X_n)$. Тук X_i е кодът на отговора му на i-тата задача, т.е. X_i приема стойности 1 при правилен отговор (решена задача) или 0 в противен случай. Разглежданият метод намира онази стойност от скалата *способности*, в която постигнатият резултат X върху целия тест е с най-голяма вероятност.

Да означим с $P(X/\theta)$ вероятността за получаване на резултата X, като функция на θ. Като имаме пред вид предположението за независимост, получаваме,че $P(X/\theta) = \prod_{i=1}^{n} \left(P(X_i = 1/\theta)\right)^{X_i} \left(1 - P(X_i = 1/\theta)\right)^{1-X_i}$. Целта е да се намери най-голямата стойност на тази функция. Понеже всеки множител в дефиницията на функцията $P(X/\theta)$ е между 0 и 1, стойностите на произведението са „близки" до 0 числа. Това не е удобно от техническа гледна точка. (При компютърния софтуер, числа „близки" до 0, понякога се закръгляват на 0.) Затова се търси най-голямата стойност на функцията

$$L(\theta) = \ln P(X/\theta) = \sum_{i=1}^{n} \left(X_i \ln P(X_i = 1/\theta) + (1 - X_i) \ln\left(1 - P(X_i = 1/\theta)\right)\right).$$

Оказва се, че в голяма част от случаите тази функция има локален максимум, който е и най-голямата стойност на функцията. Функциите, които съответстват на ученици, решили правилно или всички задачи, или нито една задача, нямат локален максимум. За тях най-голямата стойност се достига в $+\infty$ (ако са решили вярно всички задачи) или в $-\infty$ (ако не са решили вярно нито една задача). Особеност се наблюдава и при ученици, които са решили вярно по-трудните задачи, а са дали грешен отговор на по-лесните. За тях съответната функция има локален максимум, но той не е най-голямата стойност на функцията. Всички такива случаи трябва да се изследват самостоятелно.

7. Идеята за относителност на измерването.

Да се върнем сега на двата типа обекти, за които говорихме в началото на тази глава: тези, които измерваме (ученици), и тези, с които правим измерването (тестови задачи). Известно е, че както задачите от даден тест могат да служат за измерване на някои характеристики на учениците (например, ученически постижения), така и учениците могат да служат за измерване на някои характеристики на задачите.

В класическата теория на тестовете качествата на задачите (трудност, разграничителна сила) зависят от учениците, които са ги решавали, а също така качествата на учениците (постиженията им) зависят от задачите, с които ги оценяваме. Тази взаимна зависимост е също едно ограничение на класическата теория. В единната скала на вероятностното моделиране този проблем не съществува, т.е. качествата на единия тип обекти не зависят от качествата на другия. Например, както вече знаем, параметрите на която и да е задача (т.е. нейните качества) могат да се определят, без да е нужно всички ученици да са работили по нея; а също така, *постиженията* на който и да е ученик могат да се оценят, без да е нужно той да е работил върху всички задачи.

8. Моделиране на ученическите постижения

Една от силните страни на вероятностното моделиране е да предлага различни математически модели за обяснение на данните от тестове за постижения (Hambleton, Swaminathan, Rogers, 1991). Ако даден модел не се съгласува с дан-

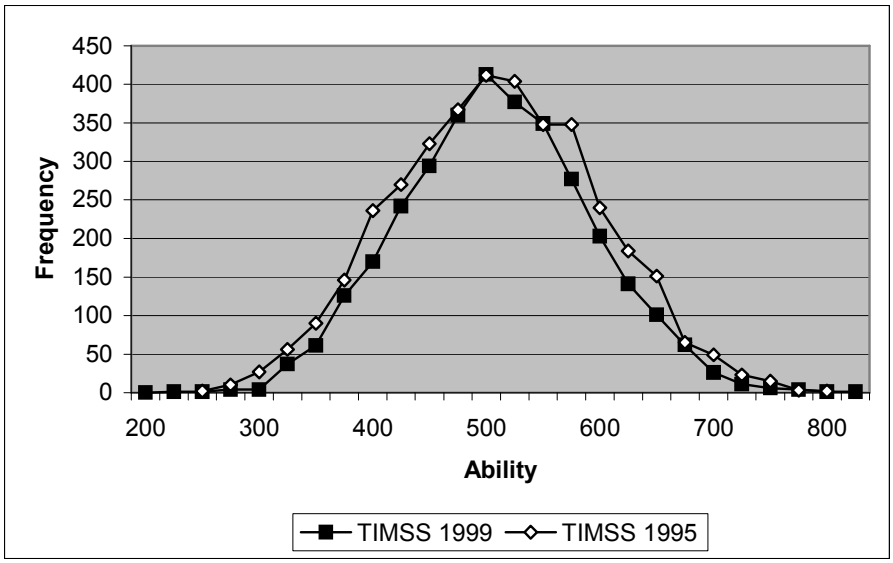
ните, изводите, които се правят, трябва да се използват предпазливо, защото те може да не са точни. При такова положение трябва да се изпробват други от моделите, които предлага вероятностното моделиране. Има различни методи за оценка на това дали даден модел се съгласува с данните, като най-разпространени са тези с използването на „хи-квадрат" теста и с графична техника.

Тук ще направим оценка на съгласуваността на модела с данните за България за 6 задачи по математика с избираем отговор, които са общи за изследванията TIMSS 1995 и TIMSS 1999, като използваме графична техника. Тези задачи са единствените математически задачи, които са давани на *всички* ученици, участвали в изследванията през 1995 и 1999 година (т.е. чрез тях се осъществява директна връзка между различните книжки). Те се появяват на едно и също място във всички осем на брой тестови книжки, използвани през тези години. (Изследванията TIMSS 1995 и TIMSS 1999 използват непълен свързан тестов дизайн, като един от клъстерите задачи е общ за всичките тестови книжки и този клъстер е на едно и също място във всичките книжки.)

Да отбележим най-напред, че данните за ученически постижения за TIMSS 1995 и TIMSS 1999 са скалирани чрез дву- и три-параметричните модели на вероятностното моделиране (Yamamoto, Kulick, 2000) като са използвани данните за всички участващи държави. Избраният модел е общ за всички данни и публикуваните международни резултати са базирани на изводите от него. След пилотната фаза, част от задачите са елиминирани поради лоша съгласуваност на модела с данните за някои от участващите държави. Ето защо трябва да се предполага, че останалите задачи дават „задоволителна" съгласуваност за всяка държава. Разбира се, всяка държава може да направи оценка за съгласуваността на общия модел със собствените си данни, за всички или някои задачи. Това дава известна представа в каква степен трябва да се доверяваме на направените изводи.

За разбиране на написаното по-долу, трябва да се знае, че скалата на ученическите постижения за TIMSS 1995 и TIMSS 1999 е със средна стойност 500 и стандартно отклонение 100. На фиг. 10 е представено разпределението на постиженията по математика на българските ученици в изследванията за TIMSS 1995 и TIMSS 1999. Както се вижда от графиката, броят на учениците с бал под 300 и на тези с бал над 700 е много малък. Понеже методът за оценка на съгласуваността на модела с данните е чувствителен към „малки" извадки (т.е. той не

е точен за интервали от скалата, в които има малко на брой ученици), най-подходящият интервал за интерпретация на съгласуваността е от бал 350 до бал 650.



Фиг. 10. Разпределение на постиженията по математика на българските ученици в за TIMSS 1995 и TIMSS 1999.

Използването на графичната техника се базира на следната процедура. За всяка задача се пресмята емпиричната и теоретичната (според използвания модел) вероятност за получаване на правилен отговор на задачата „във всяка точка от скалата на способностите". В случая, скалата на способностите е разделена на интервали от по 50 единици и учениците са групирани в тези интервали според техните постижения за всяко от изследванията. За всяка група се пресмята частта (процента) на учениците, които са решили правилно съответната задача – това е така наречената емпирична вероятност да се реши правилно задачата в съответната точка (интервал). Теоретичната вероятност за всяка точка (интервал) се пресмята от формулата, описваща математическия модел, като се използват получените оценки за параметрите на модела. Разбира се, емпиричната и теоретичната вероятност трябва да се трансформират в скалата на TIMSS 1995 и TIMSS 1999. След като тези вероятности са изчислени и съответните точки нанесени, точките в съседните интервали за съответната вероятност се свързват с отсечки. Така се получават две криви – емпирична и теоретична. Ако те са „близки" една до друга, може да считаме, че има добра съгласуваност на модела

с данните. Ако кривите показват „голямо отклонение”, съгласуваността не е добра.

По-долу са представени шестте задачи и графиките, които се получават чрез описания метод.

M012001. *Още колко малки квадратчета от фигурата трябва да бъдат запълнени, така че* $\frac{4}{5}$ *от всички малки квадратчета да бъдат запълнени?*

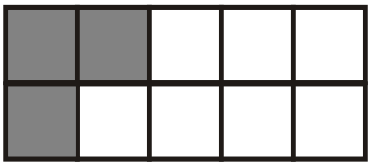
A. 5

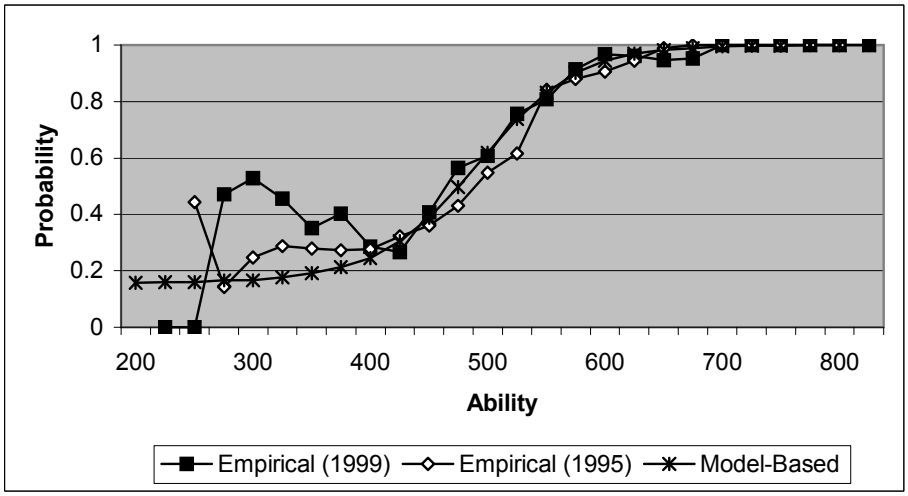
B. 4

C. 3

D. 2

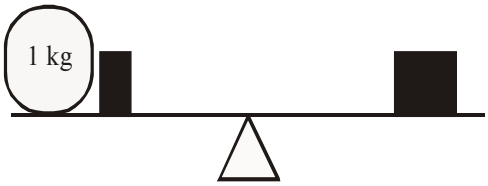
E. 1





Фиг. 11. Емпирични и теоретични вероятности за получаване на правилен отговор на задача M012001 – България, TIMSS 1995 и TIMSS 1999.

M012002. *Предметите на везната я балансират. На лявото рамо има тежест от 1 kg и половин тухла. На дясното рамо има цяла тухла.*



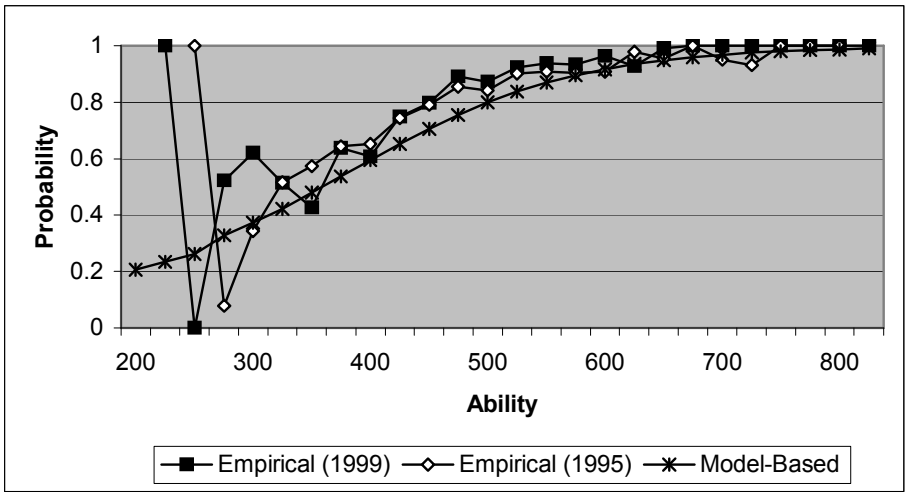
Колко тежи една цяла тухла?

A. 0,5кг

B. 1 кг

C. 2 кг

D. 3 кг



Фиг. 12. Емпирични и теоретични вероятности за получаване на правилен отговор на задача M012002 – България, TIMSS 1995 и TIMSS 1999.

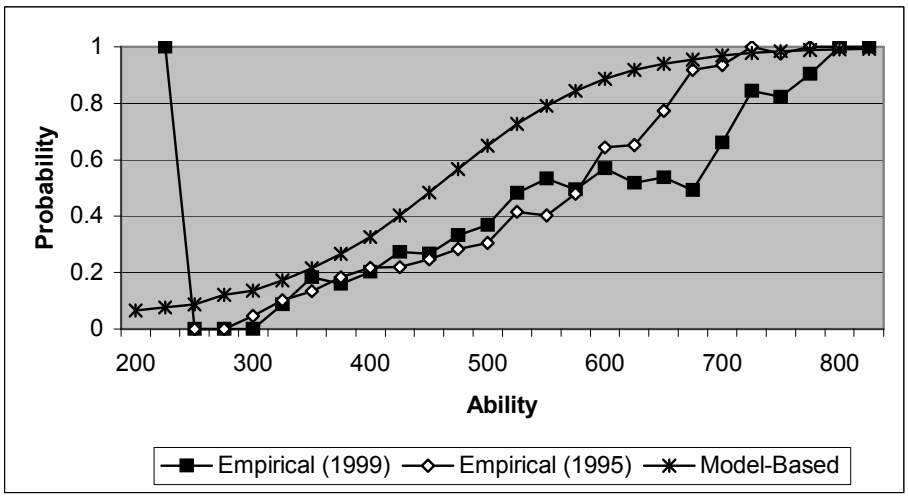
M012003. *Дължината на кутия е 9 см с точност до един сантиметър. Кои от тези стойности могат да бъдат действителна дължина на кутията?*

A. 10 см

B. 9,9 см

C. 9,6 см

D. 8,6 см



Фиг. 13. Емпирични и теоретични вероятности за получаване на правилен отговор на задача M012003 – България, TIMSS 1995 и TIMSS 1999.

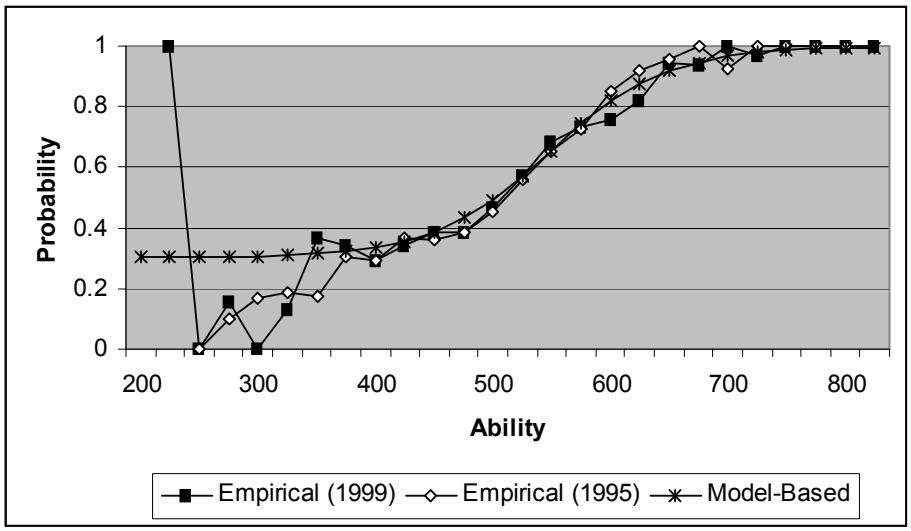
M012004. *Ани прави 4 обиколки на стадиона за същото време, за което Катя прави 3 обиколки. Колко обиколки ще направи Ани за времето, за което Катя прави 12 обиколки?*

A. 9

B. 11

C. 13

D. 16



Фиг. 14. Емпирични и теоретични вероятности за получаване на правилен отговор на задача M012004 – България, TIMSS 1995 и TIMSS 1999.

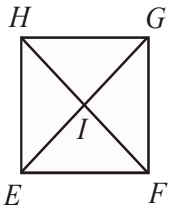
M012005. *Кои от твърденията са ГРЕШНИ за квадрата EFGH?*

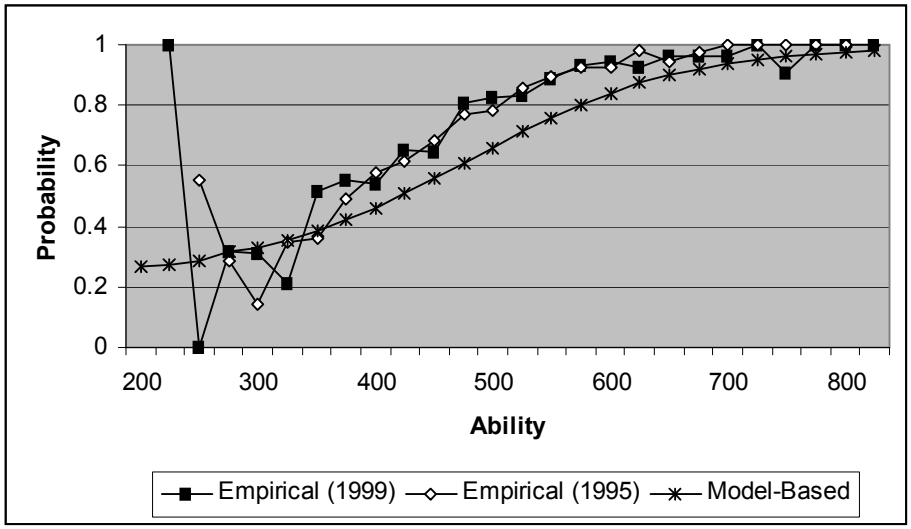
A. $\triangle EIF$ *и* $\triangle EIH$ *са еднакви.*

B. $\triangle GHI$ *и* $\triangle GHF$ *са еднакви.*

C. $\triangle EFH$ *и* $\triangle EGH$ *са еднакви.*

D. $\triangle EIF$ *и* $\triangle GIH$ *са еднакви.*





Фиг. 15. Емпирични и теоретични вероятности за получаване на правилен отговор на задача M012005 – България, TIMSS 1995 и TIMSS 1999.

***<u>M012006</u>**. От три различни теста Иван събрал 78, 76 и 74 точки, а Мария - съответно 72, 82 и 74 точки. Какъв е средният резултат на Иван в сравнение с този на Мария?*

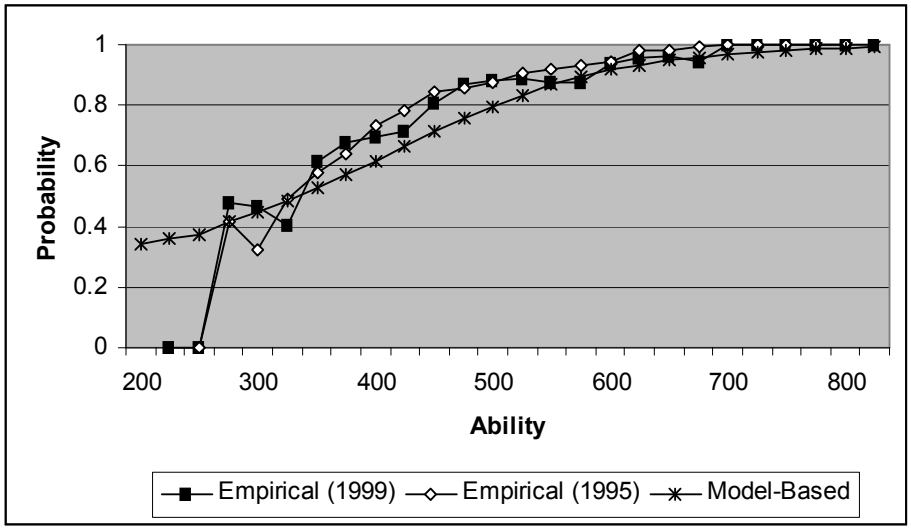
A. Иван има с 1 точка повече.

B. Иван има с 1 точка по-малко.

C. Двамата имат еднакъв среден резултат.

D. Иван има с 2 точки повече.

E. Иван има с 2 точки по-малко.



Фиг. 16. Емпирични и теоретични вероятности за получаване на правилен отговор на задача M012006 – България, TIMSS 1995 и TIMSS 1999.

Както вече казахме, интересът към горните фигури е съсредоточен върху интервала от 350 до 650 от скалата на способностите. Оказва се, че за задачи M012001 (фиг. 11) и M012004 (фиг. 14) съгласуваността на модела с данните е доста добра (в казания интервал) и за двете изследвания TIMSS 1995 и TIMSS 1999. За останалите задачи, обаче това не е така. Липса на съгласуваност се забелязва в интервала, в който попадат голяма част от учениците. В някои случаи теоретичният модел предсказва по-високи постижения на учениците, отколкото те действително са получили (емпиричните данни). Така е например със задача M012003 (фиг. 13), за която ученици със способности 500 се очаква да имат около 60% вероятност да отговорят правилно на задачата. За емпиричните данни тази вероятност е около 30% (малко по-малка за TIMSS 1995). Двете емпирично получени криви се различават съществено от теоретичната, което говори за сериозно несъответствие между модела и данните. В други случаи постиженията на българските ученици (емпиричните данни) са по-добри от онези, които предсказва теоретичния модел. Например за задачи M012002 (фиг. 12),

M012005 (фиг. 15) и M012006 (фиг. 16) учениците със способности в интервала от 400 до 500 имат между 5% и 15% по-добри постижения от теоретично предсказаните.

Подобен анализ е направен и за други математически задачи, които са общи за изследванията TIMSS 1995 и TIMSS 1999 (Kelvin, G., K. Bankov, 2006). Тези задачи, общо 47, имат за цел да свържат (да бъдат мост между) двете изследвания, т.е чрез тях се „приравняват" скалите им. Оказва се, че емпиричните данни са статистически значимо по-високи от теоретичните за 16 от тези задачи в TIMSS 1995 и 12 задачи за TIMSS 1999. Обратното, емпиричните данни са статистически значимо по-ниски от теоретичните, се наблюдава в 4 задачи в TIMSS 1995 и за 7 задачи в TIMSS 1999. Така, в около 40% от общите за двете изследвания математически задачи се наблюдава лоша съгласуваност на данните с модела. Това означава, че избраният модел не е задоволителен за анализа на българските данни от TIMSS 1995 и 1999.

ПЕТА ГЛАВА
ТЕСТОВИ ЗАДАЧИ – КЛАСИФИКАЦИЯ, ПРАВИЛА ЗА ПИСАНЕ

Тестовите задачи са градивните елементи (тухличките) на всеки тест. Затова, колкото „по-добри" задачи са налице, толкова „по-хубав" тест ще се получи. В тази глава се разглеждат различните видове задачи по отношение на техния формат. Отделя се внимание на писането на тестови задачи. Целта е да се дадат практически насоки за съставяне на множество „хубави" задачи. Смисълът на думата „хубави" става по-ясен в контекста на обясненията на насоките за писане.

1. Видове тестови задачи

Два са основните вида на тестовите задачи: (1) *задачи с избираем отговор* (ще използваме съкращението ИО); (2) *задачи със свободен отговор* (ще използваме съкращението СО). Основното различие между двата вида е, че при задачите ИО от ученика се иска да посочи правилния (според него) отговор от няколко дадени възможни отговора; при задачите СО ученикът трябва сам да напише (конструира) отговора. Всеки от двата вида има различни формати на задачите (Haladyna, 1999).

От гледна точка на измерването, много важен въпрос е дали интерпретацията на това, което измерваме, зависи (или се променя) от използването на задачи ИО или СО. Отговорът не може да се даде само с „да" или „не". В някои случаи видът на задачата има значение за измерването, а в други – този въпрос е без значение (дори няма смисъл да се поставя). Тук са разгледани предимствата и недостатъците на всеки от видовете, както и кога е удачно да се използва единият или другият вид. Препоръките са от гледна точка на: (а) съдържанието, което измерва задачата; (б) познавателната област на задачата, т.е. познавателните процеси и/или дейности, с които трябва да се ангажира ученикът, за да реши задачата.

Задачите ИО (някои ги наричат *затворени задачи*) се състоят от: (1) условие; (2) няколко възможни отговора; (3) указание как ученикът да посочи правилния (правилните), според него, отговор (отговори). Най-общо казано, тези задачи трудно се съставят, но лесно се провеждат и проверяват. Трудностите в съставянето са разгледани в следващия параграф. „Лесното" провеждане

се дължи факта, че за отговора на ученика не се изисква нищо друго, освен отбелязване. Времето за отговор е сравнително малко. Проверката на отговора е „автоматизирана" в смисъл, че може да стане от технически персонал чрез въвеждане на ученическите отговори в компютър. Това означава, че не са необходими висококвалифицирани специалисти по предмета, които да проверят ученическите отговори. Автоматизираният начин за проверката е гаранция, че в оценяването на ученическите отговори няма субективизъм.

Задачите СО (някои ги наричат *отворени задачи*) се състоят от: (1) условие; (2) поставен въпрос или задание; (3) указание какво се очаква от ученика да даде като отговор и в какъв вид. Ученикът самостоятелно съставя и записва отговора на тестовата задача. За тези задачи се казва, че лесно се съставят, но трудно се провеждат и проверяват. Въпреки че условието и поставянето на въпроса/заданието изглежда по-комплексно и „сложно" от това при задачите ИО, счита се, че писането на тези задачи са по-лесни за авторите. Обикновено авторът е наясно със същността на задачата, която иска да постави. Написването изисква само достатъчно ясна формулировка. Трудностите в провеждането идват от нуждата от повече време за отговор, както и от предвиждане на място и начин за написване на отговора. Най-трудната част при задачите СО е проверката. Всеки ученически отговор трябва да бъде прочетен, проверен и кодиран от висококвалифициран специалист по предмета според написани правила. Това изисква: (1) авторите на задачите да напишат ръководства за кодиране за всяка задача; (2) екипът специалисти, който проверява ученическите отговори, да премине на специална подготовка. Това е един от сериозните проблеми при използването на задачите СО, не само защото проверката е скъпа дейност, която изисква време, но и защото тя винаги е (в известна степен) субективна. Тези въпроси се разглеждат по-подробно в последния параграф на тази глава.

Счита се, че задачите ИО са подходящи за измерване на „по-ниски" познавателни способности, главно на ниво знание. Това твърдение трябва да се възприема с голяма доза условност. Има прекрасни задачи ИО, които измерват познавателни способности от високо ниво. Специалистите по измерване предпочитат използването на задачи ИО винаги, когато това е възможно. Причините са: (1) сравнително малкото време, което е нужно на ученика, за да отговори на задача ИО; (2) липсата на субективност при проверката; (3) добре развитите методи и скали за представяне на резултатите. От тази гледна точка, ако една за-

дача може да се зададе като задача ИО, а не като СО, добре е да се зададе като ИО. Редица изследвания показват, че в повечето случаи учениците, когато не налучкват отговора, ползват една и съща стратегия за решаване, независимо дали задача е ИО или СО.

Задачите СО са подходящи за измерване на „по-висши" познавателни способности. При тях може да се иска учениците да представят аргументи за решението, което дават, да композират текст, да дадат логически разсъждения. Така се оценява не само получен отговор, но и начинът, по който той е получен, както и разсъждения, аргументации и т.н., съпътстващи решението. Тези елементи са важни при оценяване на високи познавателни способности. От педагогическа гледна точка, не се препоръчва за оценяване да се използват само задачи ИО, а да се включват и задачи СО. Систематичното оценяване на ученическите постижения само със задачи ИО води до повърхностно запаметяване на факти, до формално отношение към „попълването" на теста с голяма доза налучкване. Това, без съмнение, води до същия начин на преподаване в процеса на обучение на учениците.

Сериозен недостатък на задачите СО е, че за тях е нужно повече време за решаване. Така, при избора за използване на такъв вид задачи трябва да се задава въпросът „заслужава ли си времето, отделено в теста за такава задача, за информацията, която ще получим от нея?" „Дълги" задачи СО, изискващи „простичък" отговор, от който малко научаваме за знанията и уменията на ученика, не са за предпочитане. Друг недостатък е, че проверката на задачите СО е „скъпа" дейност от страна на време и човешки ресурси, т.е. отново стигаме до въпроса кога това си заслужава. Важен недостатък е и субективността, която винаги съпровожда проверката на ученическите отговори.

Един от най-сериозните недостатъци, които се изтъкват при използването на задачите ИО е, че правилният отговор може да се налучква. Това, без съмнение, е така. Приемайки този факт, трябва да се използват съобразени с него начини на скалиране и оценяване на резултатите. Разработени са различни методики в това отношение. Ще да дадем само един пример, в който се вижда как може да се „компенсира" ефекта от налучкването чрез прилагане на така наречената коригираща формула за изчисляване на тестов бал.

Да предположи, че имаме тест, в който всички задачи са ИО с по N възможности за отговор всяка. Нека даден ученик е посочил R правилни и W

грешни отговора (предполага се, че сборът $R+W$ е равен на броя на задачите в теста). Коригиран тестов бал на ученика се изчислява по формулата $C = R - \frac{W}{N-1}$. Ето един конкретен пример, за да илюстрираме смисъла на коригирания бал. Да предположим, че имаме тест с 30 задачи с по 3 възможности за отговор (т.е. $N = 3$). Нека действителните знания на Стоян по този тест са такива, че той може да реши 18 задачи. Това, което прави Стоян в реална ситуация е, че той решава 18-те задачи, които знае как да реши, и налучква отговорите на останалите 12 задачи. Тъй като вероятността за налучкване на правилен отговор е 1/3, той ще налучка правилните отговори на средно 1/3 от тези 12 задачи, т.е. ще получи средно още 4 точки от налучкване. Така за Стоян имаме $R = 18 + 4 = 22$, $W = 8$ и неговият коригиран бал е $C = 22 - \frac{8}{2} = 22 - 4 = 18$, което е точно отражение на действителните му знания. Тази коригираща формула може да се прилага, когато на учениците е дадена инструкция да налучкват, в случай че не знаят как да решат задачата. Използването на формулата има някои недостатъци. Ако в горния пример ученикът действително знае да реши 22 задачи и е сбъркал (посочил е грешни отговори) на останалите 8, той ще бъде „ощетен” с 4 точки.

Резюме на казаното до тук е, че крайностите (използване само на задачи ИО или само задачи СО) не са за предпочитане. Двата вида задачи имат място в тестовото оценяване. Всеки вид има своите предимства и недостатъци. Авторите на тестове трябва добре да ги познават и да преценяват кога е най-добре да ползват единия или другия вид. Таблица 5 систематизира някои характерни свойства на двата вида задачи.

Таблица 5. Свойства на задачите ИО и на задачите СО

Свойство	**Задачи с избираем отговор (ИО)**	**Задачи със свободен отговор (СО)**
Измервани познавателни способности	По-конкретни, от по-ниско равнище, по-тривиални.	По-абстрактни, от по-високо равнище, най-ценни като резултат от обучението.
Съставяне	Условието се пише лесно, но дистракторите се съставят трудно.	По-лесно от задачите ИО, въпреки че условието е по-комплексно и изисква повече указания.

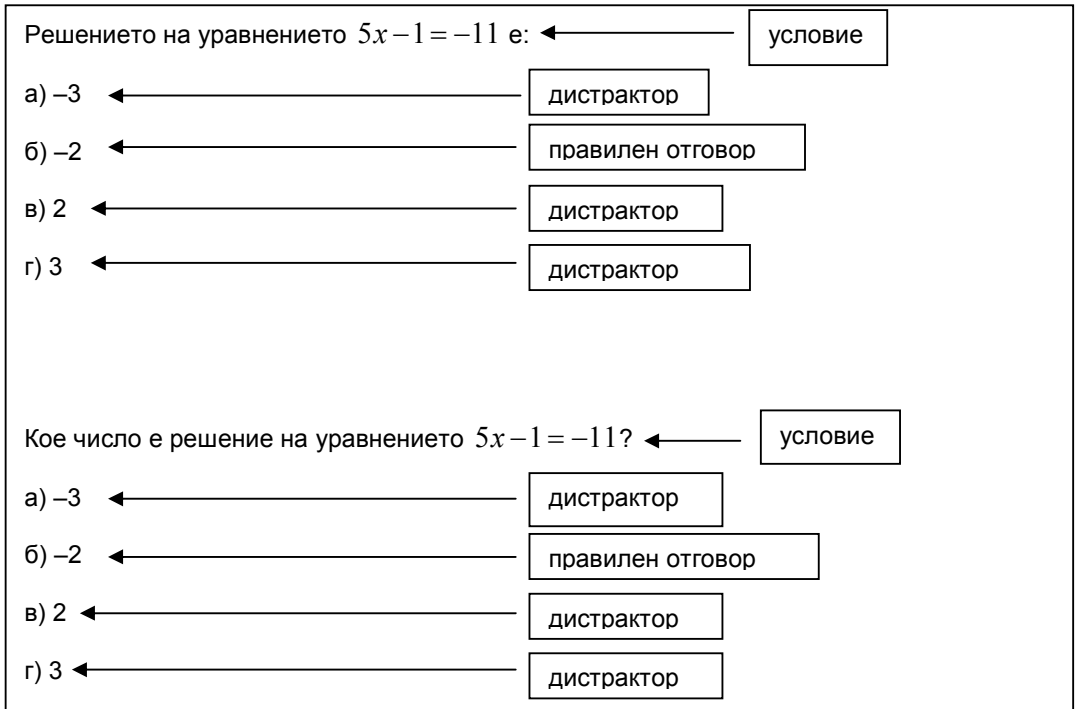
Свойство	Задачи с избираем отговор (ИО)	Задачи със свободен отговор (СО)
Провеждане	Лесно, не се изисква време или допълнителни средства.	По-трудно, нужно е повече време, както и място за написване на отговора.
Проверка	Автоматична, не е нужен висококвалифициран персонал.	Трудна, нужно е ръководство за оценяване, висококвалифицирани проверители и обучение.
Надеждност	Висока.	По-ниска, поради наличието на проверители.
Обективност	Обективна оценка.	Субективна оценка, поради наличието на проверители.

2. Формати на задачи с избираем отговор

Тук са разгледани различни формати на задачи с избираем отговор (ИО), подкрепени с примери. Коментирани са предимствата и недостатъците на форматите.

Използването на различни формати обогатява възможностите за оценяване. То внася известно разнообразие в теста и пречи на тествания да се отегчи. От друга страна, не трябва да се прекалява с различните формати в рамките на един и същ тест, главно поради трудности в точкуването и скалирането на резултатите.

Класически формат. Задачите от този формат се състоят от: (1) условие; (2) няколко възможности за отговор. Условието съдържа заданието/същността/идеята на задачата. Броят на възможностите за отговор обикновено е между 3 и 5. От тях точно една е *правилен отговор*. Останалите (неправилните) отговори се наричат *дистрактори*. От ученика се иска да посочи правилният, според него, отговор. Посочването става по указан начин (ограждане, зачертаване, попълване в лист за отговори и др.). Условието може да се постави като въпрос или като незавършено изречение (завършването му става във всяка възможност за отговор). Двата начина са еднакво приемливи. На фиг. 17 е показан една и съща тестова задача с условие най-напред като незавършено изречение, а после – като въпрос.



Фиг. 17. Класически формат на задача ИО

Правилният отговор е обективно правилен, т.е. неговата правилност не зависи от гледната точка на отговарящия или друг субективен критерий. Дистракторите са без никакво съмнение грешни отговори. Те би трябвало да са такива, че всеки дистрактор да е възможен отговор за ученици, които не притежават знания/умения да решат задачата и да не е възможен отговор за онези, които имат такива знания/умения. Един начин да се постигне това изискване е да поставят като дистрактори типични грешки, които учениците допускат при решаване на задачата.

Задачи за съответствие. Те се състоят от няколко опции, няколко условия и инструкция как да се отговори (фиг. 18). Броят на опциите и на условията при тези задачи не е задължително да е един и същ (такъв е случаят в показания пример). Опциите трябва да са хомогенни (в примера всички те са имена на планети).

Предимствата на този формат е, че с една инструкция се поставят няколко задания, вместо инструкцията да се повтаря няколко пъти. Форматът е удобен за оценяване на разбиране на понятия и принципи.

На всяка чертичка в дясната колонка напишете съответната буква на планетата от лявата колонка, така че да се получат правилни твърдения.

А. Меркурий	_____ е най-далече от Слънцето
Б. Венера	_____ е най-близко до Слънцето
В. Земя	_____ е най-голямата планета
Г. Марс	_____ е най-малката планета
Д. Юпитер	_____ има повече от две луни
Е. Сатурн	_____ има точно две луни
Ж. Уран	_____ се върти най-бавно
З. Нептун	_____ е обитаема
И. Плутон	

Фиг. 18. Задача за съответствие

Задачи с алтернативен избор. По същество, това са тип класически задачи с две възможности за отговор (фиг. 19).

За всяко изречение изберете вярната от двете алтернативи и запишете съответната буква на празното място в началото.

____ 1. Едночленът $3xy^2 2x$ (**А**-е, **Б**-не е) в нормален вид.

____ 2. Едночлените $-5xy^2$ и $\frac{1}{5}xy^2$ са (**А**-подобни, **Б**-противоположни).

____ 3. В пропорцията $\frac{a}{b}=\frac{c}{d}$ числата *a*, *b*, *c* и *d* се наричат (**А**-членове, **Б**-коефициенти).

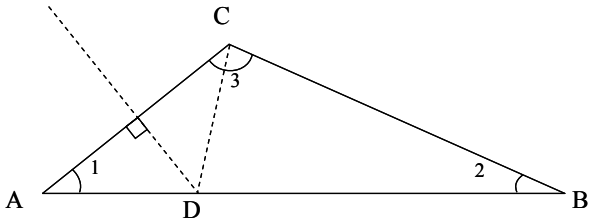
Фиг. 19. Задача с алтернативен избор

Форматът е подходящ, когато не може да се намери повече от един смислен дистрактор. Ученикът трябва да сравни две алтернативи. Това е принципната разлика от задачите „вярно-невярно” (виж по-долу), в които ученикът трябва да работи по метода „контра пример”.

Предимствата на този формат са: (1) задачите лесно се съставят; (2) изискват „малко” време за отговор, което означава, че могат да се дадат много задачи за малко време; (3) с тях добре може да се покрие обширна тематика; (4)

показват добра надеждност. Форматът е подходящ за текущо оценяване в клас. Най-сериозният недостатък е, че вероятността за налучкване на правилната алтернатива е 50%. Това трябва да се има предвид при анализа на измерителните качества на задачите от този формат.

В триъгълника *ABC*, симетралата на *AC* пресича *AB* в точка *D*, така че *AD=DB*. Кое от следните твърдения е вярно?



*Отбележете **едно квадратче** на всеки **ред**.*

	Вярно	Невярно
$\triangle ADC$ е равнобедрен	☐	☐
$\sphericalangle CDB = \sphericalangle ACB$	☐	☐
D е центърът на описаната окръжност за $\triangle ABC$	☐	☐
$\frac{AD}{DB} = \frac{AC}{BC}$	☐	☐
$\sphericalangle 1 + \sphericalangle 2 = \sphericalangle 3$	☐	☐
$\sphericalangle 3 = 90^0$	☐	☐

(Задачата е от международното изследване MT21.)

Фиг. 20. Задача „вярно-невярно

Задачи „вярно-невярно". До скоро такъв формат задачи се използваше предимно при текущо оценяване в клас, но напоследък се срещат все повече в стандартизирани тестове за масови оценявания. Разпространеното мнение, че с този формат се оценяват само тривиални знания не е вярно, както това може да се види на фиг. 20.

Предимствата и недостатъците на този формат са сходни с тези на задачите с алтернативен отговор. Форматът създава известни трудности при включването му в скалите за оценяване, особено, ако някои от условията са логически зависими (последните две в дадения пример).

Комплексни задачи с избираем отговор. При тези задачи са дадени няколко твърдения/въпроси. За всяко от тях има по няколко възможности за отговор, от които точно една е правилна. От ученика се иска да избере правилния, според него, отговор (фиг. 21).

С този формат могат да се оценяват и високи познавателни способности. Предимството му е, че с едно и също условие могат да се зададат няколко въпроса. Най-сериозният недостатък е включването на такива задачи в скалите за оценяване, особено ако сред твърденията/въпросите има логически зависими.

За всяко от следващите уравнения, отбележете най-малкото множество от числа, на които принадлежат *всичките* му решения.

*Отбележете **едно квадратче** на всеки **ред**.*

	Естествени числа	*Цели числа*	*Рационални числа*	*Реални числа*	*Комплексни числа*
$x+2=3$	☐	☐	☐	☐	☐
$x^2-4=0$	☐	☐	☐	☐	☐
$x+3=2$	☐	☐	☐	☐	☐
$x^2-2=0$	☐	☐	☐	☐	☐
$2x-3=0$	☐	☐	☐	☐	☐
$x\ \ +2=0$	☐	☐	☐	☐	☐

(Задачата е от международното изследване MT21.)

Фиг. 21. Комплексна задача с избираем отговор

Задачи с повече правилни отговора. Понякога е удачно да дадем задача от класически тип, в която правилните отговори са повече от един. Това противоречи на класическия формат. За такива цели се използва формата, показан на фиг. 22.

Особеностите на този форма са: (1) задачите се решават по-трудно от класическия вид задачите с избираем отговор; (2) по-трудно се съставят и редактират; (3) заемат много място на страницата; (4) изискват повече време за решаване; (5) имат сравнително слаба разграничителна сила.

Съставни задачи „вярно-невярно". Състоят се от условие и от известен брой фрази/изрази, някои от които са смислени/правилни, а други не са (фиг. 23). По същество имаме множество от задачи, като броят им може да е доста голям.

Кои от изброените човешки дейности НЕ вредят на природата? Отбележете А, Б, В или Г.

1. Добив на полезни изкопаеми
2. Залесяване
3. Създаване на паркове

А. Само 1 и 2

Б. Само 2 и 3

В. Само 1 и 3

Г. Всичките 1, 2 и 3

Фиг. 22. Задача с повече правилни отговора

Особеностите на този формат са: (1) позволява да се дадат много задачи наведнаж; (2) лесно се съставят; (3) времето за решаване е сравнително малко; (4) вероятността за налучкване на правилния отговор за всеки израз е 50%; (5) подходящи са за оценяване на разбиране, понеже предлагат предимно примери и контра примери.

След всяко от написаните живи същества отбележете **Р** ако е реалистично и **А** ако е абсурдно.

1. Водно млекопитаещо ____

2. Риба с бял дроб ____

3. Летящо млекопитаещо ____

4. Амеба с определено място за уста ____

5. Топлокръвно влечуго ____

Фиг. 23. Съставна задача „вярно-невярно”

Поредица от зависими задачи. При този формат най-напред се представя някакъв сценарий, част от разказ, проблемна ситуация и др. След това се дават редица задачи, свързани с постановката.

Поредицата от зависими задачи става все по-популярна форма. Фигурите от 24 до 27 представят четирите основни типа на този формат.

Прочетете текста и отговорете на въпросите след него.

Костенурка и заек спорели кой е по-бърз. Те определили мястото и разстоянието и се разделили. Заекът, като се осланял на вродената си пъргавина, не се постарал да бяга бързо, а легнал край пътя и заспал. Костенурката била наясно, че се движи бавно, затова тичала без почивка. Така надминала спящия заек и победила.

Баснята показва, че често пъти с труд човек може да надмогне други, които нехаят, разчитайки на вродените си дарби.

1. Какъв е видът на текста?
   а) разсъждение
   б) описание
   в) повествование

2. Кое е най-подходящото заглавие на текста?
   а) Костенурката и заекът
   б) На разходка
   в) Костенурката

3. Какво е значението на думата ПЪРГАВИНА от текста?
   а) работливост
   б) спешност
   в) бързина

4. Коя от изброените по-долу думи е синоним на думата ОПРЕДЕЛИЛИ от текста?
   а) обсъждали
   б) уточнили
   в) разговаряли

Фиг. 24. Поредица от зависими задачи – разбиране на текст

В термос е поставена смес от ферментиращи бактерии, захар и да при температура 15^0С. След 24 часа наблюдаваме какво е станало със сместа.

1. Какво е станало с температурата?
   а) повишила се е
   б) останала е същата
   в) понижила се е

2. Каква е причината за този резултат?
   а) бактериите дишат
   б) бактериите не дишат
   в) бактериите отнемат топлина, за да живеят
   г) топлината не може да се приема или отдава през термоса

3. Какво е станало с броя на бактериите?
   а) повишил се е
   б) намалил се е
   в) останал е почти същия

4. Какво е станало с количеството захар?
   а) повишило се е
   б) намалило се е
   в) останало е почти същото

Фиг. 25. Поредица от зависими задачи – проблемна ситуация

For each numbered pair of choices, circle the letter next to the correct spelling of the word.

There (**A-our, B-are**) many ways to invest money. You can earn (**A-interest, B-interest**) by buying saving bonds. Or you can become a (**A-part-owner, B-partowner**) of a company by owing stock in a company. As a shareholder in a company, you can share in company (**A-profits, B-prophets**).

Фиг. 26. Поредица от зависими задачи – линейно наредени задачи

СПОРТ	ПОСТРАДАЛИ	ПРАКТИКУВАЩИ (на милион)
Баскетбол	646 678	26,2
Колоездене	600 649	54,0
Футбол	453 684	13,3
Плуване	130 362	66,2
Волейбол	129 839	22,6
Вдигане на тежести	86 398	39,2
Риболов	84 115	47,0
Данните са от National Sporting Goods Association, USA.		

1. Кой спорт има най-много практикуващи?
   а) баскетбол
   в) колоездене
   в) риболов

2. В кой спорт има най-малко пострадали?
   а) баскетбол
   а) футбол
   в) риболов

3. В кой спорт има най-висок процент пострадали според броя практикуващи?
   а) баскетбол
   б) колоездене

Фиг. 27. Поредица от зависими задачи – използвана на нагледни материали

На края да направим резюме на форматите, които разгледахме. Те са:

1. Класически задачи с избираем отговор
2. Задачи за съответствие
3. Задачи с алтернативен избор
4. „Вярно-невярно”
5. Комплексни задачи с избираем отговор
6. Задачи с повече от един правилен отговор
7. Съставни „вярно-невярно” задачи

8. Поредица от зависими задачи
    - разбиране на текст
    - линейно наредени задачи
    - проблемна ситуация
    - задачи, използващи нагледни материали

Най-използвани и ясни за оценяване са:

- Класически задачи с избираем отговор
- Задачи за съответствие
- Задачи с алтернативен избор
- Съставни „вярно-невярно" задачи

По-малко се използват:

- „Вярно-невярно"
- Задачи с повече от един правилен отговор
- Комплексни задачи с избираем отговор

Поради засиления интерес към оценяването на високи познавателни равнища, все повече се използват поредица от зависими задачи.

3. Правила за писане на задачи с избираем отговор

Писането на тестови задачи е трудоемка, важна и отговорна дейност. Обикновено за нея се организира екип от съставители на тестови задачи. Всеки съставител трябва да е специалист в областта на оценяване на теста, но това не е достатъчно. Често пъти прекрасни идеи за задачи не могат да се реализират като тестова задача поради недостатъчната подготовка и/или опит на съставителя. В този смисъл писането на тестови задачи, особено на тези избираем отговор, е донякъде въпрос на опитност, която се придобива с практика.

В този параграф е систематизиран практически опит при съставяне на тестови задачи ИО във вид на правила (Haladyna 1999). Те могат да се възприемат като основни насоки към съставителите на тестови задачи. Въпреки че е добре да се спазват, правилата не са канони или закони, т.е. тяхното прилагане трябва да е разумно и творчески, а не консервативно. Изброените правила са толкова много, че на съставителите може да им се стори невъзможно да пишат тестови задачи като ги спазват всичките. Това не е далеч от истината. За да се съставят интересни и смислени задачи често някои правила се нарушават. Със-

тавителите трябва да са наясно със смисъла, който стои зад всяко правило и какъв е рискът (какво следва) при неговото нарушаване.

В таблица 6 са систематизирани правилата, които са групирани в пет категории: съдържание, формат, стил, условие, възможности за отговор. Всяко правило е обяснено по-долу с коментари и примери.

Таблица 6. Правила за писане на задачи с избираем отговор

Правила, свързани със съдържанието	1. Задачата да измерва точно едно определено съдържание и познавателна област от съдържателната рамка.
	2. Задачите да са логически независими една от друга.
	3. Да не се пита за прекалено специфично или прекалено общо съдържание.
	4. Задачата да измерва единствено знание или умение, а не навързани последователно няколко от тях.
	5. Да не се пита за съдържание, което е свързано с личностно мнение.
	6. Да не се използват похвати за нарочно „подлъгване” на ученика.
Правила, свързани с формата	7. Да се използват форматите, разгледани в параграф 8.
	8. Разположението на възможностите за отговор да е вертикално.
Правила, свързани със стила	9. Задачите трябва да се редактират и подобряват.
	10. Използваният език трябва да е съобразен с възрастта на учениците.
	11. Да се използва граматиката на книжовния български език (или съответния книжовен езика на теста).
	12. В задачата да няма излишни думи или текст, за да се минимизира времето за четене.
Правила, свързани с условието	13. Условието на задачата да е или като въпрос, или като незавършено изречение.
	14. Указанията/информацията за учениците трябва да са пределно ясни.
	15. Основната идея на задачата да се включи в условието, а не във възможностите за отговор.
	16. Да се избягва неадекватен контекст.
	17. По възможност, условието да е в позитивна форма. Ако се налага да е в негативна, думи като НЕ, НИКОГА и др. да са главни или по-тъмни.
Правила, свързани с възможностите за отговор	18. Броят на възможностите за отговор да е толкова, колкото добри такива могат да се намерят за задачата. Обикновено той е между 3 и 5.
	19. Възможностите за отговор да се поставят в логичен ред.
	20. Възможностите за отговор да са логически независими.
	21. Възможностите за отговор да са хомогенни по съдържание.
	22. Възможностите за отговор да са почти еднакво дълги.
	23. Да не се използват „никое от горните”, „всички по-горе”, „не знам” и др. подобни.
	24. Възможностите за отговор да са в позитивна форма.
	25. Да се избягват хумористични възможности за отговор.
	26. Точно един от възможностите за отговор да е правилен.

	27. Правилният отговор да се слага на различни места.
	28. Да не се дават неволни указания за правилния отговор, като: • Специфични думи като “винаги”, “никога”, “напълно”, “абсолютно” и т.н. • Предизвикване на асоциации, показващи правилния отговор • Граматически различен правилен отговор • Очебиен правилен отговор • Получаване на правилния отговор по погрешен начин • Групиране на дистракторите, за да с “види” правилният отговор • Очевидни дистрактори
	29. Всички дистрактори да са правдоподобни.
	30. За дистрактори да се използват типични грешки на учениците.

Правила, свързани със съдържанието:

1. Задачата да измерва точно едно определено съдържание и познавателна област от съдържателната рамка (виж глава 6). Целта на всяка тестова задача е да измери определено съдържание на определено познавателно равнище. Тези параметри са зададени в съдържателната рамка. Тя определя какви задачи по съдържание и познавателно равнище са нужни за теста. Затова съставителите пишат задачи само по указанието от съдържателната рамка и „свързват” всяка задача с точно една клетка от нея. Това правило не трябва да се нарушава. Задачи, които не са написани според изискванията на съдържателната рамка, нямат място в теста.

1. Коя формула е за намиране на лице на триъгълник?

а) $S = \frac{ah_a}{2}$

б) $S = ab\sin\alpha$

в) $S = \frac{pr}{2}$

2. Коя формула е за намиране на лице на успоредник?

а) $S = \frac{ah_a}{2}$

б) $S = ab\sin\alpha$

в) $S = \frac{pr}{2}$

Фиг. 28. Пример на зависими тестови задачи

2. Задачите да са логически независими една от друга. Не е допустимо решението на някоя задача от теста да зависи от решението на някоя друга задача от същия тест. Пример на зависими задачи е даден на фиг. 28. В него, ако

по някакъв начин е отговорено на задача 1, същият отговор не може да се даде и за задача 2. Това правило също не трябва да се нарушава.

3. Да не се пита за прекалено специфично или прекалено общо съдържание. Прекалено специфичното съдържание обикновено съответства на периферни и не важни знания/умения (фиг. 29). Не е разумно да се отделя време за оценяване на такива знания/умения. Прекалено общото съдържание (фиг. 30) често е толкова общо, че няма еднозначен правилен отговор или има толкова много изключения, че задачата изглежда абсурдна. Това правило би могло да се наруши.

Аликвотни дроби са тези, които:
а) имат числител 1 и знаменател естествено число
б) имат знаменател 1 и числител естествено число
в) са по-малки от 1 и по-големи от 0

Фиг. 29. Задача с прекалено специфично съдържание

Кой е най-сериозният проблем в света?
а) Глад
б) Образование
в) Болести

Фиг. 30. Задача с прекалено общо съдържание

4. Задачата да измерва единствено знание или умение, а не навързани последователно няколко от тях. Ако една задача измерва няколко неща, не може да се установи какво не знае ученик, който не я е решил, т.е. губи се възможността за диагностика. Ако е нужно да се оценят няколко свързани знания/умения, може да се използва поредица от зависими задачи. Вариант 1 от фиг. 31 е задача за 4 клас и оценява познаването на последователността за извършване на аритметични действия. Едновременно с това действието умножение в нея е сложно (за 4 клас) и може лесно да се сбърка. Така, ако някой не е посочил правилния отговор, причината може да не е, че не знае реда на аритметичните действия, а че е сбъркал самите действия. Във вариант 2, този недостатък на задачата е (до някъде) премахнат.

Вариант 1:
$5990 + 3541.7 =$
а) 30 777
б) 66 717
в) 84 687

Вариант 2:
$5990 + 4010.7 =$
а) 34 060
б) 70 000
в) 87 970

Фиг. 31. Задачата да измерва единствен елемент

5. Да не се пита за съдържание, което е свързано с личностно мнение. Такива задачи обикновено нямат обективно правилен отговор (фиг. 32). Правилото не трябва да се нарушава.

Кой е най-рационалният начин за пресмятане стойността на числовия израз 40.25.1024?

а) (40.25).1024
б) 40.(25.1024)
в) (40.1024).25

Фиг. 32. Задача, в която се пита за личностно мнение (няма обективно правилен отговор)

6. Да не се използват похвати за нарочно „подлъгване" на ученика. Целта на всяко оценяване на ученическите постижения е да поискаме от ученика да покаже какво знае, т.е. „да измъкнем" от него знанието/умението. Всяко нарочно подлъгване действа в обратната посока. Освен това, такива подлъгвания разрушават доверието на учениците в оценяването. Става въпрос за „трикове" от вида на: (1) нарочно заблуждаване на ученика; (2) поставяне на задачата в контекст, който или няма нищо общо с нея, или я представя в неестествен вид; (3) представяне на познати факти/принципи по начин, който няма нищо общо с този, по който те се изучават; и др. Правилото не трябва да се нарушава. В примера на фиг. 33 „подлъгването" е в първото изречение. Ако дадем задачата без него, такова „подлъгване" няма.

Някои месеци имат 31 дни. Колко месеца имат 30 дни?

а) 4
б) 6
в) 11
г) 12

Фиг. 33. Задача, която нарочно подлъгва ученика

Правила, свързани с формата:

7. Да се използват форматите, разгледани в предишния параграф. Това правило не се нуждае от коментари.

8. Разположението на възможностите за отговор да е вертикално. Вертикално разположените възможности за отговор са по-ясни за възприемане. Така се улеснява четенето и се съкращава времето за решаване на задачата. Хоризонтално разположените възможности за отговор намаляват мястото, на което се

отпечатва задачата, като в крайна сметка се намалява разходът на хартия. Въпреки, че се възприемат по-трудно, хоризонталното разположение на възможностите за отговор също е допустимо. Тогава трябва да осигурим достатъчно място между възможностите за отговор, за да е ясно коя къде започва и свършва. Недопустимо е при четири възможности за отговор те да се групират „две по две" (фиг. 34).

Кой курорт е разположен на Черно море?
а) Банско
б) Велинград
в) Пампорово
г) Созопол
← правилно

Кой курорт е разположен на Черно море?
а) Банско б) Велинград в) Пампорово г) Созопол
← допустимо

Кой курорт е разположен на Черно море?
а) Банско б) Велинград
в) Пампорово г) Созопол
← неправилно

Фиг. 34. Разположение на възможностите за отговор

Правила, свързани със стила:

9. Задачите трябва да се редактират и подобряват. Това е задължителна и важна част от съставянето на теста. Дори и да има притеснения от липса на време, трябва да се намери начин задачите да се редактират. Понякога и „добре редактирани" задачи допускат още по-добра редакция.

Вариант 1: Не добре редактирана задача
В кой ред съществителните имена от мъжки род са членувани правилно?
а) Конят е помощник на човекът.
б) Коня е помощник на човека.
в) Коня е помощник на човекът.
г) Конят е помощник на човека.

Вариант 2: Редакция на вариант 1
В кое изречение всички съществителни имена са членувани правилно?
а) Конят е помощник на човекът.
б) Коня е помощник на човека.
в) Коня е помощник на човекът.
г) Конят е помощник на човека.

Фиг. 35. Редактиране на задача

Вариант 2 от фиг. 35 е редактирана версия на вариант 1, а именно: (1) думата „ред" е заменена с „изречение" (ако възможностите за отговор са по-дълги има опасност да не се съберат на един ред); (2) заменено е „съществител-

ните имена” с „всички съществителни имена” – така е още по-ясно за учениците, че акцентът е за ***всички***; (3) махнато е излишното „от мъжки род” – всички съществителни имена в изречението са от мъжки род; наличието на „мъжки род” затруднява диагностиката – ако някой ученик е посочил отговор а), например, това може да е по причина, че си мисли, че „човек” не е от мъжки род.

10. Използваният език трябва да е съобразен с възрастта на учениците. Целта на теста е да измери и оцени това, което е заложено в спецификацията му, а не способността на ученика да чете текст, който евентуално не напълно разбира. Затова езикът на задачите трябва да е съобразен с най-слабата група ученици, за които е предвиден тестът. Ако четенето обърква учениците, резултатът от теста ще е отражение на някаква смес от знания/постижения, които измерваме, и възможности за разбиране на текст. Вариант 2 на фиг. 36 използва разбираемото „се внася в”, вместо трудната за разбиране дума „импортна” в задача за 4 клас.

Вариант 1: Език, несъобразен с четвъртокласници
Коя стока е импортна за България?
а) въглища
б) медна руда
в) нефт

Вариант 2: Подобрение на езика
Коя стока се внася в България?
а) въглища
б) медна руда
в) нефт

Фиг. 36. Език, съобразен с възрастта на учениците

11. Да се използва граматиката на книжовния български език (или съответния книжовен езика на теста). Граматичните грешки разсейват учениците. Тест с граматични грешки прави лошо впечатление, дори у учениците. Освен това понякога граматичните грешки могат да доведат до двусмислие в задачата или дори да я направят грешна.

12. В задачата да няма излишни думи или текст, за да се минимизира времето за четене. Времето за прочитане на задача ИО трябва да е малко. Така може да се дадат повече задачи за предвиденото време за теста. Премахването на излишни думи спомага за тази цел. Това, разбира се, не трябва да става за сметка на лоша или неразбираема формулировка. Вариант 2 на фиг. 37 е с по-малко думи от Вариант 1, като формулировката е също толкова разбираема.

Вариант 1: Излишни думи в условието
Кое от следните твърдения е вярно?
а) Правилните дроби са по-малки от 1.
б) Правилните дроби са по-големи от 1.
в) Правилните дроби са равни на 1.

Вариант 2: Без излишни думи
Кое е вярното твърдение?
а) Правилните дроби са по-малки от 1.
б) Правилните дроби са по-големи от 1.
в) Правилните дроби са равни на 1.

Фиг. 37. Премахване на излишни думи

Правила, свързани с условието:

13. Условието на задачата да е или като въпрос, или като незавършено изречение. Това е коментирано в параграф 2.

14. Указанията/информацията за учениците трябва да са пределно ясни. Ученикът винаги трябва да е наясно точно какво го питат и какво се очаква от него (фиг. 38).

Вариант 1: Неясно условие
За звездите и планетите:
а) движат се около други планети
б) светят със собствена светлина
в) всички те са небесни тела

Вариант 2: Ясно условие
Кое е общото между звездите и планетите?
а) движат се около други планети
б) светят със собствена светлина
в) всички те са небесни тела

Фиг. 38. Неясно и ясно условие

15. Основната идея на задачата да се включи в условието, а не във възможностите за отговор. Понякога авторите на задачи, в стремежа си да направят кратко условието на задачата, оставят основната идея на задачата във възможностите за отговор. Това не само че прави дълги всички възможности за отговор, но и разсейва учениците (фиг. 39).

Вариант 1: Основната идея не е в условието.
Маслодайни рози:
а) растат предимно по долината на Струма
б) са гордостта на Софийското поле
в) се отглеждат в Задбалканските котловини

Вариант 2: Основната идея е в условието.
Къде в България се отглеждат маслодайни рози?
а) долината на Струма
б) Софийското поле
в) Задбалканските котловини

Фиг. 39. Основната идея да е в условието на задачата

16. Да се избягва неадекватен контекст. В стремежа си да напишат „реалистични" задачи някои автори поставят контекст, който няма нищо общо със задачата. Това ненужно удължава условието и разсейва учениците (фиг. 40). Тук не става въпрос за задачи, които естествено се нуждаят от контекст и той е неразделна част от тях.

Вариант 1: Излишен контекст
Високи температури и много валежи характеризират влажния климат. Хората дишат трудно. Дори денят с умерени температури е неприятен. На кое място климатът е такъв?
а) саваната
б) тропическата гора
в) тундрата

Вариант 2: Без излишен контекст
Къде климатът се характеризира с високи температури и много валежи?
а) в саваната
б) в тропическата гора
в) в тундрата

Фиг. 40. С и без излишен контекст

17. По възможност, условието да е в позитивна форма. Ако се налага да е в негативна, думи като НЕ, НИКОГА и др. да са с главни или по-тъмни букви. Поставянето на условието в негативна форма често води до затруднения за учениците да го разберат правилно. В някои случаи се налага да се постави условието в негативна форма (например, ако целта на задачата е да оцени разбиране на отрицание). За да се избегне използването на негативна форма, понякога е добре да се промени формата на задачата (фиг. 41).

Вариант 1: Условие в негативна форма
Броят на стените на призма НЕ можа да бъде
а) 4
б) 7
в) 13

Вариант 2: Условие в позитивна форма и промяна на формата на задачата
На колко може да е равен броят на стените на призма? За всяка възможност отбележете едно от квадратчетата "да" или "не".

*Отбележете **едно квадратче** на всеки **ред**.*

	ДА	НЕ
3	☐	☐
4	☐	☐
5	☐	☐
7	☐	☐
13	☐	☐
15	☐	☐

Фиг. 41. Промяна на условието от негативна в позитивна форма

Правила, свързани с възможностите за отговор:

18. Броят на възможностите за отговор да е толкова, колкото добри такива могат да се намерят за задачата. Обикновено той е между 3 и 5. Намирането на подходящи дистрактори на задачите е едно от трудните неща при писане на задачи ИО. Дистракторите трябва да са такива, че да „работят", т.е. да бъдат избирани от ученици, които нямат необходимите знания/умения да решат правилно задачата. Понякога, за да се намерят такива дистрактори, в началото авторите може да поставят повече възможности за отговор. След анализ на данните от апробацията да оставят само онези дистрактори, които действително работят. Да се напишат повече от 4 работещи дистрактора (т.е. 5 възможности за отговор) е често пъти невъзможно. Редица изследвания показват, че 3 възможности за отговор са напълно достатъчни (в голяма част от случаите). Традицията да се пишат повече от 3 възможности за отговор е доста силна и затова много автори избират „златната среда" от 4 възможности за отговор.

19. Възможностите за отговор да се поставят в логичен ред. Така те се възприемат/осъзнават по-лесно и се намалява времето за решаване на задачата. „Логичен ред" може да е нарастващ или намаляващ (за числа) азбучен ред (за

думи или изрази) и др. В примера на фиг. 42 лявата колона съдържа възможности за отговор в произволен ред, а дясната – в нарастващ ред.

Петър за 2 часа прочита 30 страници. За колко минути ще прочете 20 страници?

а) 60	а) 60
б) 100	б) 70
в) 70	в) 80
г) 80	г) 100

Фиг. 42. Възможности за отговор в логичен ред

20. Възможностите за отговор да са логически независими. Ако някои от възможностите за отговор са логически свързани, може да се получи повече от един правилен отговор. Освен това по логически съображения могат да се изключат някои от дистракторите като неправилни отговори, без ученикът да знае да решава задачата по същество. Това правило не трябва да се нарушава. Примерът на фиг. 43 показва задача, в която не е ясно как да се отговори, ако физическият връх се достига например при 24 годишна възраст. Примерът на фиг. 44 е често срещана грешка за логически зависими отговори, когато едно понятие се включва в някое друго. В случая, всеки равностранен триъгълник е равнобедрен, следователно, ако а) не е правилният отговор, то и б) не е правилен.

Кой интервал е решение на неравенството $2x - 6 < x + 3$ за положителни x?

а) (0; 9)
б) (0; 3)
в) (9; 18)

Фиг. 43. Пример с припокриващи се интервали

В един триъгълник медианата към една от страните му е равна на половината от тази страна. Видът на триъгълника е:

а) равнобедрен
б) равностранен
в) правоъгълен

Фиг. 44. Изключване на дистрактор по логически съображения

21. Възможностите за отговор да са хомогенни по съдържание. Нехомогенните по съдържание възможности за отговор често съдържат неволна подсказка за правилния отговор, особено, ако точно той се различава по съдържание от останалите. На фиг. 45 единствената възможност за отговор, която не е свързана с въртенето на Земята, е правилният отговор. Правилото може да се нарушава, но трябва много да се внимава.

Коя е причината за различната дължина на деня и нощта през различните сезони в България?

а) въртенето на Земята около Слънцето
б) въртенето на Земята около оста й
в) наклона на земната ос

Фиг. 45. Нехомогенни по съдържание възможности за отговор

22. Възможностите за отговор да са почти еднакво дълги. Смисълът на това правило е подобен на смисъла на предишното. В примера на фиг. 46 правилният отговор в) е по-къс от дистракторите.

Сборът от два ъгъла в един триъгълник е по-малък от третия. Видът на триъгълника е

а) правоъгълен разностранен
б) равнобедрен, но не равностранен
в) тъпоъгълен

Фиг. 46. Различно дълги възможности за отговор

23. Да не се използват „никое от горните”, „всички по-горе”, „не знам” и др. подобни. Всяка задача в теста има правилен отговор и няма основателна причина той да не се даде като възможност. Поради това „никое от горните” не е смислено да се поставя. Възможността „всички по-горе” показва, че са дадени повече от един правилен отговор, което не е редно. Целта на „не знам” е да се ограничи налучкването. Но учениците не са склонни да си признават, че не знаят. Освен това налучкването не е проблем, стига да се вземе пред вид при анализа на резултатите.

24. Възможностите за отговор да са в позитивна форма. Смисълът на това правило е както смисъла на правило 17. Ако се налага да се използва отрицание, по-добре то да е в условието на задачата, отколкото във възможностите за отговор.

25. Да се избягват хумористични възможности за отговор. Наличието на „хумористични” възможности за отговор води до разсейване. Учениците не обръщат сериозно внимание на такава задача. Освен това, хумор за някои може да доведе до отрицателни емоции за други.

26. Точно един от възможностите за отговор да е правилен. Това е характерна черта на задачите ИО. Авторите трябва много да внимават за стриктното спазване на правилото. Примерът от фиг. 47 е типичен в това отношение. Така зададена задачата има три правилни отговора: а), в) и г) (например, числото 5 за а) и в) и числото 1,5 за а) и г)). Идеята на автора е да се реши неравенството.

Ако условието на задачата се формулира като „Да се реши неравенството”, тогава единственият правилен отговор е само а).

Ако $2 - x < 1$, то
а) $x > 1$
б) $x < 1$
в) $x > 2$
г) $x < 2$

Фиг. 47. Задача с повече от един правилен отговор

27. Правилният отговор да се слага на различни места. Това намалява възможността учениците да забележат някаква закономерност за появата на правилния отговор. Добре е в рамките на целия тест правилният отговор да се появява почти равномерно на всички възможни места.

28. Да не се дават неволни указания за правилния отговор, като:

- Специфични думи като “винаги”, “никога”, “напълно”, “абсолютно” и т.н. (на фиг. 48 правилният отговор е единствената възможност без дума за абсолютизиране).

Коя дейност може да повиши ефикасността от домашната работа?
а) Никога не давайте домашна работа в петък
б) Домашната работа да е по взетия урок
в) Винаги проверявайте домашната работа за другия ден

Фиг. 48. Неволно указание с думи за абсолютизиране

- Предизвикване на асоциации, показващи правилния отговор. Примерът на фиг. 49 не се нуждае от коментари.

Всеки равнобедрен триъгълник има:
а) прав ъгъл
б) тъп ъгъл
в) равни бедра

Фиг. 49. Асоцииации, показващи правилния отговор

- Граматически различен правилен отговор. На фиг. 50, правилният отговор в) има различна от дистракторите граматична структура.

За какво служат хрилете на рибите?
а) подпомагане на движението им
б) отделяне на излишната вода от тялото
в) да извличат кислорода, разтворен във водата

Фиг. 50. Различна граматична структура на правилния отговор

- Очебиен правилен отговор. Отговорът на примера от фиг. 51 трябва да е число, завършващо на нула, а такъв отговор е само един.

Стойността на израза $25.8 + 60$ е:
а) 161
б) 260
в) 1605
г) 1702

Фиг. 51. Очевиден отговор

- Получаване на правилния отговор по погрешен начин. На фиг. 52 е показан пример, в който правилият отговор може да се получи освен като вярно се реши задачата, още и като се допусне често срещаната грешка, че учениците бъркат понятията обиколка и лице.

На колко квадратни сантиметра е равно лицето на квадрат със страна 4 см.?
а) 4
б) 16
в) 32
г) 64

Фиг. 52. Получаване на правилен отговор по погрешен начин

- Групиране на дистракторите, за да с "види" правилният отговор. На фиг. 53, дистракторите се отнасят до иглолистни растения, а само правилният отговор е за широколистни.

Какво е растението, от корена на което излизат няколко твърди и здрави стъбла, покрити с кора, и листата му падат през есента?
а) широколистен храст
б) иглолистно дърво
в) иглолистен храст

Фиг. 53. Групиране на дистракторите

- Очевидни дистрактори. В лявата колона на примера от фиг. 54 веднага се забелязва, че а) и б) са дистрактори, защото в тях няма синус. Значително по-добре стоят нещата в дясната колона.

Кое тъждество е синусовата теорема?

а) $a^2 + b^2 = c^2$	а) $\sin\alpha = \frac{a}{c}$
б) $h_c^2 = a_1 b_1$	б) $S = \frac{1}{2} bc \sin\alpha$
в) $\frac{a}{\sin\alpha} = \frac{b}{\sin\beta}$	в) $\frac{a}{\sin\alpha} = \frac{b}{\sin\beta}$

Фиг. 54. Очевидни и не толкова очевидни дистрактори

29. Всички дистрактори да са правдоподобни. Няма смисъл да се поставят дистрактори, които никой не избира. В примера на фиг. 55 градовете, които не са в България, са неправдоподобни дистрактори.

Къде е обесен Васил Левски?
а) край Лондон
б) край Париж
в) край София
г) край Берлин

Фиг. 55. Неправдоподобни дистрактори

30. За дистрактори да се използват типични грешки на учениците. Това значително увеличава възможностите да се пишат правдоподобни дистрактори. На примера от фиг. 56, учениците „знаят", че или трябва да съберат, или да извадят двете числа 58 и 47. Разбира се, действието изваждане в този случай е грешно. Тогава е редно да дадем дистрактор „11", защото това е много типична грешка в този случай.

Ако $\square - 47 = 58$, то в $\square$ трябва да се запише числото:
а) 95
б) 105
в) 115

Фиг. 56. Липса на типична грешка в дистракторите

4. Задачи със свободен отговор

Задачи СО са тези, при които ученикът самостоятелно съставя и записва отговора на тестовата задача. Счита се, че те са подходящи за оценяване на високи познавателни способности. За разлика от задачите ИО, при задачите СО няма толкова разнообразни формати. Самият вид предоставя достатъчно разнообразие относно начините за представяне на задачата. Основните формати на задачите СО са три и са разгледани по-долу.

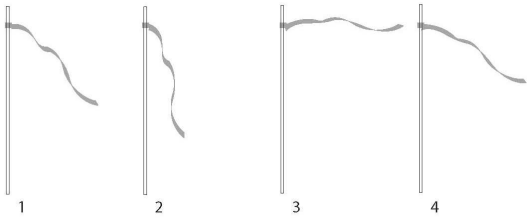
Задачи с кратък затворен отговор. При тези задачи от ученика се иска да даде отговор, без да описва (обосновава) как е достигнал до него. Допустими са *малки* вариации в начина на представяне на отговора (например, еднакво правилни са отговорите 2/5, или 0,4, или 40%). От ученика може да се иска да напише число или дума, да подчертае част от текст, да допълни диаграма и др. Може да се поиска също ученикът да подбере няколко от изброено множество

възможности, които удовлетворяват дадено условие, или да съпоставя части от изречения или диаграми. На фиг. 57 са дадени няколко примера на такива задачи.

1) След увеличение цената на една стока с 6 процента стоката струва 15,90 лв. Каква е била цената на стоката преди увеличението?
Отговор: _____лв.

2) В изречението по-долу подчертайте думата, която е написана неправилно.
През зимата тук идват много скийори и кънкьори.

3) Лентичка е закачена на пилон, за да измерва силата на вятъра.



Напишете числата 1, 2, 3 и 4 в правилния ред, така че да показват силата на вятъра от най-силен към най-слаб.
Отговор: ___;___;___;___

(Задачата е от международното изследване TIMSS.)

4) Според агрегатното състояние веществата могат да са твърди, течни или газообразни.
В таблицата са написани няколко вещества, групирани според агрегатното им състояние.
В третата колона на таблицата напишете агрегатното състояние на всяка група.

Група	Вещества	Агрегатно състояние
1	Вода и сок	
2	Въздух и кислород	
3	Скала и злато	

Фиг. 57. Примери на задачи с кратък затворен отговор

Задачи с кратък отворен отговор. При тези задачи от ученика се иска само да генерира отговор, без да описва (обосновава) как е достигнал до него. Правилният отговор може да се даде в *различни варианти*. Например, можа да се поиска да се опише процес/алгоритъм, да се даде кратко обяснение за нещо, което се е случило, да се направи чертеж или диаграма и др. На фиг. 58 са дадени няколко примери на такъв вид задачи.

1) Как може да се определи възрастта на едно дърво след като то е отрязано?
_____
_____

2) Учителката поставила задача на Карл да раздели един правоъгълник на 5 части, от които 4 са триъгълници и една е квадрат. Начертайте линии в правоъгълника по-долу, за да покажете как Карл би могъл да направи това.

3) Ледът и течната вода се използват по различен начин от човека. Напишете по един начин за използване на леда и на течната вода.

Използване на леда: _____

Използване на течната вода: _____

Фиг. 58. Примери на задачи с кратък отворен отговор

Задачи с разширен свободен отговор. При тези задачи от ученика се иска да даде не само отговор, но и да напише обяснение как той е стигнал до него; да създаде аргументиран текст (есе) и др. Решението на тези задачи понякога е дълго и сложно. Правилното решение (обяснение, текст) можа да стане по различни начини. На фиг. 59 са дадени няколко примера за такива задачи.

1) Ако n е цяло число, кое от числата n+2 и 2n е по-голямо? Обяснете отговора си.
(Задачата е от международното изследване TEDS-M.)

2) В един селски двор има овце и кокошки – общо 27 животни. Броят на краката на всички овце и кокошки е 78. Колко са овцете и колко са кокошките?
Решете задачата по два различни начина.

Първи начин:

Втори начин:

3) Една лятна сутрин Галя оставила чиния с вода до слънчев прозорец. Вечерта в чинията нямало вода.

Обяснете защо водата от чинията е изчезнала.
(Задачата е от международното изследване TIMSS.)

4) Съчинете случка в рамките на пет-шест изречения, като използвате зададеното начало.
Един ден Петьо реши да отиде в Луна-парк. Беше слънчев и топъл ден. Докато вървеше по улицата, той срещна своите приятели Митко и Ваньо, които също отиваха към Луна-парка. Тръгнаха заедно. Когато стигнаха там,
.....
.....
.....

Фиг. 59. Примери на задачи с разширен свободен отговор.

5. Проверка и кодиране на задачи със свободен отговор

Една от най-трудоемките и скъпи (на човешки и финансови ресурси) дейности при използване на задачи СО е тяхната проверката и кодирането. За всяка задача това означава, че отговорът на всеки ученик по нея трябва да се прочете от специалист по съответната област, да се провери и кодира съгласно предварително написано ръководство за кодиране. В зависимост от същността на задачата, както и от нейния формат, тези ръководства са различно дълги и подробни. Екипът от специалисти, който извършва тази дейност, преминава на задължителна предварителна подготовка с цел да се намали субективността на кодирането.

Задачите с кратък затворен отговор понякога не се нуждаят от проверка, ако на ученика е казано ясно и точно как да представи отговора, който е число или дума. Например, ако инструкцията към отговора на задачата е „Напишете отговора като естествено число с не повече от три цифри.”. Дори и да е необходима, проверката на тези задачи е сравнително лесна и не изисква много време. Ръководствата за кодиране са кратки и ясни. Подготовката на екипа за кодиране на тези задачи става бързо.

Задачите с кратък отворен отговор винаги се нуждаят от проверка, защото правилният отговор може да се даде по различни начини. Проверката им не е трудна, но изисква повече време от задачите с кратък свободен отговор. Ръководствата за оценяване са по-сложни. Подготовката на екипа е по-продължителна.

Задачите с разширен свободен отговор се нуждаят от сериозна проверка. Ръководствата за кодиране на тези задачи са по-дълги и подробни. Съставянето им е трудоемка дейност. Често те съдържат примерно кодирани решения. Подготовката на екипа по тези задачи е продължителна, разглеждат се конкретни ученически решения, коментира се тяхното кодиране. Самата проверка изисква време. Нужно е да се измерва надеждността на кодирането, както и съгласуваността между проверителите.

Ръководството за кодиране на дадена задача описва кода, който се дава на всеки вид ученически отговор върху задачата. Кодът обикновено е естествено число, което е свързано с броя точки, давани за съответния вид решение. Според броя точки, които може да се получат, задачите са: (1) едноточкови – те

се оценяват с 1 точка за правилен отговор и с 0 точки във всички останали случаи; (2) многоточкови – те се оценяват с максимален брой точки при правилен/изчерпателен отговор, с 0 точки при неправилен/неуместен/несвързан или не даден отговор и с някакво междинно число точки при отговор, който показва частичен смислен напредък по задачата (например, използване на подходящ метод за решаване, но допусната техническа грешка; или смислени, но непълни разсъждения и др.).

Таблица 7 е ръководство за кодиране на едноточкова задача – това е първата задача от фиг. 57.

Таблица 7. Ръководство за кодиране на едноточкова задача

Код	**Отговор**	**Задача No. 1 от фиг. 57**
	Правилен отговор	
1 точка	15, или 15 лв., или 15,00, или 15,00 лв.	
	Неправилен отговор	
0 точки	Написано е нещо различно от правилния отговор.	
	Без отговор	
9	Празно	

Код 9 (празно) се дава само когато ученикът не е писал нищо по задачата, т.е. няма видим белег, че той е правил опит да работи върху нея. Ако ученикът е сложил някакъв белег или знак, това не се кодира с код „празно”. Наличието на код „празно” дава възможност да се проследи дали времето, определено за теста, е било недостатъчно. Ако се окаже, че процентът на „празните” кодове в последните задачи е „голям”, това е сигурна индикация, че част от учениците не са стигнали до края на теста, т.е. времето за тях е било недостатъчно.

Фиг. 60 представя двуточкова задача и нейното ръководство за оценяване.

Понякога (особено за диагностични цели) се използва двуцифрена кодова система. (Например, изследванията TIMSS, PIRLS, PISA, TEDS-M използва такава.). *Първата цифра* показва степента на коректност (правилността) на отговора. В този смисъл тя е аналогична на броя точки, както е в предишните примери. *Втората цифра* е диагностична. Тя дава информация за евентуалните различни подходи (методи), които учениците са използвали за решаване на задачата или за типични грешки при неправилните (частично правилните) отговори.

Задача. Таблицата показва някои свойства на три вещества X, Y и Z. Едното от тях е *желязо*, другото е *вода*, третото е *кислород*.

Вещество	Точка на топене/замръзване (^{0}C)	Точка на кипене (^{0}C)	Добър проводник на електричество
X	-218	-183	Не
Y	1535	2750	Да
Z	0	100	Не

Определете всяко от веществата.

Веществото X е: _____

Веществото Y е: _____

Веществото Z е: _____

(Задачата е от международното изследване TIMSS.)

Ръководство за кодиране.

Код	**Отговор**	**Задача No.**
	Правилен отговор	
2 точки	X = кислород; Y = желязо; Z = вода.	
	Частично правилен отговор	
1 точка	Само едно или две от веществата са правилно определени.	
	Неправилен отговор	
0 точки	Нито едно вещество не е правилно определено.	
	Без отговор	
9	Празно	

Фиг. 60. Кодиране на двуточкова задача

За правилен или частично правилен отговор първата цифра определя броя на точките за съответния отговор. Тя варира от 1 до максималния брой точки за задачата. Обикновено максималният брой точки е 2, 3, рядко стига до 4. Понеже няма двуцифрени числа, започващи с цифрата 0, за удобство първата цифра за неправилен отговор е 7. Първата цифра за „празен” отговор е 9. Втората цифра носи диагностична информация. Тя може да е от 0 до 8 (т.е. кодовете изглежда, например, така: 20, 21,..., или 10, 11,..., или 70, 71,...). Втора цифра 9 означава „други”. Тя се използва за диагностични категории, които не са описани в дадените кодове. Например, с кодове 70 и 71 може да са описани две типични ученически грешки при решението на дадена задача. Код 79 се използва за всякакви неправилни отговори, които не са описани в кодовете 70 и 71. Особено „внимателни” трябва да бъдем при използването на „други” при частично правилните отговори, за да не се допусне „злоупотреба” с това, кое се счита за

частично правилно. Код 99 означава „празно” в смисъла, в който по-горе беше коментиран код 9 като „празно”.

Таблица 8 представя ръководство за двуцифрено кодиране на задачата от фиг. 60. С тези кодове има възможност да се диагностицира частично правилният отговор – кои от веществата са по-трудни или по-лесни за разпознаване.

Таблица 8. Двуцифрено кодиране на задачата от фиг. 60

Код	**Отговор**	**Задача No.**
	Правилен отговор	
20	X = кислород; Y = желязо; Z = вода.	
	Частично правилен отговор	
10	Точно две от веществата са правилно определени.	
11	Само кислородът е правилно определен.	
12	Само желязото е правилно определено.	
13	Само водата е правилно определена.	
	Неправилен отговор	
79	Всеки грешен отговор (включително зачертани, изтрити, неясни бележки и др.)	
	Без отговор	
99	Празно	

Понякога към ръководствата за оценяване се включват допълнителни бележки, примери и други упътвания. Например, към таблица 8 може да се включат следните бележки:

<u>Към код 20</u> – Бележка: Думата *кислород* НЕ можа да се замени с *въздух*. Ако се използва *въздух*, кодът е 10.

<u>Към код 10</u> – Примери: Въздух, Желязо, Вода;
Кислород, Желязо, Течност;
Кислород, ___ (празно място), Вода;
___ (празно място), Желязо, Вода

<u>Към код 11</u> – Примери: Кислород (въздух), Вода, Желязо;
Кислород (въздух), ___ (празно), ___ (празно).

Бележка: Тук се допуска използването на думата *въздух* вместо *кислород*.

Аналогични примери могат да се дадат и към кодовете 12 и 13.

Кодиране по критерии. Това е също начин за кодиране на задачи СО. Критериите са ясно описани гледни точки, които са важни за изпълнението на задачата. Всеки критерий се кодира сам за себе си (отделно и независимо от другите критерии). Като пример е разгледан начин на кодиране с използване на критерии за задача 4 от фиг. 59.

Описание на критериите:

I критерий – Дали продължението е логическо следствие от даденото начало.

II критерий – Съответства ли продължението на целта на общуване.

III критерий – Съответства ли продължението на условията на общуване.

IV критерий – Съответства ли продължението на предмета на общуване.

V критерий – Оформени ли са микротемите в текста.

VI критерий – Има ли смислова свързаност.

VII критерий – Има ли граматическа свързаност.

За **всеки критерий** по отделно се приписват:

- код 2 за напълно покрити изисквания на критерия.
- код 1 за частично изпълнение на критерия.
- код 0 за неизпълнени изисквания, които критерият поставя.

На фиг. 61 е представено как би изглеждала писмената работа на ученика с кодове по всеки от критериите. Ако всеки код съответства на броя точки по критерия, има смисъл най-отдолу да се даде общия сбор от точките.

Работата на ученика

I критерий: код _____

II критерий: код _____

III критерий: код _____

IV критерий: код _____

V критерий: код _____

VI критерий: код _____

VII критерий: код _____

Общо точки: _____

Фиг. 61. Оценяване по критерии

Ръководствата за кодиране се съставят от авторите на задачите. Те, както и самите задачи, минават през апробация, т.е. в реална ситуация ръководствата се използват за кодиране на получени ученически решения. Често, в процеса на реално кодиране по време на апробацията ръководствата търпят промени. Причината е, че не винаги може да е обхванат възможните видове решения, които учениците предлагат. Ако някакъв вид често срещано решение не е включено в

ръководството, то трябва да се добави. Понякога се оказва, че вид решение, което го има в ръководството, въобще не се среща сред ученическите работи. Тогава то се „обединява” с някой от сходните му видове. (Не е целесъобразно въобще да се премахне, понеже при реалното провеждане на теста биха могли да се появят, макар и малък брой, такива решения.)

В началото на тази глава беше казано, че колкото „по-добри” задачи са налице, толкова „по-хубав” тест ще се получи. В края на главата ще напишем някои белези на „хубави” задачи от гледна точка на съдържание и смисленост.

„Хубави” задачи може да са такива, които:

- фокусират върху важни области от обучението;
- са конструктивни и смислени задания;
- са безпристрастни;
- наблягат на централни/важни знания и умения, а не на тривиални детайли;
- ясно казват на учениците какво се очаква от тях като задание;
- са самостоятелни и не зависят от това дали учениците са изпълнили или не някоя(и) от предишните задачи.

В такива задачи не трябва да има:

- труден за разбиране изразен стил;
- дълги изречения;
- сложни изречения;
- логически труден за разбиране език;
- неясно поставени въпроси;
- противоречивост;
- неясни думи или термини, които не са познати за учениците, за които е предназначен теста;
- грешни и/или подвеждащи чертежи;
- неточни или неправилно направени диаграми.

ШЕСТА ГЛАВА
КОНСТРУИРАНЕ НА ИЗМЕРИТЕЛЕН ИНСТРУМЕНТАРИУМ

Конструирането на измерителен инструментариум за ОДИ става по определен процес, който се базира на специфична технология. Нейното систематично следване и преминаване през всеки етап е задължително условие за получаване на измерителен инструмент (тест) с „хубави" измерителни качества. Ето основните етапи:

- **Планиране и подготовка на общ дизайн на теста** – формулиране на целите и учебното съдържание, съставяне на тестовите спецификации, определяне на броя и формàта на задачите.
- **Съставяне на тестови задачи** – съставяне на тестовите задачи, подреждането им за апробация, подготовка на инструкции за изпълнението на теста.
- **Предварително изследване** – проверка на валидността, корекция на тестовите спецификации и съответните задачи.
- **Апробация** – избор на извадка от ученици, които са колкото е възможно поеднотипни с учениците, за които е предназначено основното измерване. Провеждане на предварителния тест (експериментално тестиране).
- **Изследване на качествата на задачите и теста** – изследване на качествата на отделните задачи, оценяване на качествата на теста като цяло – надеждност, валидност.
- **Окончателно оформяне на теста** – подбор на задачите за основния тест, корекции на инструкциите, задачите, външно оформление, подготвяне на помощни материали.
- **Определяне на оценъчните скали** – изработване на инструкция за оценяването.

Таблица 9 представя дейностите при създаване на тест, приблизителното време за изпълнението им и хората, които би трябвало да отговарят за тях. Както се вижда от таблицата създаването на един тест за ОДИ изисква от 50 до 62 седмици (около една календарна година) екипна работа.

Таблица 9: Основни етапи в създаването на тест

Етап	**Приблизително време**	**Изпълнител**
Изясняване на целта на теста	1 седмица	Образователни политици, спонсорите, координаторът на екипа за създаване на тест
Написване на рамката и спецификациите на теста. Консултации за широкото им одобрение	4 до 6 седмици	Координаторът на екипа за създаване на тест, експертни групи, опитни учители, специалисти по предмета на теста, статистици, автори на тестови задачи, образователни политици, спонсори
Писане на тестови задачи	12 до 14 седмици със скорост по 20-30 задачи на автор за седмица	Координаторът на екипа за създаване на тест, експертни групи, автори на тестови задачи, спонсори
Композиране на теста за апробация	4 седмици	Координаторът на екипа за създаване на тест, автори на тестови задачи, специалисти по предпечатна подготовка, редактори
Отпечатване на теста за апробация	2 седмици	Екипът за създаване на тест, автори на тестови задачи
Пакетиране и разпространяване на теста за апробация	2 до3 седмици	Екипът за създаване на тест
Администриране на теста за апробация и връщане на материалите	2 до 3 седмици	Администраторите на теста в училищата и Екипът за създаване на тест
Проверка на задачите със свободен отговор (ако е нужно) от апробацията	1 седмица	Координаторът на екипа за създаване на тест и авторите на задачи
Въвеждане на данните от апробацията	1 седмица	Статистик и технически персонал
Анализ на данните от апробацията	2 седмици	Статистик, авторите на задачите, координаторът на екипа за създаване на тест
Подбор на задачи за основното тестване	2 седмици	Координаторът на екипа за създаване на тест, автори на задачи, спонсори
Композиране на крайния тест	2 седмици	Координаторът на екипа за създаване на тест, автори на тестови задачи, специалисти по предпечатна подготовка, редактори
Отпечатване на крайния тест	2 седмици	Екипът за създаване на тест, автори на тестови задачи
Пакетиране и разпространяване на крайния тест	2 до3 седмици	Екипът за създаване на тест
Администриране на крайния тест и връщане на материалите	2 до 3 седмици	Администраторите на теста в училищата и Екипът за създаване на тест
Проверка на задачите със свободен отговор (ако е нужно) теста	2 до 3 седмици	Координаторът на екипа за създаване на тест и авторите на задачи
Въвеждане на данните	2 до 3 седмици	Статистик и технически персонал
Анализ на данните	2 до 3 седмици	Статистик, авторите на задачите, координаторът на екипа за създаване на тест
Написване на доклади и отчети	3 до 4 седмици	Статистик, автори на тестови задачи, координаторът на екипа за създаване на теста

По-долу са описани някои от основните стъпки при създаване на измерителен инструментариум, които не са обект на други глави в този труд.

1. Определяне целите на теста и измервания конструкт

Всеки тест се създава с ясно и точно определена цел. Тя трябва да бъде заявена в самото начало на процеса за създаване на тест. Обикновено, тестът има една основна цел. Например, тест за правилно и бързо извършване на аритметични действия може да има различни цели: (1) да диагностицира ученици в определен клас (например четвърти) като определи онези, които имат затруднения с тази дейност и какви по-точно са тези затруднения; (2) да бъде част от процедура за подбор на касиер в банка.

Още в самото начало на планиране на теста трябва да се определи основната му цел. Това е важно за всяка от следващите стъпки в подготовката му. Типични цели на тестове за измерване на постижения на ученици в са: диагностика на знанията в обучението на многобройни популации; диагностика на отделни ученици, показващи затруднения при обучението; постижимост на образователни изисквания; проверка на поставени изисквания за входно или изходно ниво; поставяне на оценка; класиране; др.

Резултатите от теста и тяхната интерпретация трябва да се използват единствено и само за целите, за които е създаден теста. Използването на резултати от тест за цели, които не са предвидени в неговото конструиране, обикновено води до неадекватни тълкувания, достигане до обърквaщи или противоречиви изводи и в крайна сметка до вземане на неправилни решения.

Целите на теста, както и използването на резултатите, могат да са от различно естество. Thorndike (1982) ги систематизира в следните основни видове:

1. *Цели, свързани с обучението* – основният въпрос е дали отделни ученици, или групи ученици, са изучили в достатъчна степен определен учебен материал и имат (или нямат) готовност да преминат на следващ учебен материал.
2. *Диагностични цели* – задачата е да се установи наличието на затруднения при изучаване на определен учебен материал, за да се вземат своевременно мерки за неговото усвояване.

3. *Селективни цели* – свързани са с класиране за прием в учебно заведение или за други нужди, в които е нужно да се направи подбор на кандидати.
4. *Цели, свързани с позиционирането* – задачата е да се определи правилното място на отделни ученици или на групи ученици в рамката на изучаване на даден предмет. Например, когато се подбират курсисти в курсове за изучаване на чужд език, добре е всеки курсист да попадне в група ученици, които имат почти еднакви знания за езика. За различните групи знанията може да са различни (начинаещи, малко напреднали, напреднали и т.н.)
5. *Насочващи цели* – свързани с вземане на решения за бъдещето на учениците, като например към какво профилирано училище да се насочат, каква бъдеща професия е подходяща за тях и др.

Определянето на измервания конструкт е важно, за да може този конструкт да се прояви по подходящ начин с тестови задачи. Така, концептуалната идея за конструкта се операционализира в конкретни задачи. Трудностите в ясното определяне на конструкта идват от неговия абстрактен характер. Колкото по-конкретно е описан конструкта, толкова по-лесно е да се съставят задачи, които го измерват. По-долу са дадени няколко примерни описания на конструкти.

- Езикови умения – познаване значението на голямо количество думи/изрази и използване на най-подходящите думи/изрази в даден контекст.
- Граматични умения – познаване граматиката на езика и правилното й използване при писане на текст.
- Умения за извършване на аритметични действия – изразяват се в бързо и точно извършване на „прости” аритметични действия.
- Умения за преобразуване на алгебрични изрази – изразява се в познаване на правилата и формулите за преобразуване на алгебрични изрази и прилагането им в дадени ситуации.

Такива описания на конструкти служат за поне две основни цели: (1) дават допълнително обяснение за това, какво се има пред вид под съответния конструкт; (2) подсказват вида на тестовите задачи, с които може да се оцени конструкта.

2. Рамка и тестови спецификации

Тестовите спецификации са писмено формулиране на идеята за теста като цяло: какво ще измерва, каква ще е структурата му, кои ще са основните му компоненти и др. Това е първият документ, в създаването на всеки тест. Тестовите спецификации са основата, върху която се пишат тестовите задачи и се конструира цялостния тест. Добре направени, те са предпоставка за създаване на измерителен инструментариум (тест) с „хубави" измерителни качества.

Всеки тест се прави с някаква цел (виж предния параграф). Първата важна стъпка в създаването на теста е изясняване на целите, за които той е предназначен, както и за какво ще се използват данните и резултатите. Едновременно с това се дава и отговор на въпроси като: Защо е нужно планираното оценяване/измерване? Кого ще оценява/измерва теста (обекта на измерване)? Какво ще оценява/измерва теста (конкретизиране на целите)? Кога ще стане провеждането на теста и колко време ще отнеме то? На какъв език ще се проведе теста (ако популацията е многоезична)? Какъв формат задачи ще съдържа теста? По какви параметри ще бъдат класифицирани задачите в теста, за да се направи смислен анализ на резултатите? Отговорите на тези и други подобни въпроси се описват в така наречените тестови спецификации.

След като е уточнено какво ще измерва тестът, нужно е да се направи подробно описание на съдържателните области. Всеки елемент от съдържанието на теста се описва подробно и ясно, за да няма двусмислие в съдържанието, което е обект на оценяването. Нещо повече, трябва да се балансират отделните съдържателни единици като се посочи пропорционалното им съотношение в целия тест. Добре е да се посочи още приблизителния брой задачи, които ще има в теста по всяка от съдържателните области.

Подходящ ориентир за такова пропорционално разпределение може да е съответното пропорционално разпределение на времето за изучаване на отделните съдържателни единици според учебния план. Например, ако в учебния план се предвиждат 4 часа за изучаване на тема А и 6 учебни часа за тема Б, добре е в теста отношението на броя задачи от тема А към броя задачи от тема Б да е приблизително $2:3$. Ако е възможно да се направи добра оценка за времето за решаване на всяка задача, може да се иска пропорционалност по отношение

на времето за решаване на задачите от всяка съдържателна област, а не на броя задачи.

Ако тази пропорционалност не се планира в самото начало, може да се получи не балансиран тест. Това означава, че в него има повече задачи по учебно съдържание, което няма особена тежест според учебния план, както и обратно, „важно" според учебната програма съдържание да се оценява със сравнително малко на брой тестови задачи.

Освен съдържателна област, всяка тестова задача се класифицира и по друг параметър, който се нарича познавателна (когнитивна) област. Той е свързан с познавателните процеси и/или дейности, с които трябва да се ангажира ученикът, за да реши задачата. Ще обясним това с един пример. Разгледайте следните две описания на една и съща съдържателна област от равнинната геометрия: (1) *Познаване на основните понятия, свързани с окръжността (радиус, диаметър, хорда, централен ъгъл, вписан ъгъл и др.);* (2) *Пресмятане на лица и обиколка на кръг и/или негови части, както и мерки на ъгли, свързани с окръжността.* Очевидно е, че първото описание изисква само разпознаване на изучени елементи (понятия), докато при второто се иска не само това, но и свързване на тези понятия в принципи и прилагане на процедури. Следователно, за да дадем достатъчно информация на авторите на тестовите задачи за целия тест е нужно не само да изброим съдържателните области, които оценяваме, но и познавателното ниво, в което искаме да ангажираме учениците, т.е. да опишем и познавателните области.

Първите сериозни опити за описание на познавателни области са публикувани през 1956 година от Блум (Bloom, 1956). Той се опитва да структурира познавателните равнища на обучението, започвайки от най-ниското (да се репродуцира дадена информация) до най-високото (да се приложи научено знание или придобито умение в качествено нова ситуация). Блум разглежда шест познавателни равнища: *познаване, разбиране, приложение, анализ, синтез, преценяване*. Тази класификация е известна като таксономия на Блум. Ето какво съдържание влага Блум във всяко равнище.

Познаване – припомняне/разпознаване на факти, поднесени по начин, който е подобен на познатото в училище. Например, разпознаване на столиците на някои държави, които са изучавани в училище.

Разбиране – интерпретация/обяснение на понятия в ситуация, която е малко по-различна от познатата с училище. Например, разпознаване на съществителни имена в изречения, които не са използвани като примери в клас; посочване на прости числа, различни от тези, които са дадени в учебника.

Приложение – решаване на „нова" задача чрез използване на известни принципи или обобщения. Например, изчисляване на лице на триъгълник чрез използване на подходяща формула, без в условието на задачата да се споменава за тази формула (ученикът сам трябва да се досети да я използва).

Анализ – разделяне на проблема (задачата) на части/елементи, които имат познати връзки/принципи между тях. Например, идентифициране на биологичните родове на нов ботанически екземпляр чрез характеристиките на листата и цветовете му.

Синтез – комбиниране на елементи в едно цяло чрез използване на оригиналната му структура или решаване на задача, която изисква последователно комбиниране на различни принципи в нова (непозната за учениците) ситуация. Например, написване на компютърна програма за изчисления с голямо количество данни, в която да има условия за вход, изход и т.н.

Преценяване – прилагане на вътрешен (самостоятелно изработен) или външен критерий за критична преценка по отношение на точност, прецизност, логическа съвместимост, или артистичност, или философска гледна точка и др. Например, написване на критична статия за публикация, описваща емпирично социологично изследване.

Таксономията на Блум не е свързана с конкретен предмет. Тя се отнася за произволна интелектуална дейност. След него са правени редица допълнения, редакции, промени и т.н. в таксономията, които са съобразени със спецификите на съответните предмети или предметни области. Специално за целите на тестовете, се оказва, че не са нужни толкова много познавателни равнища. За тестовата практика се използват от 2 до 4 познавателни равнища (области). Когато целите и обекта на измерване са определени, колективът, който подготвя теста, изготвя по собствена преценка и познавателните области (равнища) като дава подробно и ясно описание за всяко от тях.

Поради голямото разнообразие в системите за познавателни равнища, нужно е за всеки тест да се направи описание на познавателните области, които ще бъдат застъпени в него. Както и при съдържателните области, така и тук,

трябва да се определи пропорционалното разпределение на задачите по формулираните познавателни области.

Така, класификацията на задачите за даден тест обикновено става в две размерности: съдържателна и познавателна. Описанието на областите в тези две размерности е удобно да стане в таблица, която понякога се нарича съдържателна рамка на теста (или таблица от спецификациите).

В таблица 10 е представена примерна съдържателна рамка по предмета „Човекът и природата", IV клас (по разработка на Людмила Зафирова от Факултета по начална и предучилищна подготовка на СУ „Св. Климент Охридски" за национално оценяване в IV клас ръководено от автора на дисертационния труд). В нея са показани тежестите (в проценти) на всяка от съдържателните области и на техните подобласти. Сборът от процентите по последните три стълба дава относителната тежест в проценти на трите познавателни области, които са заложени в теста. Разбира се, сборът от процентите в последните три стълба трябва да дава 100%.

Таблица 10. Примерна съдържателна рамка (в проценти и брой задачи) за тест по „Човекът и природата", IV клас

Основни съдържателни области	**Съдържание**	**Познавателни области**		
		Познаване	Разбиране	Приложение
I. Вещества, тела и организми 35% (14 задачи)	1. Различава вещества по свойства и употреба.		7% (3)	
	2. Разграничава по основни признаци неживите тела от живите организми.		3% (1)	
	3. Класифицира живите организми според средата им на живот.		5% (2)	
	4. Посочва приспособленията на живите организми за живот в дадена среда.	5% (2)		
	5. Групира растенията на дървета, храсти, треви.		5% (2)	
	6. Групира животните на бозайници, птици, влечуги, риби и насекоми.		5% (2)	
	7. Разпознава Земята като планета от Слънчевата система и Слънцето като основен източник на светлина и топлина.	5% (2)		

Основни съдържателни области	**Съдържание**	**Познавателни области**		
		Познаване	Разбиране	Приложение
II. Природни явления и процеси 30% (12 задачи)	1. Назовава основни жизнени процеси в организмите.	3% (1)		
	2. Илюстрира с примери необходимостта от енергия за организмите, бита и промишлеността.		5% (2)	
	3. Проследява измененията в агрегатното състояние и кръговрата на водата в природата (по схема).	7% (3)		
	4. Разпознава движението на тела по дадено описание или картина.		5% (2)	
	5. Разбира последиците от движението на Земята около оста й и около Слънцето.		5% (2)	
	6. Разпознава измененията в природата и човешката дейност през различните сезони.	5% (2)		
III. Човекът и неговото здраве 25% (10 задачи)	1. Назовава органи в човешкото тяло (по схема).			5% (2)
	2. Обяснява как протичат основните жизнени процеси при човека.		6% (2)	
	3. Изброява хигиенни правила за здраво тяло.	7% (3)		
	4. Назовава дейности на човека, водещи до нарушаване на равновесието в природата.	4% (2)		
	5. Разпознава вредни за човека вещества и въздействия и основни замърсители на околната среда.		3% (1)	
IV. Наблюдения, експерименти и изследване 10% (4 задачи)	1. Коментира последиците от извършването на дадени наблюдения или опити с тела, вещества и растения.		5% (2)	5% (2)
Общо		36% (15)	54% (21)	10% (4)

Когато се фиксира броят на задачите в целия тест, може да се напише същата таблица, като на мястото на процентите се постави броя на задачите, пресметнати според предвидените проценти. Например, при общо 40 задачи в теста, съответният брой е написан в скоби след всеки процент. Броят на задачите във всяка клетка от таблицата се закръглява до цяло число.

3. Създаване на банка от тестови задачи и тяхната експертна оценка

В глава 5 е посветена на писането на тестови задачи. Тук ще отбележим само някои моменти, касаещи създаването на измерителния инструментариум.

Тестовите задачи се пишат въз основа на тестовите спецификации. Всяка тестова задача трябва да е свързана с точно една определена съдържателна и познавателна област от съдържателната рамка. Посочването към коя точно от тези области принадлежи задачата е работа на нейния автор. За улеснение, писането на тестовите задачи става на специално подготвена бланка (фиг. 62), която се попълва от автора на задачата. За задачите със свободен отговор авторът предава и съответното ръководство за оценяване. В последствие то може да претърпи промени, но съставянето му е част от написването на задачата.

Математика	Съдържателна област No:	Познавателна област No:	Очакван процент верни отговори
Задача – условие и възможности за отговор (ако има такива)			
Правилен отговор_____ (за задача с избираем отговор) Решение (за задача със свободен отговор): Обяснения за дадените дистрактори (ако има такива)			
Предложил: _____ Подпис: _____ Дата: _____			
Служебен номер на задачата (НЕ СЕ попълва от този, който предлага задачата.)			

Фиг. 62. Примерна бланка за тестова задача

Множеството от всички съставени задачи за даден тест се нарича банка от тестови задачи. Броят на задачите в банката трябва да е от порядъка на около 3 пъти повече от броя на задачите, които се очаква да съдържа теста. Това е така, защото практиката показва, че след апробацията (виж следващия параграф) само около една-трета от задачите показват сравнително добри измерителни качества, за да бъдат включени в окончателния вариант на теста.

Написаните задачи се дават за експертна оценка на назначен за целта експертен съвет. Този съвет се състои от няколко души, експерти в предметната област, разбиращи от обучението в областта, с опит в преподаването, оценяването и владеенето на езика. Целта на експертната оценка е чрез градивна критика да се повиши качеството на задачите. Експертният съвет преценява доколко всяка задача е подходяща да оцени конструкта, като вземат отношение по следните основни въпроси:

- Спазени ли са на основните правила за писне на тестови задачи?
- Взети ли са пред вид всички възможни начини за решаване на задачата?
- Взети ли са пред вид често допускани грешки при решаване на задачата?
- Правилна ли е класификацията по съдържателни области?
- Правилна ли е класификацията по познавателни области?
- Има ли задачата единствен обективен правилен отговор? (За задачи с избираем отговор)

Задача No:			
Характеристики на задачата	Да	Не	Предложение за корекции *
Ясно ли е поставено условието? (Няма двусмислие.)			
Правилна ли е класификацията по съдържание?			
Правилна ли е класификацията по познавателно равнище?			
Задачата има ли правилен отговор?			
Единствен ли е правилният отговор? В последната колона напишете правилния отговор.			
Добре ли са подбрани дистракторите?			
Логически правилно ли са подредени дистракторите?			
Правилно ли са направени чертежите и/или диаграмите?			

* При липса на място, напишете предложенията за корекции на отделен лист.

Фиг. 63. Примерна бланка за експертна оценка на тестова задача

- Добре ли са подбрани задачите? (За задачи с избираем отговор)
- Смислено ли е ръководството за оценяване? (За задачи със свободен отговор)
- Може ли задачата да се подобри редакционно?
- Може ли задачата да се подобри езиково?

Желателно е експертите да попълват бланка (примерна такава е дадена на фиг. 63) за всяка задача, за да подредят информацията в нагледен вид.

Понякога при експертната оценка се налага да се проверява съгласуваността на експертите. Следващият пример показва как се извършва такова съгласуване за класификация на задачите по съдържателна област.

Нека в спецификациите на теста са предвидени 4 съдържателни области с условни имена „Тема 1”, „Тема 2”, „Тема 3” и „Тема 4”. Разполагаме с авторската класификация на задачите по тези области. Всеки от трима експерти E1, E2 и E3 дава индикация дали е съгласен или не с класификацията на автора, като използва следната тристепенна скала: 1 – задачата е правилно класифицирана; 2 – не съм сигурен; 3 – задачата е неправилно класифицирана. Част от резултатите са показани в таблица 11.

Таблица 11. Експертна оценка за класификация на задача

Задача No	Класификация от автора	Мнение на експертите		
		E1	E2	E3
...	...	...	...	...
82	Тема 1	3	3	3
83	Тема 2	1	1	1
84	Тема 4	1	2	1
85	Тема 4	2	3	2
86	Тема 3	1	1	1
...	...	...	...	...

Ясно е, че експертите имат отлична съгласуваност по задачи 82, 83 и 86. Задача 84 вероятно не е правилно класифицирана от автора. За задача 85, обаче, трябва да се постигне консенсус. Това обикновено става чрез разговори между експертите, които понякога се налага да чуят и съображенията на автора.

4. Апробация на тестови задачи

Апробацията е дейност, при която тестовите задачи се дават за решаване от ученици, които са колкото е възможно по-еднотипни с учениците, за които е предвиден основният тест. Целта на апробацията е да се изследват измерителните качества на написаните тестови задачи (трудност, разграничителна сила, ефективност на дистракторите), за да може от тях да се подберат за основния тест онези с възможно най-добри качества. От друга страна, ако поведението на популацията (учениците, за които е предвиден основният тест) по отношение на измервания конструкт не е познато от предишни измервания, апробацията дава възможност да се „изучи" това поведение. В този случай може да се каже, че апробацията дава информация до колко тестовите задачи и учениците, които трябва да ги решават, са „подходящи" един за друг.

Апробацията не трябва да се използва, за да се правят изводи за постиженията на учениците. Апробацията е тест за тестовите задачи (тест за самия тест), затова основното внимание от нейните резултати е насочено към свойствата на задачите, а не към резултатите на учениците. В този смисъл, апробацията има значение за „оразмеряване" на измерителния инструмент (теста).

Важен момент за апробацията е тя да се проведе с ученици, които са колкото е възможно по-еднотипни с учениците, за които е предвиден основният тест. Например, ако тестът е предназначен за измерване на ученици в края на четвърти клас, подобни на тези ученици имаме една година по-рано – тези, които тогава са в четвърти клас.

Друг важен въпрос е броят на учениците, които трябва да участват в апробацията. Това в известна степен зависи от общия дизайн на теста и от методите, които ще се използват за скалиране и анализ на резултатите. За средствата, използвани в така наречената класическа теория на тестовете, достатъчно е всяка задача да бъде решена от около 150 до 200 ученика по време на апробацията.

По време на апробацията могат да се пробват различни формати на една и съща задача. Например, една и съща задача може да се даде в някои тестови книжки като задача с избираем отговор, а в други – като задача с кратък свободен отговор (фиг. 64).

13 + 17 + 8 =

А) 28 Б) 30 В) 38 Г) 110

ИЛИ

13 + 17 + 8 = _____

Фиг. 64. Пример на една и съща задача в два различни формата

Може да се пробват и различни редакции на една и съща задача.

Изучаването на статистиката от резултатите е свързана най-вече с анализ на измерителните качества на задачите (трудност, разграничителна сила, ефективност на дистракторите). В резултат на това за всяка задача се получава информация дали нивото на трудност и разграничителната сила са подходящи за популацията и как работят дистракторите. Често, в резултат на този анализ, в задачите се откриват грешки, които по една или друга причина са останали незабелязани до този момент. Например, причини някоя задача да има „лоши" показатели за трудност, разграничителна сила и ефективност на дистрактори може да са: (1) грешно подаден ключ (правилен отговор); (2) липса на обективно правилен отговор; (2) наличие на повече от един правилен отговор; (3) неясно зададено условие (двусмислие и др.); (4) неясни възможности за отговор; (5) лошо подбрани дистрактори.

Важно е за всяка задача да се анализира процентът на неотговорилите ученици. Това също дава информация за някои качества на задачите и на теста като цяло. Например:

- Ако повече от 15% от учениците не са отговорили на няколко от последните задачи в теста, пилотният тест е бил дълъг. За основното тестване той трябва да се съкрати.
- Ако повече от 15% от учениците не са отговорили на задача, която не е в края на теста, вероятно те не са разбрали задачата, или тя е много трудна за тях. Така или иначе, учениците са имали затруднения с тази задача и е добре тя да се ревизира внимателно или да не се използва в основното тестване.

- Ако определена група ученици (например повече от 15% от момичетата) не са отговорили на някоя задача, а останалите ученици са давали отговор, може би задачата не е „равнопоставена” за всички ученици. Може би тя не трябва да се включва в теста за основното изследване.
- Ако повече от 15% от учениците не са отговорили на определен формат задачи (например такива с разширен свободен отговор), те вероятно не знаят (не разбират) как да отговарят на такъв формат въпроси. Може да се търси възможност за тренировка на учениците по този формат задачи преди да се проведе основното тестване. Или да не се използва този формат задачи.

Практиката показва, че след апробация около една-трета от написаните задачи имат сравнително добри измерителни качества, за да бъдат включени в основния тест. Останалите се нуждаят от по-малка или по-голяма корекция. Затова, препоръката е да се апробират около три пъти повече задачи, от предвидени брой за основния тест.

5. Конструиране на окончателен вариант на теста

Най-общо казано, окончателният вариант на теста се конструира като се използва съдържателната му рамка и се подбират задачи с възможно най-добри измерителни качества, показани на апробацията. Това означава, че в този процес са важни две неща: (1) спецификациите на теста, по-точно, спазването на съотношението между съдържателните и познавателни области в теста, така, както е зададено в съдържателната рамка; (2) измерителните качества на задачите, получени от апробацията.

Една задача не трябва да бъде избирана за основния тест, ако:

- тя е извън рамката и спецификациите на теста;
- има лоши статистически данни от апробацията (лоши измерителни качества);
- концептуално безсмислена.

Ето някои от най-важните критерии за подбор на задачи за основния тест:

- Трудността на задачата (процентът ученици вярно решили задачата) да е между 20% и 80%.

- Процентът на учениците, които не са отговорили на задачата, е малък.
- Разграничителната сила (точковата бисериална корелация) на задачата е по-голяма от 0,2.
- Надеждността на теста се увеличава, ако задачата се включи в теста (и намалява, ако тя се изключи от него).
- Задачата показва добра „равнопоставеност" на големите подгрупи в популацията (например момичета–момчета).
- За задачи с избираем отговор, точковата бисериална бисериална корелация на верния отговор да е положителна и по-голяма от 0,2, а тази на дистракторите да е отрицателна.
- За задачи с избираем отговор, всички дистрактори да са „ефективни", т. е. да са избрани от поне 5% от учениците.

Фигура 65 показва типичен пример за необходимите данни за една задача с избираем отговор, които трябва да имаме след апробацията.

Categories	**А [0]**	**Б [0]**	**В [0]**	**Г [1]**
Count	90	14	20	254
Percent	23.7	3.7	5.5	67.0
Pt-Biserial	–0.26	–0.21	–0.16	0.39
Mean Ability	–0.02	–0.48	–0.14	0.54

Фиг. 65. Пример за анализ на данни за задача с избираем отговор

Първият ред показва възможностите за отговор, като Г е правилният отговор (1 точка). Вторият ред съдържа броят ученици, дали като отговор съответната възможност. Третият ред показва този брой като процент от всички ученици, работили по задачата. Числото 3,7% в този ред показва, че дистракторът Б) е „слаб", т.е. не е „привлекателен" за учениците и може да се помисли за неговата смяна. Четвъртият ред съдържа точковите бисериални корелации (мярка за разграничителната сила) на възможностите за отговор. Виждаме, че те

са отрицателни за всички дистрактори и тя е положителна само за правилния отговор. Последният ред съдържа средната стойност на способностите за учениците, избрали съответната възможност за отговор. Виждаме, че тя е сравнително висока за онези, които са избрали правилния отговор и е ниска за онези, които са отговорили грешно. Това говори, че задачата „работи” добре.

Ето някои критерии за задачи, които не са с избираем отговор:

- Ако задачата се оценява с „вярно–невярно”, разграничителната сила трябва да е по-голяма от 0,2.
- Ако задачата допуска частично правилни отговори, разграничителната сила трябва да е положителна и по-голяма от 0,3.
- Ако задачата се оценява с „вярно–невярно”, процентът на вярно отговорилите трябва да е между 20% и 80%.
- Ако задачата допуска частично правилни отговори, във всяка категория трябва да имаме не по-малко от 5% от отговорите на учениците.
- Ако задачата допуска частично правилни отговори, средната стойност на способностите на учениците трябва да намалява от верните отговори към намаляващите по степен частично правилни отговори.

Фигура 66 показва типичен пример за необходимите данни за една задача с избираем отговор, които трябва да имаме след апробацията.

Categories	**0 [0]**	**1 [1]**	**2 [2]**	**9 [0]**	**Disc = 0.47**
Count	1466	425	268	809	
Percent	49.4	14.3	9.0	27.3	
Pt-Biserial	0.09	0.11	0.45	–0.48	
Mean Ability	–1.66	–1.53	–0.90	–1.90	

Фиг. 66. Пример за анализ на данни за задача със свободен отговор, допускаща частично правилен отговор

Първият ред показва възможните категории: 0 за грешен отговор, 1 за частично правилен отговор, 2 за напълно правилен отговор, 9 за „празен отговор”. Общата разграничителна сила е 0,47, която е висока. Както се вижда, тази разграничителна сила е различна от точковата бисериана корелация на напълно

правилния отговор (0,45). Следващите редове са аналогични на тези от фигура 20. Точковата бисериална корелация нараства с нарастване на категориите 0, 1 и 2. Същото важи и за средната стойност на способностите. Тези данни говорят за добри измерителни качества на тази задача. Процентът на учениците, които не са отговорили на задачата, е 27,3. Това число трябва да се анализира в контекста на цялостния тест (дали задачата е в края на теста и др. подобни фактори, някои от които бяха обсъдени по-горе). Трябва да се има пред вид, че за задачите със свободен отговор процентът на неотговорилите често пъти надхвърля 20%.

Следователно, измерителните качества на задачите от апробацията са нещо важно, но те не трябва да се абсолютизират. Погрешна е постановката, че окончателният вариант на теста може да се направи без участието на специалисти по предметната област на теста. Ако се подбират само задачи с най-добри измерителни качества, без да се съобразяваме с тяхната съдържателна и познавателна класификация, може да се получи тест, който много се различава от онова, което предвиждат спецификациите (тест с лоша валидност). Ето защо участието на специалисти по предметната област на теста е задължително.

Сериозен проблем понякога е, че в някои от съдържателните (или познавателните) области на теста може да няма достатъчно задачи с добри измерителни качества. Тогава се правят „компромиси”, като: (1) се включват задачи с „не особено добри” измерителни качества; или (2) се правят „малки” корекции на задачи, за които се предполага, че ще подобрят измерителните качества.

Няма точно обяснение на това колко „малки” корекции са допустими. Всяка коригирана задача, колкото и малка да е корекцията, е друга задача, с други измерителни качества. Затова всяка коригирана задача би трябвало да се апробира отново, за да се получат качествата на новата задача. На практика, често няма време и ресурси, за да се направи повторна апробация. Ето защо, идеалният случай е да не се правят промени. Но на практика правилото е, че колкото по-малко промени се правят, толкова по-добре. Промените могат да са като например:

- Промяна на няколко думи в текста на задачата, за да стане тя по-ясна или да се използват по-познати думи.
- Смяна на дистрактор, който „не работи добре”.

- Граматични и стилистични корекции.
- Подобряване на оформлението (по-добри чертежи, правилно разположение на фигурите и др.)

Следният пример показва как „малко” е коригирана една задача след апробацията, за да бъде включена в основния тест.

Задача (както е дадена за апробация). Нека *S* е лицето на кръг с радиус *r*, *P* е обиколката на неговата окръжност и $R = \frac{S}{P}$. Когато *r* нараства, кое от следните твърдения е вярно за *R*?

А) *R* намалява

Б) *R* е константа

В) *R* расте като линейна функция

Г) *R* расте като квадратна функция

Д) *R* расте като степенна функция

Таблицата показва някои от измерителните качества на задачата (правилният отговор е В).

Възможности за отговор	А	Б	В*	Г	Д	Без отговор
Процент отговорили	15,4	0,9	65,4	9,6	5,8	2,9
Точкова бисериана корелация	–0,13	0,31	0,44	–0,15	–0,32	–0,23

Прави впечатление, че дистрактора Б е избран от сравнително малко ученици и има положителна разделителна сила. Това означава, че той не е подходящ дистрактор и е по-добре да се махне. Освен това условието на задачата може да се редактира, като се изясни кое е аргументът на разглежданата функция. Така, за основния тест се предлага следния вариант на задачата:

Задача (предложена за основния тест). Нека *S* е лицето на кръг с радиус *r*, *P* е обиколката на неговата окръжност и $R = \frac{S}{P}$ е функция на *r*. Кое от следните твърдения е вярно за *R*?

А) *R* намалява

Б) *R* расте като линейна функция

В) *R* расте като квадратна функция

Г) *R* расте като степенна функция

Ето някои от качествата на задачата, които са получени при основното тестване (правилният отговор е Б).

Възможности за отговор	А	Б*	В	Г	Без отговор
Процент отговорили	13,9	59,7	9,3	7,6	9,5
Точкова бисериана корелация	–0,22	0,51	–0,09	–0,15	–0,37

Както се вижда, промените в параметрите на останалите възможности за отговор са добри, което потвърждава допускането, че направените корекции в задачата са смислени и подобряват нейното качество.

6. Подготовка на ръководства

В много случаи отделните етапи при провеждане на теста се изпълняват от екипи, които са различни от тези, които са създали теста. Тогава е необходимо авторите на теста да подготвят съответни ръководства.

Ръководство за провеждане на теста описва подробно процедурата за това, къде какво и как (в каква последователност и време) трябва да се направи в стаите, в които се провежда теста, от началото до края.

Ръководство за кодиране на задачите със свободен отговор съдържа всички необходими материали и указания за това кодиране.

Ръководство за събиране на материалите и данните описва процедурата за събиране и комплектоване на всички кодирани ученически отговори и тяхното нанасяне в компютър (получаване на електронен файл за обработка на данните).

Ако е необходимо и е възможно да се направи преди провеждане на теста, се подготвя *ръководство за превръщане на суровия тестов бал в други скали*.

В зависимост от конкретната ситуация, възможно е да има нужда и от други ръководства и помощни материали за хората, които провеждат теста и използват резултатите от него.

СЕДМА ГЛАВА
ИЗМЕРВАНЕ НА ПОСТИЖЕНИЯТА ПО МАТЕМАТИКА В TIMSS

Тук ще разгледаме концепцията за оценяване на математическите постижения на учениците в рамката на изследването TIMSS. Като член на експертния съвет на TIMSS, авторът на този труд е участвал в разработването на рамката и спецификациите на математическия тест на TIMSS-2003 (TIMSS Assessment Frameworks and Specifications 2003), на TIMSS-2007 (TIMSS 2007 Assessment Frameworks) и TIMSS-2011 (TIMSS 2007 Assessment Frameworks). Той е един от авторите на тези рамки, както и на не малка част от задачите, включени в математическия тест на TIMSS.

По своя замисъл TIMSS е международно широко мащабно ОДИ на ученици в две възрастови групи: завършващи IV клас и завършващи VIII клас. Целта е да се направи сравнителен анализ на постиженията на учениците от споменатите популации в изучаване на математиката и природните науки според изискванията на учебните планове. Става ясно, че е нужно да се обхване голямо количество тематичен материал. Понеже обектът на изследването е „всички завършващи" даден клас, познавателните умения би трябвало да се дефинират на нивото на „средно статистическия ученик". България участва в TIMSS само с популацията на завършващите VIII. По-долу ще говорим само за тази популация и само за математика.

1. Рамка и тестови спецификации на TIMSS

Броят на участващите държави в различните фази на TIMSS е различен и варира между 45 и 55. Ясно е, че учебните планове по математика също се различават. Затова първата стъпка в създаването на рамката на изследването е постигане на консенсус в съдържанието по математика, която ще се оценява. За целта беше направено мащабно и подробно изследване на учебните планове по математика в участващите държави. Методиката и резултатите от това изследване (Schmidt at al., 1997), макар и да са интересни сами по себе си, не са обект на този труд.

Най-лесният начин да се определи математическото съдържание е да се вземе сечението на онова, което се изучава до VIII клас във всички участващи държави. Този подход обаче стеснява много силно съдържанието. Затова беше

приета следната стратегия. Като се основава на данните от изследването на учебните планове, експертният съвет да състави тематична рамка за математическото съдържание, което ще се оценява в TIMSS. Тази тематична рамка да отговаря на съвременните виждания за онова математическо съдържание, което международната математическа колегия счита, че трябва да се знае от 14 годишните ученици по света.

Така съдържателната рамка на TIMSS сама по себе си става основа за развитие на национални планове за обучение по математика, тъй като тя е своеобразен образец за онова, което учениците трябва да изучат до завършването на VIII клас. Съдържателните области, които се оценяват в TIMSS, са дадени в Таблица 12.

В концепцията за създаване на съдържателната рамка на TIMSS е заложена идеята, че за всяка държава може да има математическо съдържание, което не се изучава от учениците до завършване на VIII клас. Оказва се, че за всичките фази на TIMSS това е вярно в по-малка или по-голяма степен за всяка една участваща държава.

IV клас		**VIII клас**	
Числа	• Естествени числа • Обикновени и десетични дроби • Числови изрази • Зависимости	**Числа**	• Естествени числа • Обикновени и десетични дроби • Цели числа • Отношение, пропорция и процент
		Алгебра	• Зависимости • Алгебрични изрази • Уравнения, формули, функции
Геометрични фигури и измерване	• Прави и ъгли • Двумерни и тримерни геометрични фигури • Геометрични преобразувания	**Геометрия**	• Геометрични фигури • Измерване в геометрията • Геометрични преобразувания
Представяне на данни	• Четене и интерпретиране • Организиране и представяне	**Данни и вероятности**	• Организиране и представяне на данни • Интерпретиране на данни • Вероятност

Таблица 12. Съдържателни области по математика за TIMSS-2007

Таблица 13 показва процентно разпределението на тестовото време по тематика и познавателни области за IV и за VIII за TIMSS-2007.

Таблица 13. Процентно разпределение на тестовото време за TIMSS-2007

Съдържание за IV клас	**Процент**	
Числа	50%	
Геометрични фигури и измерване	35%	
Представяне на данни	15%	
Съдържание за VIII клас	**Процент**	
Числа	30%	
Алгебра	30%	
Геометрия	20%	
Данни и вероятност	20%	
Познавателни области	**Процент**	
	IV клас	**VIII клас**
Знание	40%	35%
Приложение	40%	40%
Аргументация	20%	25%

2 . Тестов дизайн.

Изследването TIMSS използва три-параметричния вероятностен модел (IRT – виж четвърта глава). За измерване на ученическите постижения се конструират две скали – една по математика и една по природни науки. Тъй като съдържанието на всяка от тези области е много голямо, не е възможно един ученик да работи по всички задачи от теста. Затова в TIMSS 2007, например, задачите по математика, както и тези по природни науки, са групирани в по 14 блока (под-теста) – общо 28 блока. Блоковете с математически задачи са номерирани с M01–M14, а тези със задачи от природните науки – със S01–S14. С помощта на тези блокове са конструирани 14 различни тестови книжки, както е показано в таблица 14.

Общото тестово време за решаване на задачите от всяка книжка е 90 минути – по 45 минути за всяка от двете части. Всеки блок е оразмерен за 22,5 минути тестово време. Всеки ученик работи само по една тестова книжка, т.е. по два математически и два природонаучни блока. Всеки блок участва в точно две книжки. Така между всеки два математически блока може да се направи връзка, макар и чрез няколко книжки. (Например, за да намерим връзка между блоковете M01 и M06 преминаваме през книжките с номера 1, 2, 3, 4 и 5.)

Таблица 14. Тестов дизайн на TIMSS-2007.

Тестова книжка	**Първа част-блокове**		**Втора част-блокове**	
Книжка 1	M01	M02	S01	S02
Книжка 2	S02	S03	M02	M03
Книжка 3	M03	M04	S03	S04
Книжка 4	S04	S05	M04	M05
Книжка 5	M05	M06	S05	S06
Книжка 6	S06	S07	M06	M07
Книжка 7	M07	M08	S07	S08
Книжка 8	S08	S09	M08	M09
Книжка 9	M09	M10	S09	S10
Книжка 10	S10	S11	M10	M11
Книжка 11	M11	M12	S11	S12
Книжка 12	S12	S13	M12	M13
Книжка 13	M13	M14	S13	S14
Книжка 14	S14	S01	M14	M01

3. Математическите задачи в TIMSS

В TIMSS се използват следните формати на задачи: (1) задачи с избираем отговор; (2) задачи с кратък свободен отговор; (3) задачи с разширен свободен отговор.

Тук са разгледани някои от задачите на TIMSS, които са предлагани или съществено редактирани от автора на този труд. Задачите са групирани по съдържателна област, като са дадени кратки обяснения какво оценява всяка от тях.

Числа

M032704. Автобус пътува с постоянна скорост. Изминатият път е право пропорционален на времето, за което той е изминат. Ако автобусът изминава 120 km за 5 часа, колко километра ще измине той за 8 часа?

A. 168
B. 192
C. 200
D. 245

Тази задача е от подобласт „Отношение, пропорция, процент” и познавателна област Приложение. Тя оценява умението на учениците да решават задачи с думи, свързани с път и време, като се намира неизвестен член в пропорция. Без съмнение, това е важно умение за учениците от 8 клас.

M032525. Кое число, като се раздели на –6, се получава 12 като резултат?
A. –72
B. –2
C. 2
D. 72

Задачата е от подобласт „Цели числа” и познавателна област Знание. Тя оценява знанието за намиране на цяло число, което след деление с отрицателно цяло число, дава предварително даден резултат. Според българската програма това се изучава в по-долни класове, но се счита, че знанието е важно за учениците от 8 клас.

M042041. Работник отрязал $\frac{1}{5}$ от една тръба. Отрязаното парче е дълго 3 m. Колко метра е дължината на цялата тръба?

A. 8 m
B. 12 m
C. 15 m
D. 18 m

Задачата е от подобласт „Обикновени и десетични дроби” и познавателна област Приложение. Тя оценява умението за намиране на число по дадена негова част, изразена в дроб с числител 1. В българската програма по математика това също се изучава в по-долни класове, но се счита, че е важно умение за учениците от 8 клас.

M032381. Броят на децата в една екскурзия е по-голям от 55, но по-малък от 65. Децата могат да се разделят в групи от по 7, но не могат да се разделят в групи от по 8. Колко са децата в тази екскурзия?

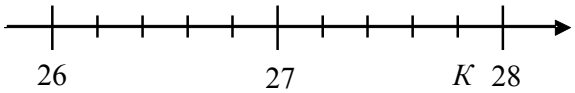
Отговор: _____

Ръководство за оценяване.

Код	Отговор	Задача No. M032381
	Правилен отговор	
10	63; 9x7; или 7x9	
	Неправилен отговор	
70	56; 8x7; или 7x8	
79	Всеки друг грешен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Естествени числа” и познавателна област Приложение. Тя оценява умението да се намери естествено число в даден интервал, което се дели на точно едно от две дадени естествени числа.

M042024.



Кое е числото, означено с *K* на числовата ос?

A. 27,4
B. 27,8
C. 27,9
D. 28,2

Задачата е от подобласт „Обикновени и десетични дроби” и познавателна област Знание. Тя оценява познаването на представянето на десетична дроб върху числовата ос между две последователни цели числа, когато само целите числа са означени. Това познание е от по-долните класова за България, но то се счита за важно за учениците от 8 клас.

M042197.

–3, 6, –12, 24, ...

Напишете правило, с което може да се намери който и да е член на тази редица, като се знае предходния член.

Правило:

Ръководство за оценяване.

Код	Отговор	Задача No. M042197
	Правилен отговор	
10	Умножаване с –2 на всеки член / Умножаване с 2 и смяна на знака / или друго еквивалентно правило	
	Неправилен отговор	
79	Всеки грешен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Това е задача от подобласт „Цели числа” и познавателна област Аргументация. Тя оценява намирането на правило за умножение с отрицателно

число за намиране на последователните членове на числова редица. За решаването на задачата не се изисква някакво специално знание, а вътрешна убеденост (аргументация), че написаното правило „работи” за всеки член на редицата.

M042031. Дробите $\frac{4}{14}$ и $\frac{\square}{21}$ са равни. На колко е равно □ ?

A. 6
B. 7
C. 11
D. 14

Задачата е от подобласт „Обикновени и десетични дроби” и познавателна област Приложение. Тя оценява прилагането за знанията за сравняване на дроби към избирането на числителя на дроб при даден знаменател, така че дробта да е равна на друга дроб с различен знаменател. Това е знание от по-долните класове в българското училище, но се счита, че приложението му е важно за осмокласниците.

M042039. Едно палто струва 60 зеда. Алан е купил палтото, когато цената му била намалена с 30%. Колко е спестил Алан?

A. 18 зеда
B. 24 зеда
C. 30 зеда
D. 42 зеда

Тази задача е от подобласт „Отношение, пропорция и процент” и от познавателна област Приложение. Тя оценява умението да се намали дадено количество с даден процент.

M032529. Едно палто струва 120 зеда. По време на разпродажба то се продава за 84 зеда. С какъв процент е намалена цената на палтото за разпродажбата?

A. 25
B. 30
C. 35
D. 36

Тази задача е от подобласт „Отношение, пропорция и процент” и от познавателна област Приложени. Тя оценява умението да се намери процента на

намаление, ако се знае началната и намалената цена. В известен смисъл задачата е обратна на предишната задача.

M022106. В един автобус има 36 пътници. Отношението на децата към възрастните на пътниците в автобуса е 5 към 4. Колко деца има в автобуса?

Отговор: _____

Ръководство за оценяване.

Код	Отговор		Задача No. M022106
	Правилен отговор		
10	20		
	Неправилен отговор		
70	9	[5+4 или 36:4]	
71	16	[броя на възрастните]	
72	5	[отношението на децата]	
73	27	[36-9]	
79	Всеки грешен отговор (включително зачертани, изтрити, неясни бележки и др. белези).		
	Без отговор		
99	Празно		

Задачата е от подобласт „Отношение, пропорция, процент” и от познавателна област „Приложение”. Тя измерва умението да се намери големината на една от частите при дадено цяло и отношението на двете части.

M032529. Едно палто струва 120 зеда. По време на разпродажба то се продава за 84 зеда. С какъв процент е намалена цената на палтото за разпродажбата?

A. 25
B. 30
C. 35
D. 36

Тази задача е от подобласт „Отношение, пропорция и процент” и от познавателна област Приложение. Тя оценява умението да се намери процента на намаление, ако се знае началната и намалената цена. В известен смисъл задачата е обратна на предишната задача.

M032725. Напишете $3\frac{5}{6}$ като десетична дроб, закръглена до втората цифра след десетичната запетая.

Отговор: _____

Ръководство за оценяване.

Код	Отговор	Задача No. M032725
	Правилен отговор	
10	3,83	
	Неправилен отговор	
70	3,56	
79	Всеки друг грешен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Обикновени и десетични дроби” и познавателна област Приложение. Тя оценява умението да се превръща смесена дроб в десетична със закръгляне до дадена цифра след десетичната запетая.

От прегледа на тези (а и други, не цитирани тук) задачи от областта Числа на TIMSS, могат да се направят някои изводи, касаещи преподаването на темата. За учениците от VIII клас се счита, че е важно да знаят и използват за решаване на задачи:

1. Права пропорционалност.

2. Намиране на част от число и число по дадена негова част.

3. Намаляване/увеличаване на дадено количество с даден процент; или намиране на процента, с който е увеличено/намалено дадено количество.

4. Намиране на неизвестен член в пропорция, която трябва да се състави според описана в задачата ситуация.

5. Осмисляне на понятието „обикновена дроб”.

6. Действия с рационални числа.

Изброените знания и умения естествено са основни за обучението по математика. Както се вижда, те са основни и за оценяването на постиженията на учениците от изследваната възрастова група.

Алгебра

M032540. $3(2x-1)+2x=21$
Каква е стойността на x?

A. -3

B. $-\frac{11}{4}$

C. $\frac{11}{4}$

D. 3

Тази задача е от подобласт „Уравнения, формули, функции” и от познавателна област Приложение. Тя оценява умението да се реши линейно уравнение с едно неизвестно.

M032738. Какво означава $xy+1$?

A. Прибавяне на 1 към y и след това умножаване с x
B. Умножаване на x и y с 1
C. Прибавяне на y към x и след това прибавяне на 1
D. Умножаване на x с y и след това прибавяне на 1

Задачата е от подобласт „Алгебрични изрази” и от познавателна област Знание. Тя измерва знанието на смисъла на алгебричен израз, в който има умножаване и събиране.

M042100. Коя двойка числа (x, y) удовлетворява уравнението $3x+4y=24$?

A. (0, 8)
B. (3, 4)
C. (4, 3)
D. (6, 0)

Задачата е от подобласт „Уравнения, формули, функции” и от познавателна област Знание. Тя измерва знанието за това какво означава двойка числа да удовлетворява линейно уравнение с две неизвестни.

M032698. Якетата на Христо са с 3 повече от якетата на Ана. Ако Христо има n якета, Колко са якетата на Ана, изразени чрез n?

A. $n-3$
B. $n+3$
C. $3-n$
D. $3n$

Задачата е от подобласт „Алгебрични изрази” и от познавателна област Приложение. Тя измерва умението да се открие не сложен алгебричен израз, който моделира дадена ситуация.

M042235. $x+y=12$ и $2x+5y=36$.
Кои са стойностите на x и y?

A. $x=2,\ y=10$
B. $x=4,\ y=8$
C. $x=6,\ y=6$
D. $x=8,\ y=4$

Тази задача е от подобласт „Уравнения, формули, функции” и от познавателна област Приложение. Тя измерва умението да се реши не сложна система от две линейни уравнения с две неизвестни.

M042199. Кой от изразите е еквивалентен на $4x-x+7y-2y$?

A. 9
B. $9xy$
C. $4+5y$
D. $3x+5y$

Задачата е от подобласт „Алгебрични изрази” и от познавателна област Приложение. Тя измерва умението да се открие нормалния вид на алгебричен израз с две променливи.

M042238. Коя точка лежи на правата $y=x+2$?

A. (0, –2)
B. (2, –4)
C. (4, 6)
D. (6, 4)

Тази задача е от подобласт „Уравнения, формули, функции” и от познавателна област Знание. Тя измерва знанието за това кога една точка, зададена с координати, лежи на права, зададена с уравнение, в равнината.

M032047. На колко е равен сборът на три последователни цели числа, ако средното от тях е $2n$?

A. $6n+3$
B. $6n$
C. $6n-1$
D. $6n-3$

Задачата е от подобласт „Алгебрични изрази” и от познавателна област Приложение. Тя измерва умението да се намери нормалния вид на алгебричен израз на сбор от три последователни цели числа, изразен с една променлива.

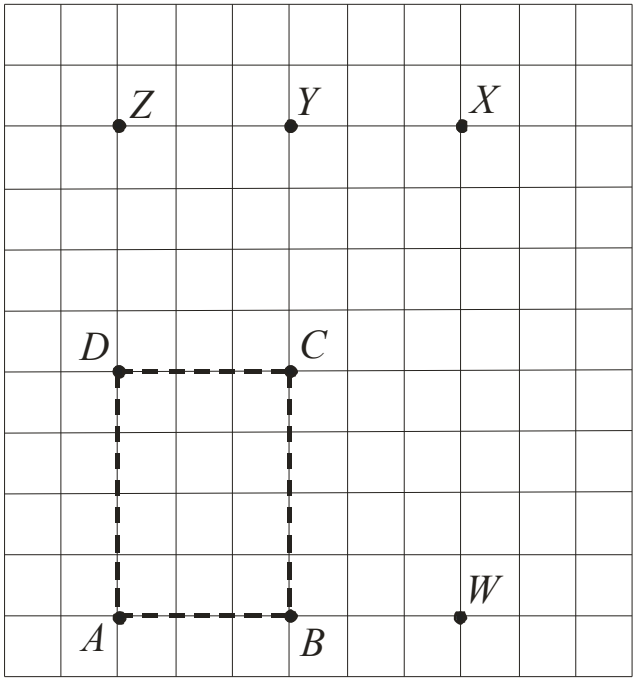
От прегледа на тези (а и други, не цитирани тук) задачи от областта Алгебра на TIMSS, могат да се направят някои изводи, касаещи преподаването на темата. За учениците от VIII клас се счита, че е важно да знаят и използват за решаване на задачи:

1. Смисъла на алгебричните изрази.
2. Преобразуване на алгебрични изрази.
3. Линейни уравнения с едно и с две неизвестни.
4. Решаване на системи от две линейни уравнения с две неизвестни.
5. Моделиране на „прости” ситуации с линейни уравнения.

Изброените знания и умения естествено са основни за обучението по математика. Както се вижда, те са основни и за оценяването на постиженията на учениците от изследваната възрастова група.

Геометрия

M042130. Като използвате отбелязаните точки, начертайте триъгълник с лице ДВА ПЪТИ по-голямо от лицето на правоъгълника *ABCD*.

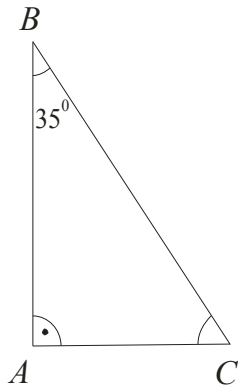


Ръководство за оценяване.

Код	Отговор	Задача No. M042130
	Правилен отговор	
10	Като се използват отбелязаните точки, е начертан триъгълник с лице 24 квадратни единици. Например: *AZW, ZWX, XAW, XZA, AYW, BZX, XWD*.	
	Неправилен отговор	
70	Триъгълник с лице 12 квадратни единици.	
79	Всеки друг грешен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Измерване в геометрията” и познавателна област Приложение. Тя оценява умението да се начертае триъгълник с лице, което е два пъти лицето на даден правоъгълник.

M032579.

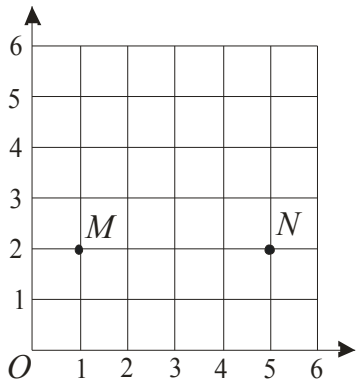


Колко е мярката на ъгъл *C* в триъгълника от чертежа?

A. 45^0
B. 55^0
C. 65^0
D. 145^0

Задачата е от подобласт „Геометрични фигури” и познавателна област Приложение. Тя оценява умението да се намери неизвестен остър ъгъл в правоъгълен триъгълник.

M032294.

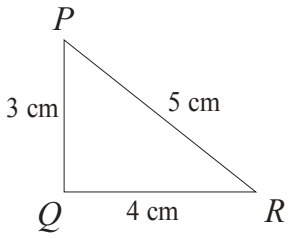


Дадени са точките *M* и *N* от чертежа. Иван търси такава точка *P*, че триъгълникът *MNP* да е равнобедрен. Коя от точките може да бъде *P*?

A. (3, 5)
B. (3, 2)
C. (1, 5)
D. (5, 1)

Задачата е от подобласт „Геометрични фигури” и познавателна област Знание. Тя оценява умението да се намери точка в координатна система, която е връх на равнобедрен триъгълник, при дадени върхове на основата.

M032402.

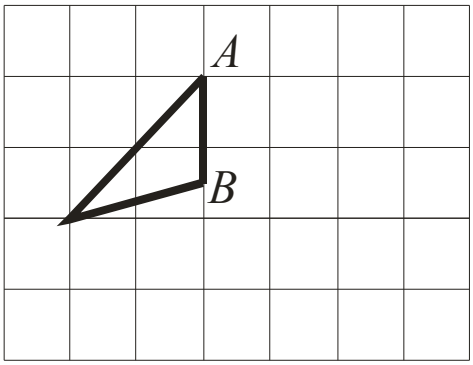


Кое ни дава основание да твърдим, че *PQR* е равностранен триъгълник?

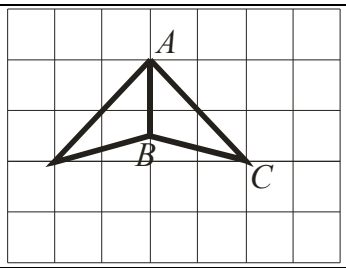
A. $3^2 + 4^2 = 5^2$
B. $5 < 3 + 4$
C. $3 + 4 = 12 - 5$
D. $3 > 5 - 4$

Задачата е от подобласт „Геометрични фигури” и познавателна област Знание. Тя оценява умението да се обосновава, че даден триъгълник е правоъгълен като се използва Питагоровата теорема.

M042151. Начертайте останалата част от фигурата, така че *AB* да е нейна ос на симетрия



Ръководство за оценяване.

Код	Отговор	Задача No. M042151
	Правилен отговор	
10	Правилен чертеж.	
		
	Неправилен отговор	
79	Неправилен чертеж (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Геометрични преобразувания” и познавателна област Приложение. Тя оценява умението да допълни дадена фигура до симетрична, ако е дадена част от фигурата и оста на симетрия.

M022055. Колко е периметърът на квадрат, който има лице 100 квадратни метра?

Отговор: _____

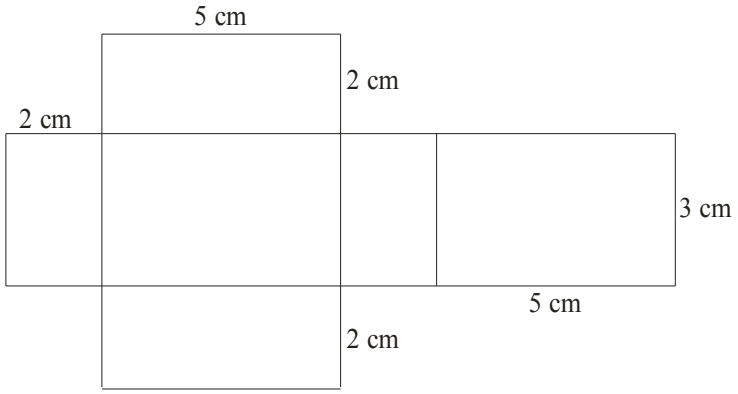
Ръководство за оценяване.

Забележка: Не се прави разлика между отговори дадени с мерни единици или без тях.

Код	Отговор	Задача No. M022055
	Правилен отговор	
10	40	
	Неправилен отговор	
70	25 [100:4 страни]	
71	10 [дължината на една страна]	
72	100 [10.10]	
73	400 [100.4 страни]	
79	Всеки друг грешен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Измерване в геометрията” и познавателна област Приложение. Тя оценява умението да се намери периметъра на квадрат като е дадено лицето му в квадратни единици.

M032344.



Когато показаната фигура се сгъне, се получава кутия с форма на правоъгълен паралелепипед. На колко е равен обемът на кутията?

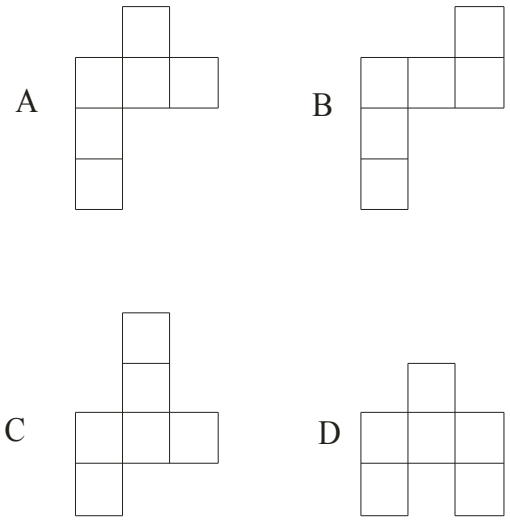
Отговор: _____ cm^3

Ръководство за оценяване.

Код	Отговор	Задача No. M032344
	Правилен отговор	
10	30 или еквивалентно на него.	
	Неправилен отговор	
79	Неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

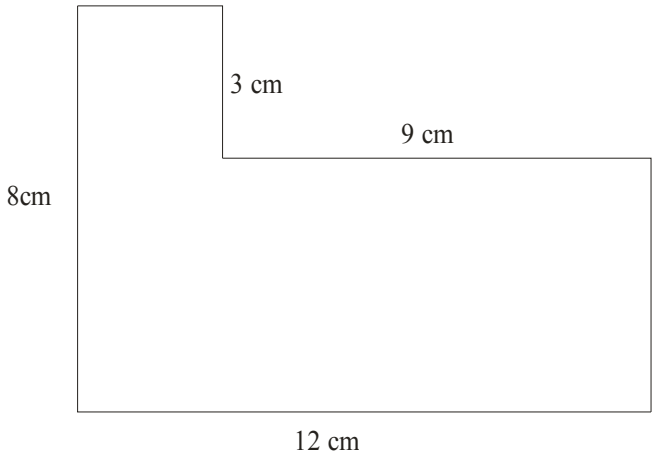
Задачата е от подобласт „Измерване в геометрията” и познавателна област Приложение. Тя оценява умението да се пресметне обемът на правоъгълен паралелепипед по дадени размери на развивката му.

M042265. Коя от фигурите като се сгъне, ще се получи куб?



Задачата е от подобласт „Геометрични фигури” и познавателна област Приложение. Тя оценява умението да се идентифицира развивка на куб.

M032575.

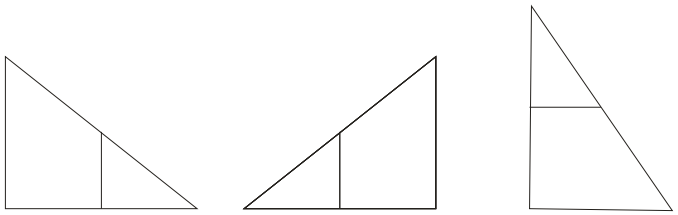


Колко е лицето на фигурата в квадратни сантиметри?

A. 66
B. 69
C. 81
D. 96

Задачата е от подобласт „Измерване в геометрията” и познавателна област Приложение. Тя оценява умението да се намери лицето на неправилна фигура, съставена от два правоъгълника.

M032397.



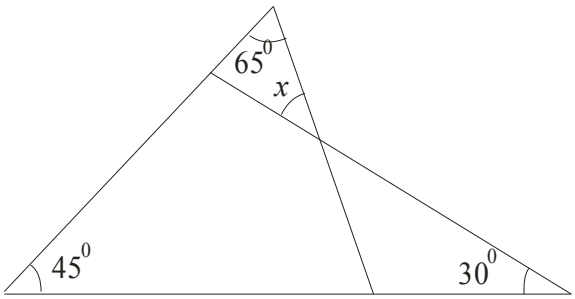
Фигура 1 Фигура 2 Фигура 3

Кое от изброените преобразувания, приложени в указания ред, преобразува Фигура 1 във Фигура 2 и след това във Фигура 3?

A. Осева симетрия и транслация
B. Осева симетрия и въртене на 90^0 по посока на часовниковата стрелка
C. Въртене на 180^0 и транслация
D. Въртене на 90^0 обратно на часовниковата стрелка и осева симетрия

Задачата е от подобласт „Геометрични преобразувания” и познавателна област Знание. Тя оценява умението да се определят преобразуванията, които преобразуват последователно една фигура няколко пъти.

M032398.



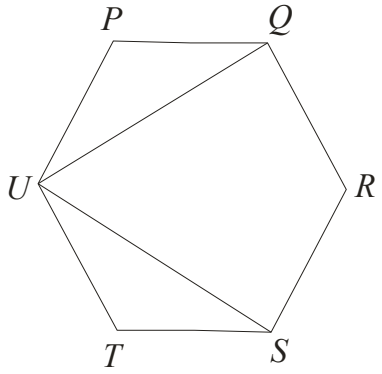
Каква е стойността на x от фигурата?

A. 30^0
B. 40^0
C. 45^0
D. 65^0

Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се прилагат свойствата на външен и вът-

решен ъгъл в триъгълника, за да се намери неизвестен ъгъл във фигура образувана от триъгълници.

M032205.

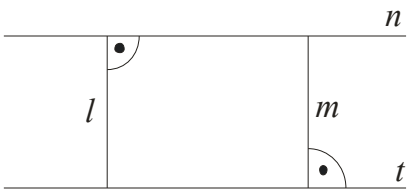


Ако *PQRSTU* е правилен шестоъгълник, на колко е равен ъгъл *QUS*?

A. 30^0
B. 60^0
C. 95^0
D. 120^0

Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се използват свойствата на правилния шестоъгълник, за да се намери мярката на ъгъл.

M042120.

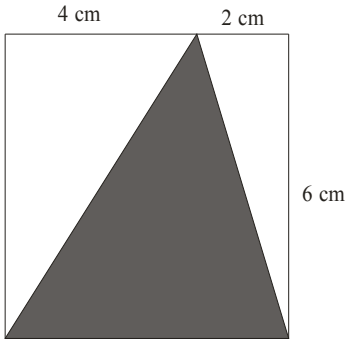


Правите *n* и *t* са успоредни и два прави ъгли са означени на чертежа. Кое от твърденията е вярно за отсечките *l* и *m*?

A. Отсечките *l* и *m* са успоредни и с равни дължини.
B. Отсечките *l* и *m* са перпендикулярни и с равни дължини.
C. Отсечката *l* е успоредна на *m*, но е по-къса от *m*.
D. Отсечката *l* е перпендикулярна на *m*, но е по-къса от *m*.

Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се определи правилното твърдение, отнасящо се до успоредни прави.

M022243. На фигурата е показан потъмнен триъгълник вписан в квадрат.



На колко е равно лицето на потъмнения триъгълник?

Отговор: _____

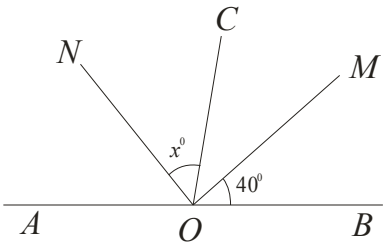
Ръководство за оценяване.

Забележка: Не се прави разлика между отговори дадени с мерни единици или без тях.

Код	Отговор	Задача No. M022243
	Правилен отговор	
10	18	
	Неправилен отговор	
70	36	
79	Всеки друг неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Измерване в геометрията” и познавателна област Приложение. Тя оценява умението да се намери лицето на триъгълник вписан в квадрат с дадени размери.

M032414.



Точките *A*, *O* и *B* лежат на една права, *OM* е ъглополовяща на ъгъла *BOC* и *ON* е ъглополовяща на ъгъла *AOC*.На колко е равно *x*?

Отговор: _____

Ръководство за оценяване.

Код	Отговор	Задача No. M032414
	Правилен отговор	
10	50 (с или без знак за градуси)	
	Неправилен отговор	
70	40 (с или без знак за градуси)	
79	Всеки друг неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

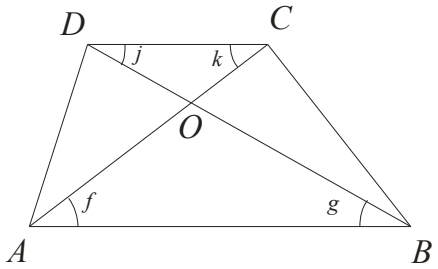
Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се решават задачи с ъглополовящи на съседни ъгли.

M032331. На колко градуса се е завъртяла минутата стрелка на часовник от 6:20 сутринта до 8:00 сутринта на същия ден?

A. 680^0
B. 600^0
C. 540^0
D. 420^0

Задачата е от подобласт „Измерване в геометрията” и познавателна област Приложение. Тя оценява умението да се прилагат знания за часовника за решаване на задачи за намиране на ъгли, свързани с часовниковата стрелка.

M042264.



На фигурата *ABCD* е трапец.

Триъгълниците *AOB* и *COD* са подобни. Напишете една двойка ъгли, които трябва да са равни, защото тези два триъгълника са подобни.

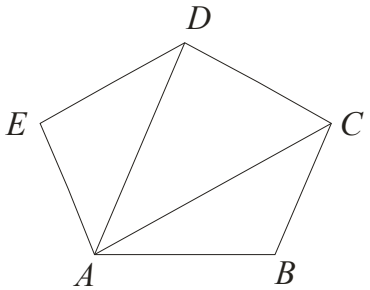
Отговор: _____

Ръководство за оценяване.

Код	Отговор	Задача No. M042264
	Правилен отговор	
10	Което и да е от равенствата $f = k$ или $g = j$.	
	Неправилен отговор	
79	Неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се използват свойствата на подобните триъгълници, за да се определят равни ъгли.

M032692.



На колко е равен сборът от всички вътрешни ъгли на петоъгълника? Покажете как го намерихте.

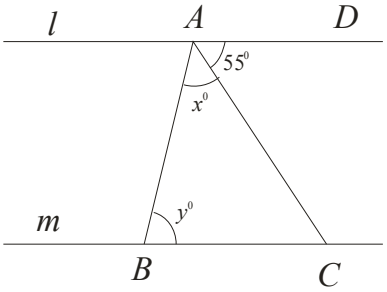
Отговор: _____

Ръководство за оценяване.

Код	Отговор	Задача No. M032692
	Правилен отговор	
20	540 градуса или 6 прави ъгъла. Правилно показан начин. (например: 3 триъгълника умножено по 2 прави ъгъла или по 180 градуса)	
	Частично верен отговор	
10	540 градуса или 6 прави ъгъла, но без да се покаже как е намерено	
	Неправилен отговор	
79	Неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се използва знанието за сбор от ъгли в триъгълника, за да се намери борът от вътрешните ъгли в петоъгълник.

M042137.

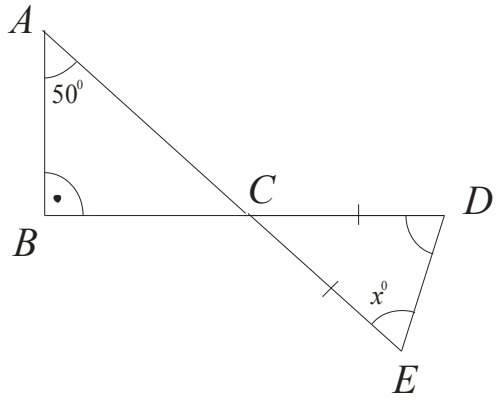


Правите l и m са успоредни. Мярката на ъгъл DAC е 55^0. Каква е стойността на $x+y$?

A. 55
B. 110
C. 125
D. 135

Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се прилагат знания за успоредни прави и триъгълници, за да се намери сбор от ъгли.

M042036.



Дадено е, че $CD = CE$. На колко е равно x?

A. 40
B. 50
C. 60
D. 70

Задачата е от подобласт „Геометрични фигури” и познавателна област Аргументация. Тя оценява умението да се използват свойствата на равнобедрен правоъгълен триъгълник, за да се намери мярката на ъгъл.

M032324. Точките A, B и C лежат на една права и B е между A и C. Ако $AB = 10$ cm и $BC = 5,2$ cm, колко е разстоянието между средите на AB и BC?

A. 2,4 cm
B. 2,6 cm
C. 5,0 cm
D. 7,6 cm

Задачата е от подобласт „Измерване в геометрията” и познавателна област Приложение. Тя оценява умението да се използва информация за дължини на отсечки върху една права, за да се намери разстоянието между средите им.

От прегледа на тези (а и други, не цитирани тук) задачи от областта Геометрия на TIMSS, могат да се направят някои изводи, касаещи преподаването на темата. За учениците от VIII клас се счита, че е важно да знаят и използват за решаване на задачи:

1. Успоредни прави.

2. Лице, периметър и връзка между тях в конкретни фигури.

3. Познаване и идентифициране на развивка на правоъгълен паралелепипед.

4. Чертане върху дадена квадратна мрежа на: фигури с определено лице, точка по дадени координати, образ на фигура при симетрия.

5. Сбор от ъглите в триъгълник.

6. Външен ъгъл в триъгълник.

7. Свойства на правоъгълен триъгълник, включително Питагоровата теорема.

8. Свойства на равнобедрения триъгълник.

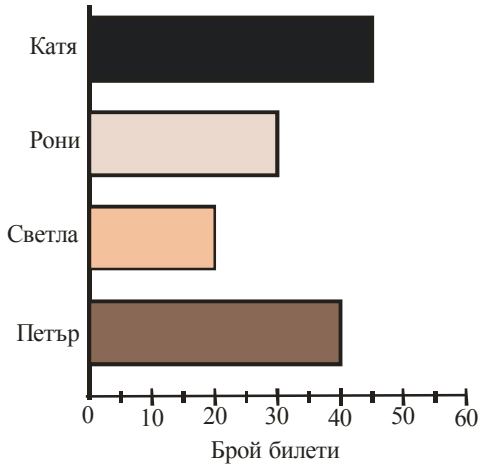
9. Идентифициране на образ на фигура при дадена еднаквост или композиция от еднаквости.

10. Установяване на съответни елементи в подобни триъгълници.

Изброените знания и умения естествено са основни за обучението по математика. Както се вижда, те са основни и за оценяването на постиженията на учениците от изследваната възрастова група. От изброената по-горе тематика, в Българското училище до осми клас малко се застъпва точка 4. Не се изучават Питагоровата теорема и подобни триъгълници – това е материал за по-горните класове. Трябва да се подчертае, обаче, че знанията, които се предполагат в TIMSS по тези теми, са на ниво познаване, т.е. не се предполага задълбочено изучаване на подобни триъгълници, както и по-сложни приложения на Питагоровата теорема от нейното директно прилагане.

Данни и вероятности

M042250



Катя, Рони, Светла и Петър продавали билети за училищен концерт. Диаграмата показва броя на билетите, които всеки от тях е продал. Двама ученици заедно са продали толкова билета, колкото е продала Катя. Кои са те?

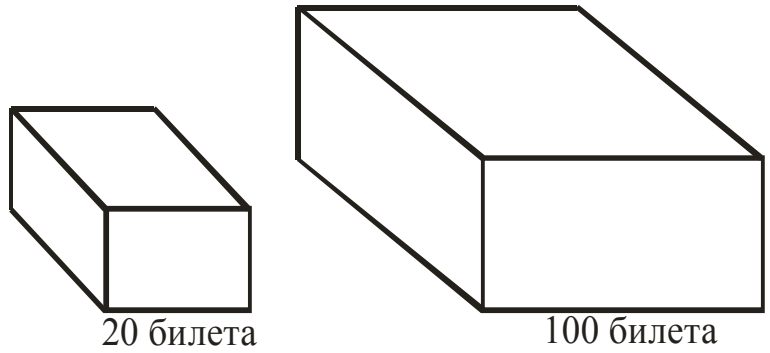
Отговор _____ и _____

Ръководство за оценяване.

Код	Отговор	Задача No. M042250
	Правилен отговор	
10	Рони и Светла.	
	Неправилен отговор	
79	Неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Организиране и представяне на данни” и познавателна област Приложение. Тя оценява умението да се прочетат данните от диаграма и да се определят количествата, които удовлетворяват дадено условие.

M022101. В малката кутия има 20 билета, номерирани с числата от 1 до 20. В голямата кутия има 100 билета, номерирани с числата от 1 до 100.



Без да гледате, избирате билет от някоя от кутиите. От коя кутия с по-голям шанс бихте избрали билет с номер 17?

A. От кутията с 20 билета
B. От кутията със 100 билета
C. С един и същ шанс от двете кутии
D. Не може да се каже

Задачата е от подобласт „Вероятност” и познавателна област Знание. Тя оценява знанието, че вероятността за благоприятен изход на едно събитие е обратно пропорционална на броя на елементите в множеството от събития.

M42260. Петър и Христо са кандидати за председател на ученическия съвет в училище. Резултатите от избора са:

За Петър са гласували 80% от учениците.
За Христо са гласували 20% от учениците.

Колко вероятно е произволно избран ученик да е гласувал за Петър?

A. Съвсем сигурно е, че ученикът е гласувал за Петър.
B. Твърде вероятно е, че ученикът е гласувал за Петър.
C. Не е много вероятно, че ученикът е гласувал за Петър.
D. Съвсем сигурно е, че ученикът не е гласувал за Петър.

Задачата е от подобласт „Вероятност” и познавателна област Приложение. Тя оценява умението в дадена ситуация да се прецени определен шанс като „твърде вероятен”.

M022257. В една купа има 36 еднакви по големина мъниста. Някои от тях са сини, някои са зелени, някои са червени и някои са жълти. Извадено е едно мънисто, без да се гледа в купата. Вероятността то да е синьо е $\frac{4}{9}$. Колко сини мъниста има в купата?

A. 4
B 8
C. 16
D. 18
E. 20

Задачата е от подобласт „Вероятност” и познавателна област Приложение. Тя оценява умението да се определи броят на благоприятните изходи, като са дадени всички изходи и вероятността за благоприятен изход.

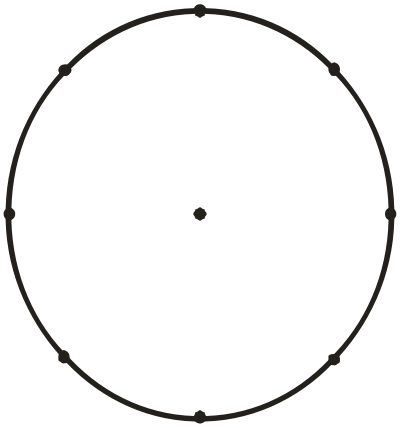
M04222. В чантата на София има 16 камъчета. От тях 8 са червени и 8 са черни. Тя изважда 2 черни камъчета и не ги връща обратно. След това София изважда трето камъче от чантата. Какво може да се каже за цвета на третото извадено камъче?

A. По-вероятно е то да е червено, отколкото черно.
B. По-вероятно е то да е черно, отколкото червено.
C. Еднакво вероятно е то да е червено или черно.
D. Не можа да се каже от кой цвят е по-вероятно да е то.

Задачата е от подобласт „Вероятност” и познавателна област Приложение. Тя оценява умението да се определи най-вероятният изход при дадена ситуация.

M032695. В едно училище има 400 абитуриент. От тях 50 ще учат в университет, 100 – в политехниката, 150 – в бизнес колеж, а останалите ще постъпят на работа.

Като използвате дадения кръг, представете с кръгова диаграма тази информация. Означете всяка област на начертаната диаграма.

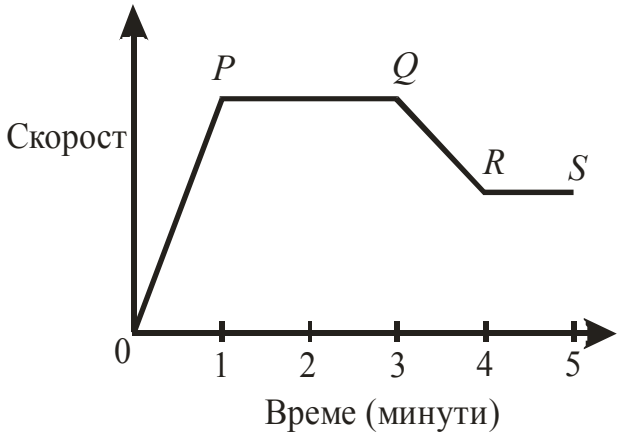


Ръководство за оценяване.

Код	Отговор	Задача No. M032695
	Правилен отговор	
20	Правилно начертана и означена кръгова диаграма. (1 сектор – университет; 2 сектора – политехника; 2 сектора – работещи; 3 сектора – бизнес колеж)	
	Частично верен отговор	
10	Четири сектора, от които поне два, но не всички, са правилно начертани и означени.	
11	Четири сектора правилно начертани, но без означение или означения 50, 100, 150, 200.	
	Неправилен отговор	
70	Четири сектора, но само един или нито един не е начертан правилно.	
79	Неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Организиране и представяне на данни” и познавателна област Приложение. Тя оценява умението да се начертае и означи кръгова диаграма по дадени данни.

M042252



Петя кара велосипед. Графиката попазва изменението на скоростта през първите 5 минути от пътуването. Кое би могло да е възможно обяснение за частта от графиката между *Q* и *R*?

A. Петя кара велосипеда нагоре по хълм.
B. Петя кара велосипеда надолу по хълм.
C. Петя спира за една минута.
D. Петя кара велосипеда обратно на посоката, в която е тръгнала.

Задачата е от подобласт „Интерпретиране на данни” и познавателна област Аргументация. Тя оценява умението да се определи възможно описание с думи на част от графика.

M042269. Два отбора се състезават на дълъг скок. Ето техните резултати:

	Средна дължина на скока
Отбор А	3,6 м
Отбор Б	4,8 м

Във всеки отбор има един и същ брой ученици.
Кое от твърденията е вярно?

A. Всеки ученик от обор Б е направил по-дълъг скок от който и да е ученик от отбор А.
B. След скока на всеки ученик от отбор А е скачал ученик от отбор Б, който е направил по-дълъг скок.
C. Като цяло, отбор Б е направил по-дълги скокове от отбор А.
D. Някои ученици от отбор А са направили по-дълги скокове от някои ученици от отбор Б.

Задачата е от подобласт „Интерпретиране на данни” и познавателна област Приложение. Тя оценява умението да се определи вярно твърдение, свързано с понятието средна стойност.

M042167. Ахмед се явява на общо 5 математически теста, на всеки от които може да получи максимум 10 точки. Точките на Ахмед върху първите 4 теста са съответно 9, 7, 8, 8. Възможно ли е Ахмед да получи среден успех 9 от всичките 5 теста?

Обосновете отговора си.

Ръководство за оценяване.

Код	Отговор	Задача No. M042167
	Правилен отговор	
10	НЕ с правилна обосновка. Например: Ахмед трябва да получи 13; или той може да получи средно 8,4; или нужни са му 45 точки, а може да получи най-много 42; или еквивалентни на тези обяснения.	
	Неправилен отговор	
79	Неправилен отговор (включително зачертани, изтрити, неясни бележки и др. белези).	
	Без отговор	
99	Празно	

Задачата е от подобласт „Интерпретиране на данни” и познавателна област Аргументация. Тя оценява разбирането на понятието средна стойност за решаване на задачи.

От прегледа на тези (а и други, не цитирани тук) задачи от областта Данни и вероятности на TIMSS, могат да се направят някои изводи, касаещи преподаването на темата. За учениците от VIII клас се счита, че е важно да знаят и използват за решаване на задачи:

1. Четене и интерпретиране на данни от таблици, диаграми и графики.

2. Графично представяне на данни.

3. Разбиране на понятието средна стойност.

4. Разбиране на понятията по-вероятно, по-малко вероятно, както и на елементарна вероятност в практически ситуации.

Изброените знания и умения са на сравнително елементарно ниво. Въпреки че в учебния план по математика в България не се акцентира върху темата Данни и вероятности, тези знания и умения не са нови за учениците. Те се изучават, макар и епизодично в различни периоди в училище.

ТРЕТА ЧАСТ
ОБРАБОТКА И АНАЛИЗ НА ДАННИ ОТ ШИРОКО МАЩАБНИ ОЦЕНЪЧНО-ДИАГНОСТИЧНИ ИЗСЛЕДВАНИЯ

При събирането на данни от широко мащабни национални или международни ОДИ често пъти не е възможно да се използват готови статистически рецепти или стандартен софтуер. Причината е, че такива оценявания работят със сложна извадка и това дава отражение на статистическите методи за анализ на данните. Използването на вероятностното моделиране също изисква специфични техники и адекватни статистически процедури.

В тази част се разглеждат някои основните методи за обработка и анализ на такива данни. Обръща се внимание на методики и технологии, които съществено се различават от класическите приложения на статистическите методи. Описаните тук процедури лесно се адаптират, за да са адекватни на моделите, използвани при конкретно оценяване. Целта е колкото се може по-добре да се моделира реалната ситуация от събирането на данните, за да бъдат обективни направените анализи.

ОСМА ГЛАВА
ТЕГЛА НА ОБЕКТИТЕ ОТ ИЗВАДКАТА

Както беше обяснено в трета глава, при широко мащабни ОДИ се използват сложни извадки с цел да се прилагат валидни и ефикасни методи за получаване на неизместени оценки на цялата популация. Точността на тези оценки в голяма степен зависи от прецизността, с която е направена извадката, и от използваните статистически методи за анализ на получените данни.

Класическо предположение при прилагане на статистически методи е, че наблюденията са независими и с едно и също вероятностно разпределение. При изследвания в образованието ситуациите са значително по-сложни и се налага използването на така наречените *устойчиви методи*, които са по-малко чувствителни към горното предположение и към наличието на изключително нетипични елементи в извадката (аутлайери). Повечето от тези методи изискват пресмятане на отделно тегло за всеки наблюдаван обект, за да се вземе пред вид различната вероятност, с която обектите попадат в извадката. Отчитането на тези тегла намалява възможността за значителни измествания на оценките на цялата популация.

1. Пресмятане на извадковите тегла

Използването на извадки при широко мащабни ОДИ е важно от технологична гледна точка. Данните получени за извадката, трябва да се обобщят за цялата популация. Извадковите тегла са числови стойности, които се приписват на всеки елемент (ученик) от извадката, за да може полученото обобщение да бъде по-точно и по-устойчиво. В тези тегла е заложена и корекция за училища, паралелки и/или ученици, които са попаднали в извадката, но по една или друга причина не са отговорили на изследването. Извадковите тегла са обратно пропорционални на вероятността, с която даден ученик попада в извадката. В този смисъл са показател за това колко ученика от популацията се представляват от даден ученик от извадката.

Описаният в трета глава метод на двустъпкова стратифицирана клъстерна извадка е удобен от гледна точка на устройството на училищната система. При него обаче учениците се избират с различни вероятности. За да се осигури по-голяма точност на получените оценки на цялата популация, нужно е това да се вземе пред вид, като се пресметне тегло за всеки ученик от извадката. Тези

тегла се наричат *извадкови тегла*. Тук се обяснява пресмятането им като се има предвид двустъпкова стратифицирана клъстрена извадка. Описани са три стъпки за пресмятане на извадковите тегла, свързани съответно с вероятността за избор на училище, паралелка и ученик.

Първата стъпка изчислява извадково тегло на всяко училище, след което се прави корекция за училищата, които са попаднали в извадката, но поради една или друга причина не са участвали в изследването. Втората стъпка изчислява извадково тегло на всяка паралелка, отделно и независимо за всяко училище, като прави корекция за паралелките от извадката, които не са участвали. Последната стъпка изчислява извадково тегло на всеки ученик, отделно и независимо за всяка паралелка. Когато се избират цели паралелки (т.е. всеки ученик от паралелката попада в извадката), извадковото тегло на всеки ученик е 1. Това не е така, ако се прави извадка на ученици в самите паралелки. Разбира се и на това ниво се прави корекция за учениците, които са в извадката, но не са участвали в изследването. На края се изчислява пълно извадковото тегло за всеки ученик, което е произведение от трите извадкови тегла: това на училището, в което учи, на съответната паралелка, в която е ученикът, и на собственото му тегло в рамките на паралелката (TIMSS 2003 School Sampling Manual).

2. Първа стъпка – училищно извадково тегло

Училищното извадково тегло е обратно пропорционално на вероятността съответното училище да попадне в извадката. Като се вземе предвид описания в трета глава PPS метод на извадка (вероятността дадено училище да попадне в извадката е пропорционална на „големината на това училище”), училищното извадково тегло може да се пресметне така: Да предположим, че цялата популация е съсредоточена в N на брой училища, от които в извадката трябва да попаднат само n от тях ($n<N$). Да означим с m_i броят на учениците от популацията в училището с номер i (големината на училището) и с $M = \sum_{i=1}^{N} m_i$. Тогава извадковото тегло на училището с номер i е равно на $BW_{sc}^{i} = \frac{M}{n.m_i}$.

След пресмятане на училищните извадкови тегла, се изчислява коригиращ множител за училищата, които са попаднали в извадката, но не са участвали в изследването. Този коригиращ множител е един и същ за всички учи-

лища и е равен на $A_{sc} = \frac{n_s + n_{r1} + n_{r2} + n_{nr}}{n_s + n_{r1} + n_{r2}}$, където n_s е броят на всички училища в извадката, n_{r1} и n_{r2} са броят на училищата, които са съответно първи и втори заместник (вж. трета глава) и n_{nr} е броят на училищата от извадката, които не са участвали в изследването.

Окончателно, училищното извадково тегло на i-тото училище е $FW_{sc}^i = A_{sc}.BW_{sc}^i$.

3. Втора стъпка – извадково тегло на паралелка

Извадковото тегло на паралелката е обратно пропорционално на вероятността, с която паралелка от вече избрано училище попада в извадката. В случая когато паралелката се избира като цялостна единица (т.е. не се прави извадка на ученици от паралелката), се счита, че всяка паралелка в училището има една има съща вероятност да бъде избрана. Тогава теглото на която и да е паралелка от училището с номер i е равно на $BW_{cl}^i = \frac{C^i}{c^i}$, където C^i е броят на всички паралелки от популацията в училището и c^i е броят на паралелките, които трябва да бъдат избрани от училището. В повечето случаи c^i има стойности 1 или 2 и много рядко 3. Когато се прави извадка на ученици от паралелката, се прилага PPS метода за извадка на паралелка и съответно се коригира формулата за извадковото тегло на паралелката.

След пресмятането на извадковите тегла на паралелките, се изчислява коригиращ множител за паралелките, които са попаднали в извадката, но не са участвали в изследването. (Паралелки с малък процент участие – примерно под 50% от учениците, могат да се считат, че не са участвали. Малкият процент на участващи ученици води до изместване на оценките на популацията.) Коригиращият множител за паралелките се пресмята отделно за всяка страта от пряката стратификация и е равен на $A_{cl} = \frac{\sum_{i=1}^{s+r1+r2} c^i}{\sum_{i=1}^{s+r1+r2} c_*^i}$, където s е броят на училищата в съответната страта, $r1$ и $r2$ са броят на училищата, които са съответно първи и

втори заместник, а c_*^i е броят на паралелките в *i*-тото училище, които действително са участвали в изследването.

Окончателно, извадковото тегло на паралелка е равно на $FW_{cl}^i = A_{cl}.BW_{cl}^i$.

4. Трета стъпка – извадково тегло на ученик

Извадковото тегло на ученик е обратно пропорционално на вероятността, с която ученикът от вече избрана паралелка попада в извадката. В голяма част от случаите паралелаката се избира като цялостна единица, което означава, че щом паралелката е избрана, всички ученици от нея също са избрани. Тогава вероятността на даден ученик от избрана паралелка да попадне в извадката е равна на 1, следователно теглото на всеки ученик от *j*-тата паралелка на *i*-тото училище е равно на $BW_{st}^{i,j} = 1$. В редките случаи, когато се налага да се избират ученици от паралелките, това тегло се коригира по съответен начин.

Както и при по-горните стъпки, така и тук се изчислява коригиращ множител за учениците от паралелката, който компенсира факта, че някои ученици не са участвали в изследването. За всеки ученик от *j*-тата паралелка на *i*-тото училище този множител е равен на $A_{st}^{i,j} = \frac{s_{st}^{i,j} + s_{nr}^{i,j}}{s_{st}^{i,j}}$, където $s_{st}^{i,j}$ е броят на учениците от *j*-тата паралелка на *i*-тото училище и $s_{nr}^{i,j}$ е броят на учениците от *j*-тата паралелка на *i*-тото училище, които са участвали в изследването.

Извадково тегло на ученик от *j*-тата паралелка на *i*-тото училище е числото $FW_{st}^{i,j} = A_{st}^{i,j}.BW_{st}^{i,j}$.

5. Пълно извадково тегло

Пълното извадково тегло на ученик е произведение от извадковите тегла, които са описани в трите стъпки по-горе, а именно за всеки ученик от *j*-тата паралелка на *i*-тото училище то е $W^{i,j} = FW_{sc}^i \cdot FW_{cl}^i \cdot FW_{st}^{i,j}$. Важно е да се отбележи, че пълното извадково тегло е едно и също за всички ученици от дадена паралелка, но е различно за ученици от различни паралелки или от различни училища. Извадковите тегла се пресмятат отделно за всяка страта от пряката стратификация.

Когато се прави извадка, всеки елемент на извадката представлява определен брой обекти от цялата популация. При проста случайна извадка всеки елемент представлява един и същ брой обекти от популацията. Затова без ограничения на общността може да считаме, че теглото на всеки елемент е равно на 1. (Всички елементи имат едно и също тегло.) Смисълът на пълното извадково тегло е, че то показва колко ученика от цялата популация представлява даден ученик от извадката. Така сборът на всички пълни извадкови тегла за учениците от извадката е равен на броя на учениците от цялата популация.

6. Един пример

Ще демонстрирам ефекта от използване на извадкови тегла върху един пример. Данните са специално подбрани така, че читателят да разбере същността на проблема, а не да бъде затруднен с изморителни пресмятания.

Примерът е с въображаема популация, в която има две страти (Таблица 15).

Таблица 15. Пример на извадкови тегла.

	Страта 1	Страта 2	Общо
Цялата популация	1000	10 000	11 000
Извадка	100	100	200
Извадкови тегла	10	100	——
Не претеглена средна стойност	500	600	550
Претеглена средна стойност	500	600	591

Първата страта има 1000 ученика, а втората – 10 000 ученика. За да имаме добра сравнимост на резултатите от двете популации се прави извадка с еднакъв брой ученици – по 100 ученика от страта. Така, всеки ученик от първата страта представлява 10 ученика от популацията, а всеки ученик от втората страта представлява 100 ученика от популацията. Това означава, че извадковото тегло на всеки ученик от първата страта е 10, а на ученик от втората страта то е 100. Да предположим, че е проведен тест и средната стойност на учениците от първата страта е 500, а на тези от втората страта е 600. Ако работим без извадкови тегла, средната стойност (не претеглена) на цялата популация би била $M = \frac{100 \cdot 500 + 100 \cdot 600}{200} = 550$. Полученото число не е представително за популацията, понеже голямото количество от учениците са от страта 2, която има средна стойност 600 и малка част от тях са от страта 1, която има средна стой-

ност 500. Бихме очаквали, че популацията има средна стойност по-малка от 600, но близка до това число. Като използваме теглата на учениците получаваме $M = \frac{100 \cdot 10 \cdot 500 + 100 \cdot 100 \cdot 600}{100 \cdot 10 + 100 \cdot 100} = 590,91$ като претеглена средна стойност на популацията. Това число е много по-приемливо за средна стойност на популацията при наличните данни.

В реална ситуация ролята на извадковите тегла е аналогична на тази в разгледания пример. Те компенсират ефекта от неравните вероятности, с които учениците попадат в извадката, т.е. от това, че различните ученици представляват различен брой обекти от цялата популация.

ДЕВЕТА ГЛАВА
ДЖАКНАЙФ МЕТОД ЗА ПРЕСМЯТАНЕ НА ИЗВАДКОВАТА ГРЕШКА

Начинът за получаване на информация за цялата популация при използване на извадки се базира на статистически процедури, които позволяват да правим оценки за популацията въз основа на известни статистики на извадката. Естествено възниква въпросът какво е влиянието на конкретната извадка. Ако се използва друга извадка дори ако тя е правена по съвсем същата методика, но в нея има други ученици, нейните статистики биха били различни. Значи ли това, че ще получим различни оценки за популацията?

Класическата статистическа теория ни предлага следното решение на проблема. Да изследваме няколко различни еднотипно направени извадки, след това да усредним получения резултат и да използваме него като оценка на популацията. Това решение е възможно, ако процедурата, която използваме за получаване на резултатите лесно може да се изпълни няколко пъти, т.е тя е сравнително лесна (не трудоемка, евтина и т.н.) за изпълнение. Това не е така в педагогическите изследвания, където измерването на дори само една извадка е изключително скъпа и трудоемка дейност, която изисква много време. На практика не е възможно да се изследват няколко извадки в кратък интервал от време. Едно от нещата, които Джакнайф метода (Jackknife method) ни предлага е, че той симулира множество различни еднотипни извадки като използва една единствена конкретно направена извадка.

Друга възможност на този метод е, че коригира пресмятането на дисперсията (съответно на стандартното отклонение) като взема пред вид факта, че ученици, които учат в една паралелка, са склонни да показват по-еднотипни знания (т.е. дисперсията на резултатите им е по-малка) от същите ученици, ако те учеха в различни паралелки или училища. В трета глава, за нуждите на извадката, този факт се отчиташе с така наречения коефициент на паралелкова корелация. В термините на дисперсията, казаното означава, че ако използваме класическия „стандартен” начин за пресмятане на дисперсията ще получим по-малка стойност от действителната. Джакнайф методът коригира тази неточност.

1. Описание на метода

Джакнайф методът (Rust, K. 1985, 1986) е статистическа процедура, която използва оригинална извадка от N елемента за оценка на даден параметър t

на популацията като симулира множество еднотипни извадки чрез последователно елиминиране на елементи от извадката и увеличаване на ефекта на някой от другите елементи. Всяка една от тези симулирани извадки дава оценка на t. Чрез комбинация от тези „много” оценки се получава неизместена оценка на t и се коригира стойността на дисперсията. Методът е сравнително лесен за разбиране и прилагане (дори и в сложни ситуации). Едно от предимствата му е, че не се правят предположения за вида на разпределението, което се анализира, и това води до голяма надеждност при приложенията в практиката.

Първото приложение на Джакнайф метода в педагогически изследвания е направено за данни от NAEP в САЩ (Johnson, E. G., Rust, K. F., 1992). След това той е използван за анализ на данните на TIMSS.

2. Конструиране на Джакнайф зони

Както е описано в трета глава, списъкът от училищата в извадката е направен така, че училищата са подредени в намаляваща големина (брой ученици от разглежданата популация). Първата стъпка е да групираме училищата от този списък в така наречените зони. Всяка зона се състои от две последователни в списъка училища, които са участвали в изследването. Ако списъкът съдържа нечетен брой училища, последната зона се състои от три училища. Групирането на училищата в зони става отделно за всяка страта от пряката стратификация, т.е. не може в една и съща зона да има училища от различни страти. Това се прави с цел училищата в една и съща зона да са колко се може по-еднотипни от гледна точка на тяхната големина.

В обясненията по-долу броят на зоните е означен с H.

3. Пресмятане на дисперсията с Джакнайф метода

Да означим с t интересуващата ни статистиката (средна стойност, процент, медиана, и др.). Да значим с $t(S)$ стойността на тази статистика, пресметната върху действителната извадката, която сме направили. За пресмятането на $t(S)$ обикновено се използват извадкови тегла, както е обяснено в осма глава.

Да разгледаме сега двете училища (с номера 1 и 2) от първата зона. По произволен начин избираме едно от тези две училища и приписваме на всеки негов ученик тегло 0. (Това означава, че данните на учениците от това училище

няма да се вземат пред вид при пресмятането на статистиката *t*.) На всеки ученик от другото училище в зоната удвояваме теглото. (Това означава, че данните на всеки ученик от това училище се удвояват при пресмятане на *t*.) За учениците от училищата в останалите зони нищо не променяме. При тези нови „симулирани" тегла означаваме с $t(J_1)$ стойността на *t*, пресметната със симулираните тегла. Всъщност, $t(J_1)$ е стойността на интересуващата ни статистика върху друга, симулирана извадка от ученици, поради което нейната стойност може да се различава от $t(S)$.

След това постъпваме по същия начин с училищата от втората зона и означаваме с $t(J_2)$ стойността на *t* пресметната с други „симулирани" тегла, т.е. върху друга симулирана извадка. Този процес повтаряме последователно *H* пъти – по веднъж за всяка от зоните. Така получаваме за *t* различни стойности $t(J_1), t(J_2), ..., t(J_H)$.

Дисперсията на *t* (по Джакнайф метода) се пресмята с формулата $Var_{jrr}(t) = \sum_{h=1}^{H} \left(t(J_h) - t(S)\right)^2$. Стандартната грешка в този случай, което е $s.e.(t) = \sqrt{Var_{jrr}(t)}$, се нарича *грешка от извадката*.

Таблица 16 илюстрира определянето на зоните и пресмятането на симулираните тегла в опростена за целта ситуация.

Представени са 10 училища, номерирани от 1 до 10 (първата колона). Във всяко училище има по двама ученика (номерирани във втората колона). Третата колона съдържа номерацията на петте зони, образувани от по две училища във всяка от тях. Следващата колона (JKR) е индикация за произволен избор на училище във всяка зона – за училищата, в които JKR е равно на 0, учениците получават тегло 0 за съответното пресмятане на симулираните тегла; за онези, в които JKR е равно на 1, теглата на учениците се удвояват. Следващата колона е балът на учениците за интересуващата ни статистика *t*. Следват теглата на всеки ученик. Колоната RWgt1 съдържа първите симулирани тегла – в нея теглата на учениците в първата зона са променени съгласно променливата JRK, а във всички останали зони те не са променени. Следващите четири колони съдържат останалите симулирани тегла – за всяка зона по една колона. В дъното на таблицата са дадени съответните средни стойности (непретеглени, претег-

лени и пресметнати със симулираните тегла), както и сборът от теглата. При този пример имаме „реална” извадка от 20 ученика. На всяка от колоните RWgt1, ... RWgt5 може да се гледа като на друга подобна „симулирана” извадка от ученици от същата популация.

Таблица 16. Тегла и симулирани тегла (RWgt…)

Училище	Ученик	Зона	JKR	Бал	Тегла	RWgt1	RWgt2	RWgt3	RWgt4	RWgt5
1	101	1	1	96	6	**12**	6	6	6	6
1	102	1	1	74	7	**14**	7	7	7	7
2	201	1	0	56	9	**0**	9	9	9	9
2	202	1	0	62	2	**0**	2	2	2	2
3	301	2	0	34	2	2	**0**	2	2	2
3	302	2	0	53	7	7	**0**	7	7	7
4	401	2	1	96	5	5	**10**	5	5	5
4	402	2	1	57	2	2	**4**	2	2	2
5	501	3	0	78	2	2	2	**0**	2	2
5	502	3	0	87	2	2	2	**0**	2	2
6	601	3	1	71	8	8	8	**16**	8	8
6	602	3	1	47	7	7	7	**14**	7	7
7	701	4	0	56	1	1	1	1	**0**	1
7	702	4	0	25	2	2	2	2	**0**	2
8	801	4	1	82	2	2	2	2	**4**	2
8	802	4	1	68	2	2	2	2	**4**	2
9	901	5	1	16	10	10	10	10	10	**20**
9	902	5	1	80	8	8	8	8	8	**16**
10	1001	5	0	20	5	5	5	5	5	**0**
10	1002	5	0	35	9	9	9	9	9	**0**

Непретеглено средно:	**59,65**	**Сбор от теглата: 98**	**100**	**96**	**109**	**99**	**102**
Претеглено средно:	**57,17**	**Претеглени средни:**	**60,69**	**59,98**	**56,6**	**58,56**	**58,71**

Описаният Джакнайф метод има разновидности. Те са свързани с даване на различни тегла на училищата в избраната зона. Например, вместо да се дават тегла 0 на учениците в едното училище и да се удвояват теглата на учениците в другото училище, може да се дават тегла 1/2 на учениците от едното училище и 3/2 на учениците от другото училище. Възможни са и друго комбинации.

4. Ефективност на извадката

Сложните извадки, за които става дума в педагогическите изследвания, имат по-голяма грешка на извадката, отколкото стандартната грешка при проста

случайна извадка. Това е така, защото елементите на клъстерите (учениците групирани в паралелки) обикновено дават по-еднородни резултати, отколкото ако те са пръснати произволно по цялата извадка. Като следствие от това получаваме, че проста случайна извадка от *n* ученика е „по-ефективна" (има по-малка грешка от извадката) отколкото извадка от *n* ученика, групирани в паралелки. Степента, в която учениците от паралелките са по-еднородни, отколкото, ако те са пръснати произволно в популацията, се измерва с коефициента на паралелкова корелация (вж. трета глава). Колкото този коефициент е по-голям, толкова по-голяма извадка от ученици, групирани в паралелки, е нужна, за да постигнем ефекта, който бихме постигнали от използването на проста случайна извадка.

Въпреки казаното по-горе, извадките с клъстери от паралелки са за предпочитане в педагогическите изследвания поради технически причини, за които говорихме в трета глава. Един начин за количествено описание на намаляването на ефективността в този случай е така наречената *ефективност на извадката* (Kish, 1965). Той се изчислява като частното от оценката на дисперсията за интересуващата ни статистика *t*, пресметната с Джакнайф метод, и на дисперсията на *t*, която би се получила предполагайки проста случайна извадка, т.е. $Eff(t) = \frac{Var_{jrr}(t)}{Var(t)}$, където $Var_{jrr}(t)$ е обяснено по-горе, а $Var(t)$ е оценка на дисперсията на *t*, предполагайки проста случайна извадка.

Друга мярка, която е свързана с ефективността на извадката, е *ефективната големина на извадката*. Тя се изчислява като отношението на действителния брой *N* на ученици от извадката (групирани в паралелки) и ефективността на извадката, т.е. $EffN(t) = \frac{N}{Eff(t)}$. Смисълът на това число е, че то показва броя на учениците, които са нужни при проста случайна извадка, за да се получи същата точност на оценката на *t*, както получената от *N* ученици, групирани в паралелки.

По-долу са дадени някои от разгледаните стойности за България, пресметнати в TIMSS-95 за средната стойност по математика. Средната стойност по математика е $t(S)$=527; броят на тестваните ученици е N=3771 (избрани са като клъстери от цели паралелки); грешката от извадката, пресметната с

Джакнайф метода е $\sqrt{Var_{jrr}(t)} = 4,6$; стандартната грешка, предполагайки проста случайна извадка, е $\sqrt{Var(t)} = 1,8$; ефективността на извадката е $Eff(t) = 6,53$; ефективната големина на извадката е $EffN(t) = 577$. Последното означава, че точността на оценката *t*(*S*), получена въз основа на клъстерна извадка от 3771 ученика, е същата, каквато е при използването на проста случайна извадка от 577 ученици.

ДЕСЕТА ГЛАВА
ПРАВДОПОДОБНИ СТОЙНОСТИ И ТЯХНОТО ИЗПОЛЗВАНЕ

Правдоподобните стойности са случайно избрани стойности от ненаблюдаемото (латентно) разпределение на постиженията на всеки ученик. Те се използват предимно в широко мащабни ОДИ от типа на TIMSS, PIRLS, PISA за по-прецизна оценка на стандартната грешка в случаите, когато структурата на извадката е сложна или когато грешката от измерването е голяма.

За правилното разбиране на методите, описани в тази глава, е важно да се помни, че при широко мащабни ОДИ интересът на изследователите е върху така наречените статистики на популацията (средна стойност, стандартно отклонение, процентили, процентно разпределение и съответните стандартни грешки), а не върху оценка на постиженията на отделните ученици.

1. Един пример

Да си представим за момент, че се интересуваме от процента на хората от дадена популация, които са по-високи от 180 cm. Изследването на този въпрос обикновено става чрез случайна извадка на достатъчно много хора от популацията. Процентът p на онези от извадката, чийто ръст е по-голям от 180 cm, се счита за неизместена оценка на процента, от който се интересуваме. Стандартната грешка на оценката в този случай е числото $\sqrt{\frac{p(1-p)}{n}}$, където n е броят на хората от извадката. В този случай тя се нарича още *извадкова грешка*.

Да разгледаме сега по-сложния случай, когато искаме да установим процента на хората от дадена популация, които могат да говорят на френски език. Съществената разлика от гледна точка на предишния пример е, че докато измерването на ръста на хората е сравнително ясна задача, то не е така лесно да се каже кой може да говори френски. Някои хора владеят езика свободно; други говорят така, че ги разбират, но правят много грешки; трети могат да кажат само няколко изречения; четвърти едва свързват някои думи. Следователно, трябва ни някаква дефиниция на „може да говори френски език”. Да предположи, че имаме такава дефиниция, която може да се приложи на практика. Изследването на въпроса, който сме си поставили, става чрез проверка на уменията да се говори френски език и прилагането на дефиницията само върху „малка част” от езика (например, разговор в магазин), понеже не може да искаме от хората да

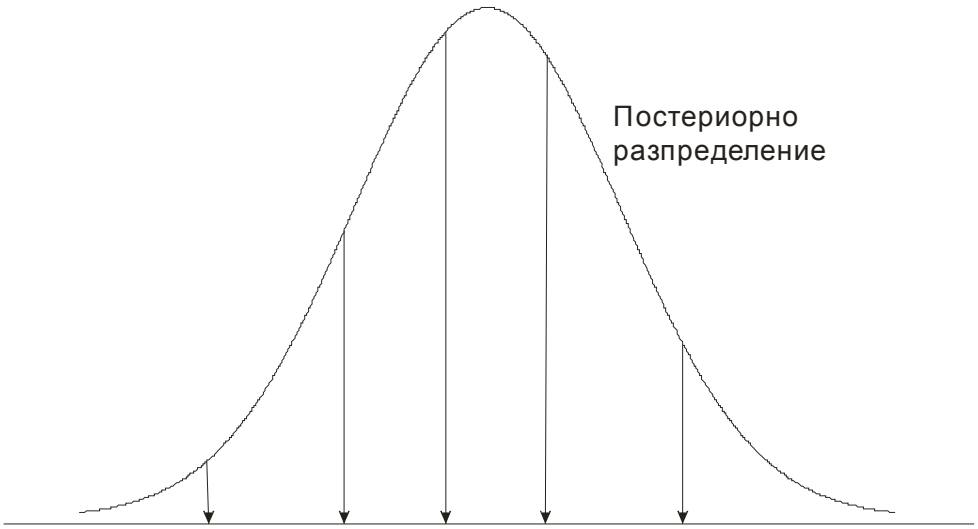
посветят часове, за да ни „покажат” умението си за говорене на френски език върху голямо разнообразие от теми. Това допълнително внася известна „несигурност” (грешка) в установяването на възможността за говорене на френски език. На статистически език това се нарича *грешка на измерването*. Като неизместена оценка на интересуващия ни процент отново може да използваме процента p на хората от случайна извадка, които могат да говорят френски език. Стандартната грешка на тази оценка, обаче, е по-голяма от $\sqrt{\frac{p(1-p)}{n}}$, понеже към извадковата грешка се добавя и грешката от измерването на всеки човек.

Този пример показва, че когато измерването на отделните индивиди от извадката съдържа „съществена” грешка, тя трябва да се вземе пред вид при пресмятане на стандартната грешка за съответната статистика на популацията. Един начин да направим това е да дадем не една, а *няколко* стойности на съответната статистика за всеки индивид от популацията. Ако грешката от измерването е „малка”, тези стойности ще са „близки” една до друга; ако грешката от измерването е „голяма”, стойностите може да са „далеч” една от друга. Тези „няколко стойности” на всеки индивид от популацията се наричат *правдоподобни стойности* (Mislevy, 1991; Mislevy, Beaton, Kaplan, & Sheehan, 1992; Rubin (1987)).

2. Определение на правдоподобните стойности

Да се върнем сега в контекста на измерване на ученическите постижения. Един нагледен начин да опишем правдоподобните стойности за даден ученик е да кажем, че те представят някакъв интервал, в който най-вероятно се намират способностите на ученика при отговорите, които сме получили от него върху определен брой задачи от теста. В четвърта глава обяснихме как се оценява стойността на променливата „способност” за всеки един ученик. Там под „способност” на ученик се разбираше една единствена стойност (точкова стойност), която се получава въз основана дадените от ученика отговори. (Аналогията с горния пример е както определянето на височина – една определена стойност.) В случая, който разглеждаме в тази глава, вместо да оценим една единствена стойност на способностите на даден ученик, ние даваме област от стойности и съответните им вероятности, т.е. задаваме вероятностно разпределение на способностите на всеки ученик. (Аналогията с горния пример е както при опреде-

ляне умението да се говори френски език – за всеки индивид то става в някакъв „интервал” с някаква вероятност.) Това разпределение се нарича постериорно разпределение за ученика. Правдоподобните стойности са случайно избрани стойности от това постериорно разпределение (фиг. 67).



Фиг. 67. Пет правдоподобни стойности

Правдоподобните стойности дават информация не само за способностите на ученика, но и за вероятността, с които тези стойности могат да се получат. Например, в горната диаграма, колкото по-високо над правдоподобната стойност е кривата на постериорното разпределение (по-дълги „стрелки”), толкова по-голяма е вероятността за получаване на тази стойност.

Колкото повече правдоподобни стойности се получат за даден ученик, толкова по-добре те описват постериорното разпределение. Практическият опит показва, че ако за всеки ученик се дават по пет правдоподобни стойности, те достатъчно точно оценяват параметрите на популацията. Има различни математически процедури за получаване на постериорното разпределение и за правдоподобните стойности на всеки ученик (Thomas (1993) and Adams and Wu (2002)).

3. Използване на правдоподобните стойности

В четвърта глава обяснихме един начин за оценка на точкови стойности на способностите на учениците – методът, известен като Maximum Likelihood

Estimation (MLE). Съществуват и други начини за такива точкови оценки. Без да се спираме на подробности, ще споменем още два такива широко разпространени метода, които ще използваме за целите по-долу. Единият от тях е Weighted Maximum Likelihood Estimates (WLE), който в известен смисъл коригира MLE за някои статистически измествания. Другият метод е Expected A-Posteriori Estimate (EAP), който на всеки ученик съпоставя като способност средната стойност на постериорното му разпределение.

Всеки от тези три метода дава средното аритметично на извадката като неизместена оценка за средната стойност на популацията. Разликите при оценката на тази средна стойност не са големи. Не стои така въпроса обаче с дисперсията (респ. стандартното отклонение). Дисперсията, пресметната по методите MLE и WLE е по-голяма от истинската, докато по метода EAP тя е по-малка. При това няма подобрение в резултатите, ако големината на извадката расте, но има слабо подобрение когато броят на задачите се увеличава.

Ще илюстрираме това с един пример (Таблица 17). За примера са избрани резултати върху тест от 20 задачи на 2000 случайно избрани ученици от популация с нормализирано нормално разпределение (средната му стойност е 0 и дисперсията е 1). В таблица 17 са дадени оценкте на средната стойност и на дисперсията на популацията по методите MLE, WLE и EAP. Използвани са още и по 5 правдоподобни стойности (PV1,…, PV5) за всеки ученик.

Таблица 17. Различни методи за оценка на средната стойност и дисперсията

Метод	WLE	MLE	EAP	PV1	PV2	PV3	PV4	PV5	Начални стойности
Оценка на средната стойност на популацията	0,003	0,004	0,003	0,004	0,004	0,003	0,004	0,003	0
Оценка на дисперсията на популацията	1,300	1,456	0,778	1,002	1,006	1,000	1,004	1,005	1

Правдоподобните стойности се използват за по-точна оценка на параметрите на *популацията*, но те не са подходящи като информация за способнос-

тите на индивидуалните ученици, понеже са случайно избрани стойности от постериорнитите им разпределения. Да си мислим, например, двама ученика, които са дали едни и същи отговори на задачите на теста, което означава, че те би трябвало да имат една и съща оценка за способностите си. Поради случайния характер на правдоподобните стойности, най-вероятно е обаче тези стойности за двамата ученика да са различни. Това показва, че използването им на ниво ученик не е удачно.

Следващият пример (Таблица 18), взет от данни на изследването TIMSS, също показва, че правдоподобните стойности не са подходящи за описание на индивидуално ниво. В скалата по математика на TIMSS са дадени по 5 правдоподобни стойности (PV1, PV2, PV3, PV4 и PV5) за всеки ученик и въз основа на тях са получени по 5 правдоподобни стойности на извадката. Таблица 18 по-долу показва тези стойности за един ученик, който е подбран така, че разликата между неговите най-голяма и най-малка правдоподобни стойности стойност е „голяма”. За цялата извадка, обаче, разликата между най-голямата и най-малката правдоподобни стойности съвсем не е така голяма.

Таблица 18. Правдоподобни стойности от TIMSS

	PV1	PV2	PV3	PV4	PV5	Max (PV) – Min (PV)
.....	.....	.....	.....	.....	.....	.....
Ученик 1430218	377,74	351,35	398,96	388,38	407,95	56,60
.....	.....	.....	.....	.....	.....	.....
Популацията (8912 ученика)	503,92	504,42	505,12	504,15	504,21	1,20

(Скалата на TIMSS е със средна стойност 500 и стандартно отклонение 100.)

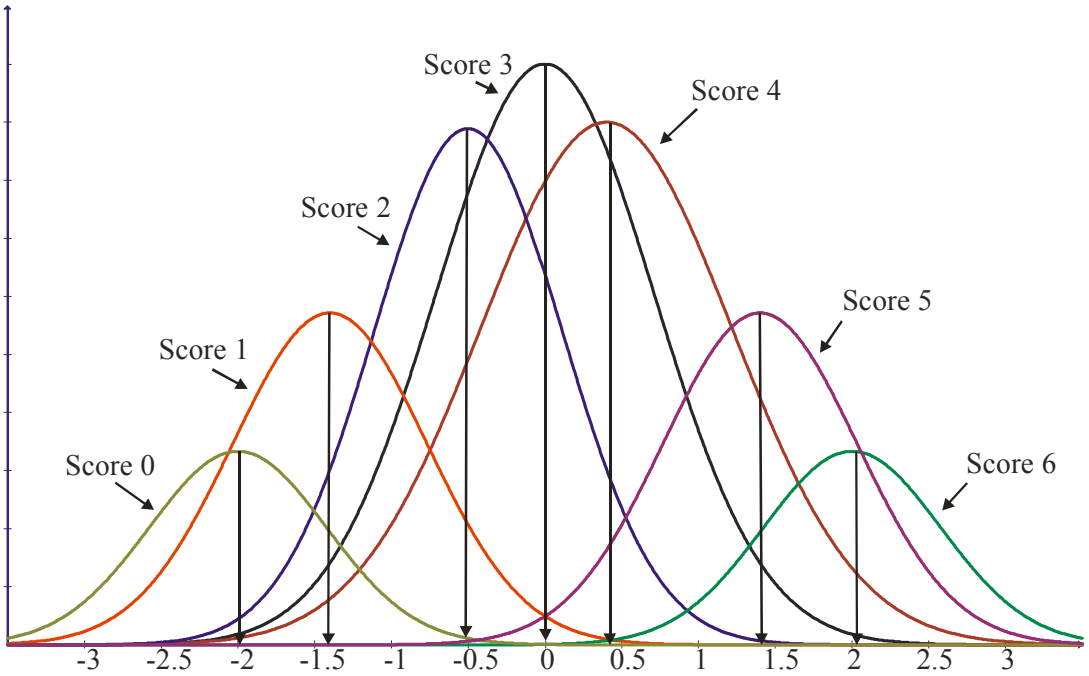
4. Описание на срезове на разпределението на ученическите постижения

В редица случаи при описание на ученическите постижения е важно да се опишат някои срезове, като например процентили или процент ученици, чиито постижения са под даден праг на баловете. По-долу е даден един пример, който показва, че за такива цели използването на правдоподобни стойности дава по-добро описание на ситуацията.

Да разгледаме тест с 6 задачи, оценени с 1=„вярно” и 0=„невярно”. Възможните стойности на скалата на баловете са 7 – всички цели числа от 0 до 6 включително. Фигура 68 показва седемте постериорни разпределения и съот-

ветните им EAP стойности (вертикалните линии, стрелката на които показва съответната стойност в скалата на способностите).

Да предположим, че се интересуваме от процента ученици, чиито способности са под –1,5. Ако използваме EAP стойностите, ще получим, че това е процентът на учениците, които са получили 0 точки на теста. Всъщност, поради дискретния характер на EAP стойностите, същият процент ученици се получава за всеки срез в интервала EAP_0 и EAP_1 (двете най-леви вертикални линии). Подобна е ситуацията и когато се използват WLE и MLE стойностите. Проблемът възниква от факта, че лицето на частта от равнината, заключено от постериорните разпределения за стойности под –1,5, е непрекъсната функция и в това лице влизат части от кривите от всички постериорни разпределения (съответстващи на седемте възможни стойности на баловете).



Фиг. 68. Седем постериорни разпределения

Тъй като правдоподобните стойности са случайно избрани от постериорните разпределения, те дават по-добра оценка за лицето на частта от равнината, заключено от съответното постериорно разпределение под даден праг на способностите. Така те по-добре описват непрекъснатия характер на лицето. С други думи, като се използват правдоподобни стойности, се отчита, че за пресмятането на процента ученици със способности под –1,5 участват не само онези,

които имат бал 0, но и такива, които имат други балове. Така правдоподобните стойности по-добре описват непрекъснатия характер на функцията лице, отколкото използването на дискретните EAP, WLE или MLE. стойности.

5. Използване на правдоподобните стойности при пресмятане на стандартната грешка на измерването

Както вече беше обсъдено в параграф 1, когато измерването на отделните обекти от извадката съдържа „съществена" грешка, тя трябва да се вземе пред вид при пресмятане на стандартната грешка за съответната статистика на популацията. Това допълнително събираемо към стандартната грешка се нарича *измерителна грешка*. Като се използват правдоподобни стойности, измерителната грешка се пресмята по следния начин.

Да предположим, че за всеки обект на извадката са дадени по n правдоподобни стойности PV_1, PV_2,..., PV_n, и тяхната средна стойност е $\overline{PV}$. Тогава измерителната грешка е равна на

$$ME = \sqrt{\frac{n+1}{n(n-1)}\sum_{i=1}^{n}\left(PV_i - \overline{PV}\right)^2} .$$

Стандартната грешка на измерването се пресмята като сбор от извадковата грешка и от измерителната грешка. Ако ползваме Джакнайф метода (седма глава) за грешката от извадката, то стандартната грешка на измерването е равна на

$$\sqrt{\sum_{h=1}^{H}\left(t\left(J_h\right) - t(S)\right)^2} + \sqrt{\frac{n+1}{n(n-1)}\sum_{i=1}^{n}\left(PV_i - \overline{PV}\right)^2} .$$

Правилното пресмятане на стандартната грешка на измерването е важно в редица случаи. Например, големината на тази грешка е от съществено значение за тестване на хипотези. Ще разгледаме пример (Таблица 19), в който неправилно пресметната стандартна грешка на измерването води до противоположно заключение от тестването на хипотеза за наличие на статистически значима разлика между резултатите на момчета и момичета. Примерът е взет от резултатите на TIMSS.

Таблица 19. Сравнение на стандартната грешка

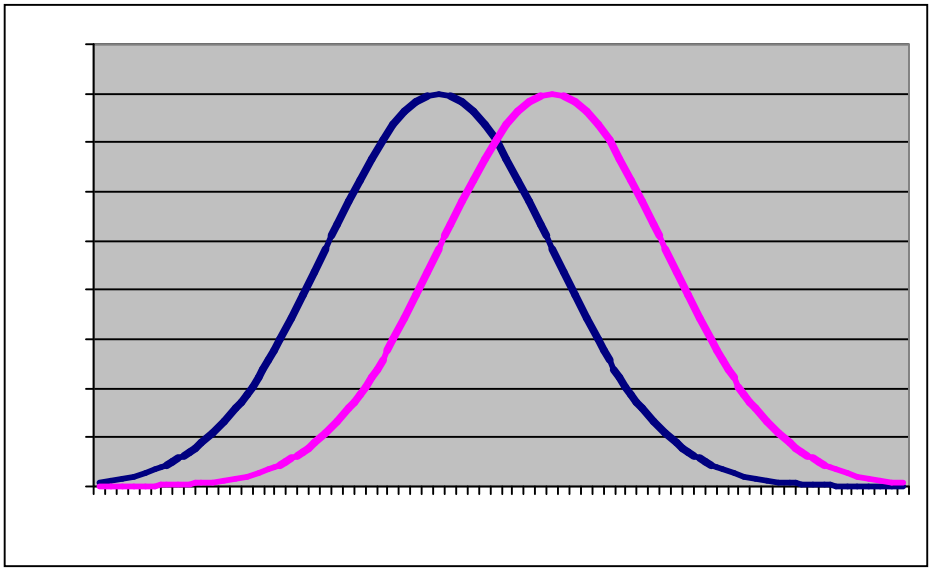
	Средна стойност	**Стандартна грешка на измерването**	
		Джакнайф + Правд. ст.	**Проста случайна извадка**
Общо	491	5,2	1,5
Момичета	495	5,5	2,0
Момчета	487	7,6	2,1
Разлика (момичета – момчета)	8	8,3	2,9

От таблицата се вижда, че ако използваме стандартните методи, предполагайки проста случайна извадка, ще получим статистически значима разлика в постиженията на момичетата и момчетата (стандартната грешка на измерването е малка). Ако вземем пред вид сложността на извадката и методите на правдоподобните стойности, стандартната грешка на измерването е значително по-голяма и в случая води до отхвърляне на хипотезата за статистически значима разлика в постиженията на момичетата и момчетата.

6. Регресионни модели с правдоподобни стойности

При широко мащабните ОДИ наред с тестовете за постижения се събират данни за произхода и семейната среда на учениците, за условията на обучение в училище, за отношението на ученици и учители към изследвания учебен предмет и други фактори, за които се вярва, че оказват влияние върху качеството на обучението и постиженията на учениците. Целта е да се изследва кои от тези предполагаеми фактори действително влияят, в каква посока, с каква сила е тяхното влияние и т.н.

В базата данни за всеки един от тези фактори се конструира по една променлива, която се нарича *условна*. Една от важните цели при анализа на данни от ОДИ е да се намери връзка между някои условни променливи и ученическите постижения. Това обикновено се прави чрез различни регресионни модели, в които условните променливи са независими променливи, а „ученическите постижения” (под формата на някакъв бал) е зависимата променлива. Възможно е регресионният анализ да се прави с правдоподобни стойности, вместо с една единствена стойност на ученическите постижения.



Фиг. 69. Популация, състояща се от две групи

Да разгледаме един пример. При анализите на данни често е интересно да се оценяват статистики на отделни подгрупи ученици от популацията (например подгрупи по отношение на пол, географски район, майчин език и др.) За целта се предполага, че цялата популация се състои от няколко, например две, под-популации („Група 1” и „Група 2”), всяка от които има нормално разпределение на ученическите постижения с една и съща дисперсия σ^2, но с различни средни стойности, например μ_1 and μ_2 (фиг. 69). В този случай, разпределението на ученическите постижения за цялата популация не е нормално, а е съвкупност от две нормални разпределения, което се означава с уравнението $g(\theta) \sim N(\mu + \alpha x, \sigma^2)$, където $x = 0$ ако разглеждаме ученик от Група 1 и $x = 1$ ако разглеждаме ученик от Група 2.

В регресионния модел се прави статистическа оценка за стойностите на μ, σ^2 и α. Да забележим още, че α е разликата между средните стойности на двете групи, т.е. средната стойност на Група 1 е μ (понеже $x = 0$),а средната стойност на Група 2 е $\mu + \alpha$ (понеже $x = 1$). Освен това σ^2 е дисперсията на всяка от двете групи, а не общата дисперсия на популацията.

В този пример променливата x е независима или условна променлива, в смисъла, в който употребихме тези думи по-горе. Тя определя принадлежността на всеки ученик към някоя от разглежданите групи. Може да имаме повече от

една независима (условна) променлива. Тогава уравнението приема вида $g(\theta) \sim N\left(\mu + \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n, \sigma^2\right)$. Нещо повече, не е задължително независимите променливи да са дискретни величини, както беше в разглеждания пример. Те може и да с непрекъснати, като например условната променлива „социално икономически статус" (SES).

Ако правим регресионен анализ с правдоподобни стойности като зависими променливи, редно е моделът, с който са получени правдоподобните стойности, да включва условните променливи. В противен случай статистическите оценки на коефициентите на регресия няма да са точни (обикновено се получават по-малки стойности от действителните). Големината на тази неточност зависи от различни фактори, като например дължината на теста, корелация между условните променливи, както и корелация на някои от тях със зависимата променлива при една и съща стойност на другите.

ЕДИНАДЕСЕТА ГЛАВА
ЙЕРАРХИЧНИ ЛИНЕЙНИ МОДЕЛИ ЗА АНАЛИЗ НА ДАННИ

Събирането на данни за големи национални и международни ОДИ става чрез сложни извадки. Това подробно е описано в трета глава. Тези извадки имат йерархична структура. Най-напред от генералната съвкупност се избират училища, след това от училищата се избират паралелки и (ако е необходимо) от избраните паралелки се избират ученици. Казано по друг начин, извадката предполага, че всеки ученик принадлежи на някаква паралелка, която пък от своя страна принадлежи на някакво училище. Йерархичното линейно моделиране (Hierarchical Linear Modeling – HLM) е специфична техника, базирана на регресионни модели, която се използва за анализ на данни, събрани чрез извадки с йерархична структура (Raudenbush, S. W. & Bryk, A. S., 2002, Snijders TAB & Bokster, RJ., 1999).

Традиционният регресионен анализ, основан на метода на най-малките квадрати (Ordinary Least Square – OLS), често се използва за изследване на връзките на една зависима променлива (например ученическите постижения) и няколко независими такива (например SES – социално икономическия статус на учениците, образованието на родителите, квалификацията на учителите и др.). Основно предположение при регресионния анализ е, че наблюденията на зависимата променлива (постиженията на учениците) са независими, т.е. постиженията на който и да е ученик по никакъв начин не зависят от постиженията на други ученици. Това предположение не е в сила, ако например някои ученици са от едно и също семейство. То не е в сила и когато някои ученици са от една и съща паралелка или училище. Това е така, защото учениците от една и съща паралелка се обучават при едни и същи условия и следователно може да се очаква, че техните постижения са „по-близки", отколкото, ако същите тези ученици биха се обучавали при различни условия (Willms, J. D., 1999.). С други думи, постиженията на всеки отделен ученик в някаква степен зависят от постиженията на групата (паралелката или училището) като цяло, както и от постиженията на отделните ученици в нея. В резултат на това, статистическата извадковата грешка, която се получава от традиционния регресионен анализ е значително по-малка от реалната. Това, от своя страна влияе на резултатите от тестове за статистическа значимост. Един

от начините да се коригира извадковата грешка е да се използва джакнайф метода, описан в девета глава.

Друга възможност за точна оценка на извадковата грешка дава използването на йерархичните линейни модели (HLM). Тези модели предоставят и някои други възможности, които ги правят предпочитани при клъстерни йерархични извадки. Чрез тях може да се прави статистическа оценка едновременно на ефекта на някои клъстерни променливи (например големината на паралелката) върху средното постижение на учениците от клъстера (паралелката), както и на ефекта на индивидуални променливи (например SES) върху него. Например, може да установим каква е разликата в средното постижения на учениците от училищата в малките населени места и тези в големите населени места при положение, че учениците имат един и същ SES. Освен това HLM моделите дават статистическа оценка на реалната дисперсия на регресионните параметри между групите (клъстерите). Например, с тяхното използване може да се отговори на въпроси като: „Колко голяма е разликата между училищата по отношение на средните постижения на учениците в тях и ефекта на SES на учениците и до колко тази разлика може да се обясни с някои от характеристиките на училището?”

В тази глава се описват HLM моделите, тяхното използване за целите на ОДИ и се представят някои резултати от анализи, в които авторът е участвал.

1. Математическо описание на HLM.

Йерархичните модели могат да се разберат най-добре, ако си мислим, че анализът се извършва на няколко степени. (Затова понякога се казва, че се прави многостепенен анализ – Multilevel Analysis.) За целите на образователните изследвания, обикновено са достатъчни две степени, които се наричат съответно ученическа (индивидуална) степен (или ниво) и училищна (или паралелкова) степен (ниво).

На първата степен анализът се прави отделно за всяко училище като се използват данни за отделните ученици. Математически това се описва с уравнения за регресионен анализ по отношение на зависимата променлива (например ученическите постижения) за всяко училище. Например, ако в извадката имаме m училища, първата степен на анализ се описва с уравненията

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij}^{(1)} + \beta_{2j}X_{ij}^{(s)} + ... + \beta_{sj}X_{ij}^{(s)} + r_{ij}, \quad j = 1, 2, ..., m$$

където Y_{ij} е стойността на зависимата променлива (постижението) на ученика i в училището j, $X_{ij}^{(1)}$, $X_{ij}^{(2)}$,..., $X_{ij}^{(s)}$ са стойностите на някои избрани независими променливи, с които искаме да обясним ученическите постижения, на ученика i в училището j, $\beta_{0j}, \beta_{1j}, ..., \beta_{sj}$ са коефициентите на регресия и r_{ij} е резидиума (отклонението на ученика i в училището j от правата на регресия). Променливите $X^{(1)}$, $X^{(2)}$,..., $X^{(s)}$ са на ученическо ниво, т.е. такива, които описват отделните ученици, например пол, възраст, SES, образование на родителите и др.

На втората степен се извършва регресионен анализ, в който зависими променливи са (някои или всички) регресионни коефициенти от първата степен, а за независими променливи се избират такива, които описват учебната среда на ниво училище (или паралелка).

$$\beta_{kj} = \gamma_{k0} + \gamma_{k1}Z_j^{(1)} + \gamma_{k2}Z_j^{(2)} + ... + \gamma_{kt}Z_j^{(t)} + u_{kj}, \quad k \in \{1, 2, ..., m\}$$

където β_{kj} са някои от регресионните коефициенти от първата степен на анализа, $Z_j^{(1)}$, $Z_j^{(2)}$,..., $Z_j^{(t)}$ са стойностите на някои избрани независими променливи за училището с номер j, $\gamma_{k0}, \gamma_{k1}, ..., \gamma_{kt}$ са коефициентите на регресия и u_{kj} е резидиума (отклонението на училището j от правата на регресия). Променливите $Z^{(1)}$, $Z^{(2)}$,..., $Z^{(t)}$ са на училищно (или паралелково) ниво, т.е. характеристики на училището, като например големината на училището (или паралелката), мярка на дисциплината в училището (или паралелката), големината на населеното място, в което се намира училището и др.

Анализът във втората степен дава възможност да се оцени големината на ефекта, който оказва всяка от променливите $Z^{(1)}$, $Z^{(2)}$,..., $Z^{(t)}$. Например, ако регресионното уравнение за втората степен е

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j^{(1)} + \gamma_{02}Z_j^{(2)} + u_{kj},$$

като $Z^{(1)}$ е големината на паралелката (брой ученици) и $Z^{(2)}$ е мярка за дисциплината, то γ_{01} показва нарастването (или намаляването) на средното постижение на паралелка при промяна на големината на класа с 1 ученик, а γ_{02} по-

казва същата промяна за една единица промяна на мярката в дисциплината на паралелката.

Както се вижда от математическото описание, основната идея на йерархичните модели е да се прави отделно анализ за всяка единица на йерархичната стълба. В резултат на обединението на регресионните уравнения за двете степени се получава сложен математически модел, който точно отговаря на структурата на извадката. Статистическата оценка на коефициентите на регресионните уравнения се извършва със специализиран софтуер.

2. Използване на йерархичните линейни модели за мониторинг на образователната система

В последното десетилетие йерархичните линейни модели намират голямо приложение за мониторинг на образователните системи като цяло или отделни техни елементи (региони, училища) (Raudenbush, S.W., & Willms, J.D., 1995., Martin, M. O., et al., 2000, Willms, J. D., 1992, Willms, J. D., 1995, Willms, J. D. & Marie-Andree Somers, 2001 и др.). Един от сериозните индикатори, който се използва за мониторинг на училищата (или регионите) е средната стойност на ученическите постижения. Той обикновено се получава чрез сравняване на постиженията, получени от специално измерване (тест). При това, винаги има училища (региони), чието средно постижение е над средното за страната и такива, за които то е под средното за страната. Основният проблем тук е, че понякога не се взема пред вид, че различните училища (региони) работят при различни условия. С други думи, ако искаме да използваме данни за средните постижения на учениците по училища (региони) трябва да вземем пред вид още поне три основни фактора за средните постижения на учениците: (1) те зависят от основните характеристики на учениците в училищата (регионите), като семейна среда, SES, предварително натрупани знания и др.; (2) те са базирани на фиксирани стандарти; (3) те зависят от промените на училището (региона), които стават с времето (Willms, J. D., 1992).

Голяма част от изследванията в тази област са насочени към (някои от) следните четири основни въпроса:

1. В каква степен училищата (или регионите) се различават в зависимост от средната стойност на ученическите постижения?

2. В каква степен ученическите постижения се различават в зависимост от статуса на учениците.
3. Какви образователни политики и практики повишават нивото на ученическите постижения?
4. Какви образователни политики и практики намаляват различията в постиженията на учениците с различен статус?

Тези въпроси са подходящи за изследване с HLM моделите, защото се вписват в тяхната структура.

Първият въпрос се отнася до *нивото* на образователната система. В най-прост вид той изисква сравнение на средните постижения на учениците по училища (или региони). В разширена форма може да се запитаме: „В каква степен училищата (или регионите) се различават в зависимост от средната стойност на ученическите постижения след като се вземат пред вид основни ученически характеристики, като семейна среда, SES, предварително натрупани знания и др.?"

Вторият въпрос е за различията в постиженията в контекста на образователната система. Използва се терминът „*градиент*" за описание на зависимостта на ученическите постижения от някакъв фактор, например SES. Ако градиентът е стръмен, това означава, че в образователната система има сериозна разлика в постиженията на учениците в зависимост от техния (SES) статус, т.е. учениците с различен статус показват сериозни различия в постиженията. Подобни различия на постиженията може да се изследва и по отношение на разликата в пола, етническата принадлежност и др. Стремеж на образователните системи е да направят градиента по-малко стръмен.

Третият и четвъртият въпрос се отнасят до това *защо* някои училища имат по-добри (по-лоши) успехи от други и *защо* някои училища по-добре се справят с понижение на градиента си. На тези въпроси се отговаря по-трудно, защото те изискват сериозно познание на дейността на училищата като цяло и на всяко отделно училище.

3. Изследване на училищната ефективност по математика в рамките на проекта TIMSS

Въпросите за училищната ефективност са важни при вземане на решения за подобряване на качеството на образованието. Затова през последните години се обръща голямо внимание на изследвания в тази област. Те имат за цел да

изучат начините, по които училището изпълнява ролята си на институция за обучение и кои са факторите, от които зависи успешното изпълнение на тази роля.

На пръв поглед нещата изглеждат прости: ефективни са онези училища, чиито ученици имат високи постижения. Както обяснихме по-горе обаче, постиженията на учениците зависят от много и различни фактори и те трябва да се вземат, предвид когато се оценява ефективността на училището. Например, училища с голям процент ученици, които имат предварителна подготовка в семейството си, с добро финансово подпомагане от общините и с налични учебно-технически средства са в сравнително добра позиция. Училища, които нямат тези предимства, срещат значително по-големи трудности в обучението. Изследванията на училищната ефективност се стремят да разделят факторите от организационно и педагогическо естество от тези, отнасящи се до подготовката и възможностите на учениците, които те имат *a priori*, т.е. преди тяхното постъпване в училище.

Нагледно казано, може да се окаже, че училище *A* е по-ефективно от училище *B*, въпреки че средните постижения на учениците от училище *B* са по-добри. Причина за това може да е, че при първоначално дадените условия, училище *A* има по-ефективни начини и процедури за обучение, отколкото училище *B*. С други думи, училище *B* има много по-добри дадености *a priori* и разчита на тях, докато училище *A* успешно се справя с неудобствата на не особено добрите първоначални условия.

В рамките на проекта TIMSS-99 е изследвана училищната ефективност като се използва специална методика и HLM модели (Martin, M. O. et al, 2000).

Първата стъпка при анализа на данни е да се отделят онези променливи, които имат значителна корелация с ученическите постижения. Някои еднотипни променливи могат да се обединят в така наречените индекси. Това става чрез анализ на главните компоненти. След това училищата във всяка участваща държава се подреждат по средната стойност на ученическите постижения по математика. Училищата в горната една-трета образуват групата на „силните” училища, а тези в долната една-трета – групата на „слабите” училища. Идеята е да се провери колко добре избраните променливи и индекси разграничават „силните” от „слабите” училища.

За всяка от избраните променливи или индекси се фиксира точка (стойност на променливата), която най-добре разграничава двете групи училища. След това за всяка държава с t-тест се проверява дали има статистически значима разлика между двете групи училища и честотата на съответната променлива или индекс. Избират се само онези променливи и индекси, за които разликата е значима за повечето държави. Например, във въпросника учениците са дали информация за броя на книгите в къщи, като са посочили един от следните отговори „от 0 до 10 книги", „от 11 до 25 книги", „от 26 до 100 книги", „от 101 до 200 книги" и „повече от 200 книги". Точката (стойността на променливата), която най-добре разграничава двете групи училища, е „100 книги". След това за всяка държава се анализира разликата в процента ученици от „слабите" училища, имащи в къщи поне 100 книги, и този от „силните" училища.

Определянето на фактори, които характеризират „силните" и „слабите" училища, е важно за по-задълбочен анализ. Както вече беше споменато, средното постижение на учениците от дадено училище зависи от началните дадености на учениците. Ако тези начални дадености биха били еднакви, тогава ефективността на училището би се измервала само с ученическите постижения. Различните дадености, обаче, трябва да се вземат пред вид, когато се оценява приносът, който училището има в обучението на учениците. Затова следващата стъпка е прилагане на многостепенен анализ, за да се изследва връзката на някои ученически и училищни фактори с постиженията на учениците, като едновременно се вземат пред вид различията, дължащи се на началните дадености на учениците.

Използвани са три двустепенни йерархични линейни модела. Първият модел изследва как ученическите постижения се различават между училищата за всички държави. Вторият установява каква част от различията в ученическите постижения се дължи на различия в семейната среда. Третият модел изследва как различията в някои битови и училищни фактори влияят на ученическите постижения, като се вземе предвид семейната среда на ученика. Следва математическото описание на трите модела.

<u>Първи модел.</u> *Училищно ниво:* $Y_{ij} = \beta_{0j} + e_{ik}$, където Y_{ij} е постижението (брой точки) на *i*-тия ученик в *j*-тото училище; β_{0j} е средната стойност на

постиженията на учениците в *j*-тото училище; e_{ik} е нормално разпределена случайна грешка с постоянна дисперсия.

Между училищно ниво: $\beta_{0j} = \gamma_{00} + u_{0j}$, където β_{0j} е средната стойност на постиженията на учениците в *j*-тото училище; γ_{00} е международната средна стойност на ученическите постижения; U_{0j} е нормално разпределена случайна грешка с постоянна дисперсия.

<u>Втори модел.</u> *Училищно ниво:* $Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ik}$, където β_{0j} е пресечната точка на регресионната права между семейната среда и постиженията на учениците в *j*-тото училище; β_{1j} е наклонът на тази регресионната права; X_{ij} е индексът на семейната среда на *i*-тия ученик в *j*-тото училище.

Между училищно ниво: $\beta_{0j} = \gamma_{00} + \gamma_{1j} Z_{1j} + u_{0j}$, където γ_{1j} е коефициентът на регресия на училищно ниво; Z_{1j} е средната стойност на индекса на семейната среда за учениците от *j*-тото училище.

<u>Трети модел.</u> В него са направени седем варианта за анализ, а именно:

1. Характеристики на паралелката – използвани са следните пет характеристики за моделиране на ученическите постижения: честота на домашната работа по математика, обем на домашната работа по математика, проверка на домашната работа от учителя, социален климат в паралелката и брой на учениците в паралелката.
2. Вариант 1 плюс характеристики на учителя – към характеристиките от първия вариант е добавен опитът на учителя по математика.
3. Вариант 2 плюс социален климат в училище – добавени са факторите административни нарушения от учениците и лошо поведение.
4. Вариант 3 плюс географско положение на училището – взема се предвид още географския район на училището (урбанизиран, селски или изолиран).
5. Вариант 4 плюс фактори като лична амбиция на ученика за по-високо образование и родителски натиск за това.
6. Вариант 5 плюс фактора семейна среда.
7. Само фактора семейна среда.

Уравнението за *училищното ниво* е същото като при втория модел. *Между училищното ниво* се моделира с уравненията $\beta_{0j} = \gamma_{00} + \gamma_{1j}Z_{1j} + \gamma_{2j}Z_{2j} + ... + u_{0j}$ и $\beta_{1j} = \gamma_{10}$, където γ_{1j}, γ_{2j}, … са коефициентите на регресия на училищно ниво и Z_{1j}, Z_{2j}, … са средните стойности на съответните променливи или индекси за *j*-тото училище.

Ето някои от основните резултати (Банков, К., 2003). Във всички държави променливите и индексите, които най-добре разграничават „силните" и „слабите" училища, са свързани със семейната среда. Те са социално-икономическият статус на семейството и родителската подкрепа за по-високо образование. Учениците от „силните" училища обикновено имат много книги в къщи, по-добри учебни пособия, по-високо образовани родители. Те самите имат високи амбиции за образование, желание да постъпят в университет и са силно подкрепяни от родителите си в тези стремежи.

Оказва се, че факторите, свързани с училището, по-слабо разграничават „силните" от „слабите" училища. От значение се оказват големината и географският район на училището, социалният климат в него и ангажираността на учениците в математически дейности. Интересно е, че за различните държави тези фактори имат различна разграничителна сила.

Резултатите от многостепенния анализ показват, че силата, с която училищните фактори влияят на ученическите постижения, варира много в различните държави. Също така голяма вариативност има в различията между училищата по отношение на семейната среда на учениците. Следователно тези фактори са свързани със социалния живот и бита на конкретната държава и това трябва да се взема пред вид при всеки опит за сравняване между държави.

Сравнително малко характеристики на паралелката оказват влияние върху постиженията на учениците. Най-значимата от тях е редовното задаване на домашна работа. Учениците, които прекарват редовно голяма част от времето върху домашна работа, имат по-високи постижения, дори и след като се „премахне" влиянието на семейната среда. Факторите, свързани с характеристиките на учителя, социалния климат в паралелката, както и нейната големина, оказват по-слабо влияние върху ученическите постижения.

Описаният тук анализ е малка част от усилията, които се полагат, за да се разбере кои фактори оказват положително влияние върху динамичния процес на

обучението. Вероятно заслужават внимание анализи, свързани с културните различия между отделните държави и влиянието на тези различия върху ученическите постижения.

4. Изследване на SES градиента на България с данните от TIMSS-2003

Практиката е доказала, че учениците, чиито родители имат по-висока образователна степен, имат по-престижни професии и по-добро заплащане, показват по-добри постижения от учениците с „по-слаби" семейни характеристики (Willms, J.D., 2002). Това означава, че има съществена зависимост на ученическите постижения от социално икономическия статус на ученика (SES). Този статус обикновено включва три основни характеристики: (1) образователната степен на родителите; (2) професията на родителите; (3) доходът на семейството.

Тази зависимост отдавна е обект на изследване в социологията на образованието (White, K.R., 1982.). Използвани са различни методи за изследване, като се започне от дескриптивна статистика и корелация между ученически постижения и SES и се стигне до използване на HLM моделите. По-долу ще изследваме така наречения SES градиент за България, получен с данни от TIMSS-2003, за да отговорим, до колкото е възможно, на следните въпроси:

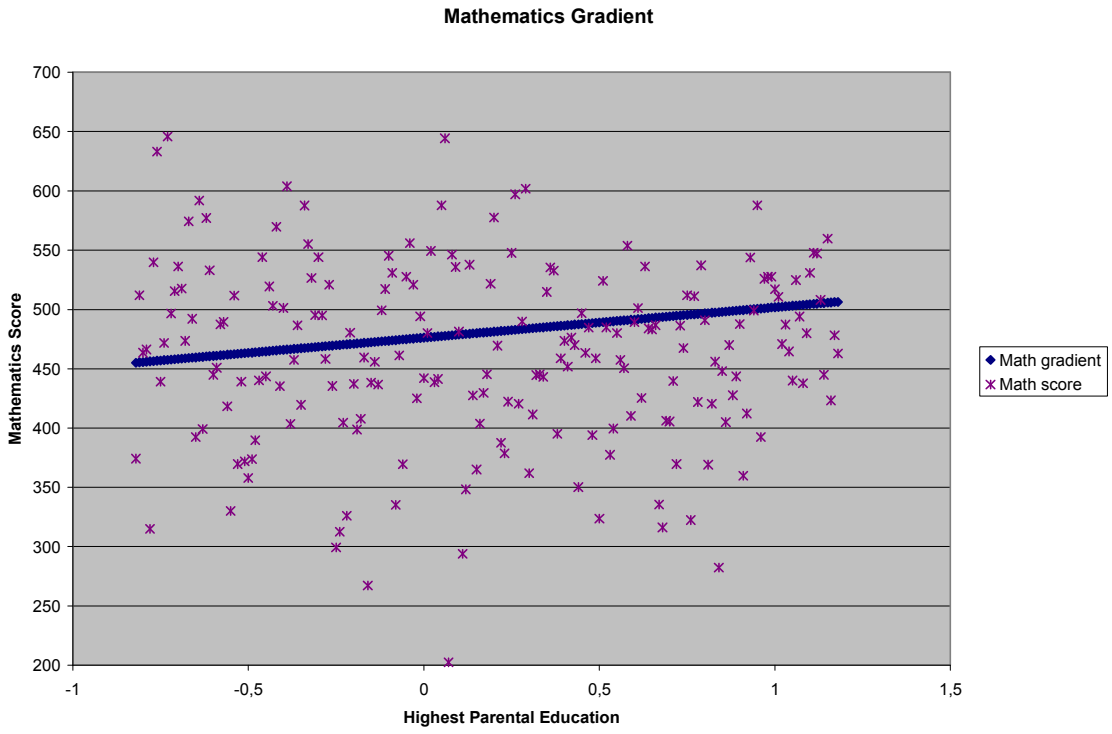
1. В каква степен се различават регионите в България по отношение на постиженията в TIMSS-2003?
2. В каква степен се различават постиженията в TIMSS-2003 за ученици с различен SES статус?
3. В каква степен се различават постиженията в TIMSS-2003 за ученици от различните видове училища (обикновени и профилирани гимназии).

За целта използваме административното разделяне на България на седем региона, а именно: 1-Северен централен (Ловеч), 2-Северозападен (Монтана), 3-Североизточен (Варна), 4-Югозападен (без София), 5-Югоизточен (Бургас), 6-Южен централен (Пловдив), 7-град София.

Социално икономически (SES) градиент е крива, която описва връзката на ученическите постижения със социално икономическия статус (SES) на учениците от дадена група (държава, регион и др.). Тук ще разглеждаме ученическите постижения (по математика или по природни науки) в TIMSS-2003. От трите характеристики на социално икономическия статус в TIMSS-2003 има

данни само за първата от тях. Затова под SES ще разбираме образователната степен на родителите.

В TIMSS-2003 се използва вероятностното моделиране, за да се конструират две скали за постижения: по математика и по природни науки (Gonzalez, E.J., Galia, J., & Li, I., 2004.). На фиг. 70 е показан SES градиента за България по математика. За описание на образователната степен се използват международните ISCED степени (http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm). Променливата „най-висока образователната степен на родителите" е центрирана (нулата е) в средната стойност за България. Градиентът е построен от 10-тия до 90-тия процентил за сраната. Използван е регресионен анализ с метода на най-малките квадрати.



Фиг. 70

Три са важните характеристики на градиента:

• *Ниво* – това е постижението на учениците със средна стойност на независимата променлива SES, в случая стойност 0 за SES, понеже така е центрирана тази скала. Нивото е характеристика за средното постижение на групата уче-

ници, като се вземе пред вид техния SES статус. Нивото на градиента за България по математика е 476,17, а за природни науки той е 478,84.

• *Наклон* – по-големият наклон говори за по-голяма зависимост на постиженията от SES статус на учениците, т.е. учениците с различен статус показват сериозни различия в постиженията. Наклонът на градиента за математика за България е 25,63, а за природни науки той е 13,94. Това означава, че повишаването на SES на ученика с една стандартно отклонение на SES скалата (което 0,88), води до увеличаване на постиженията по математика средно с 25,63 и на тези по природни науки – средно с 13,94. (Една единица на SES е с единица по-висока образователна степен според ISCED степените.)

• *Сила* – това е мярка за степента, с която отделните стойности на ученическите постижения варират над или под графиката на градиента. Големи стойности за силата показват силна връзка между ученическите постижения и SES статуса на ученика и обратно. Обикновено силата се измерва с така наречената R квадрат статистика. За България тя е 8% и 2% съответно за математика и природни науки.

Една от основните задачи на образователните системи е да повиши нивото и да намали наклона на градиента.

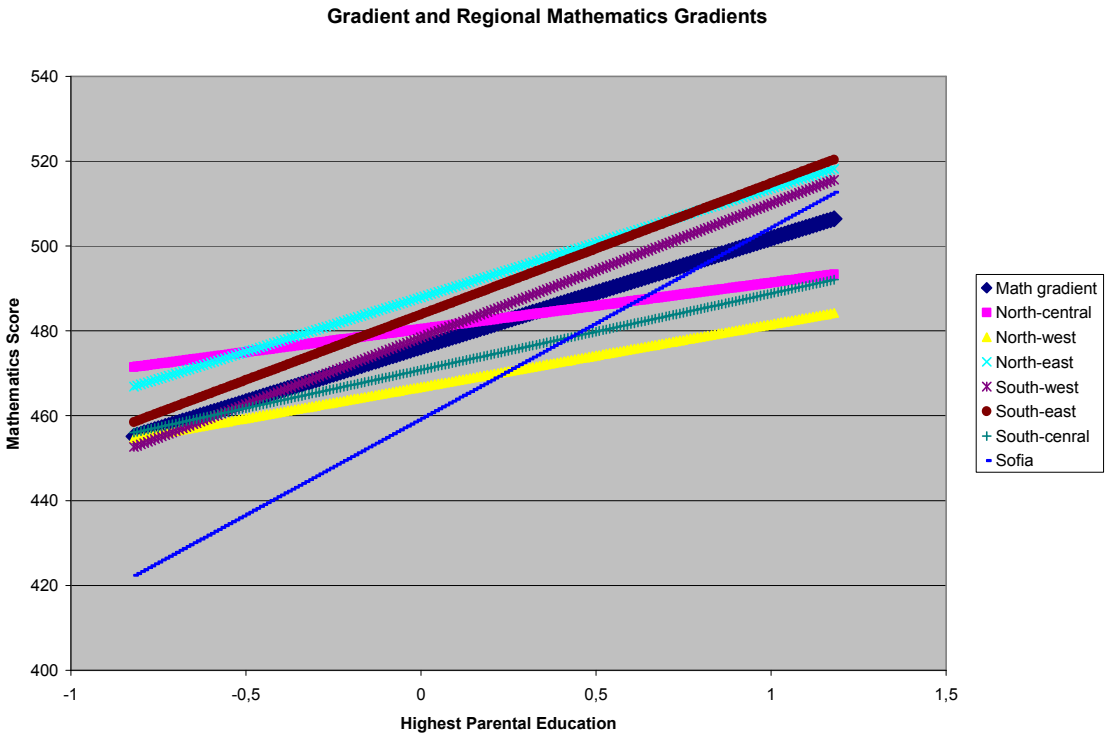
Променливата „най-висока образователна степен на родителите" е важна за разбирането на анализите. Затова в таблица 20 са дадени нейните основни статистики за цялата страна, за отделните региони и за двата вида училища (профилирани и не профилирани).

Таблица 20. Статистики на „най-висока образователна степен на родителите".

	България	**Видове училища**					
		Не профилирани			**Профилирани**		
Mean	0.00	-0.15			0.70		
Median	0.18	0.00			1.18		
Mode	0.18	0.18			1.18		
St. dev.	0.89	0.86			0.67		
	Региони						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Mean	-0.11	0.06	0.00	0.00	0.00	0.00	0.23
Median	0.00	0.18	0.18	0.18	0.18	0.18	0.18
Mode	-0.82	0.18	0.18	0.18	0.18	0.18	1.18
St. dev.	0.86	0.82	0.93	0.85	0.97	0.83	0.93

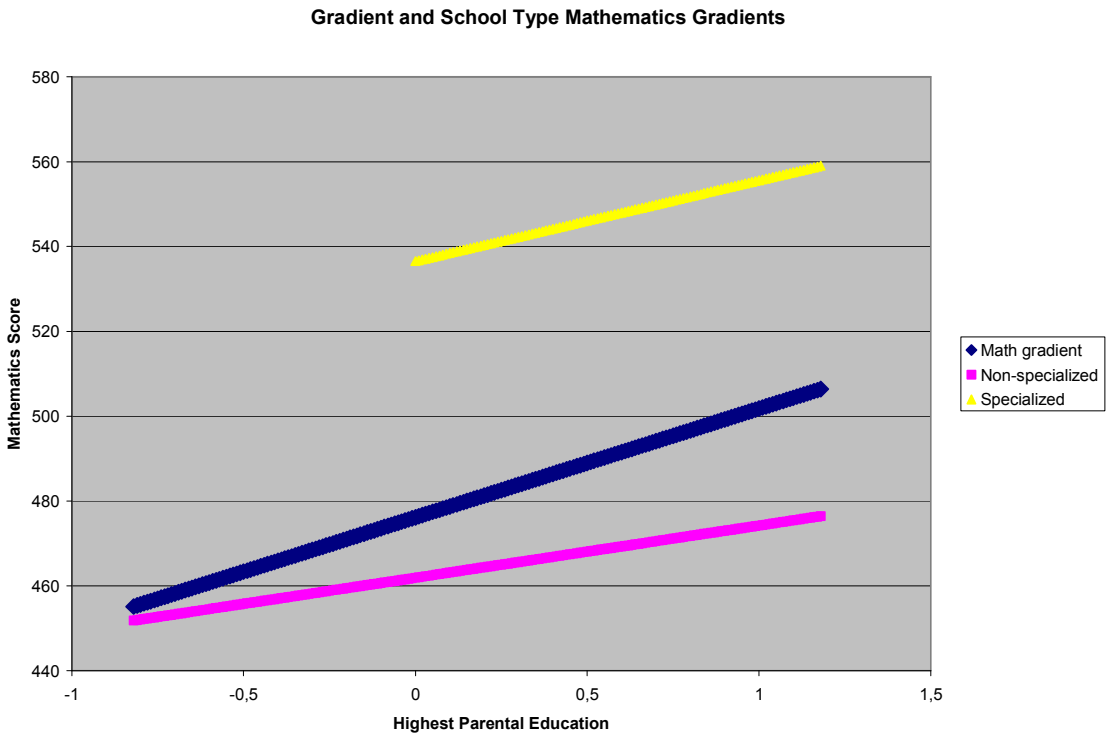
1-Северен централен (Ловеч), 2-Северозападен (Монтана), 3-Североизточен (Варна), 4-Югозападен (без София), 5-Югоизточен (Бургас), 6-Южен централен (Пловдив), 7-София.

Полезно е да се изследва градиента на различни групи ученици. Например, фигура 71 показва градиентите по математика за седемте региона в България, както и градиента за цялата страна. Най-малък е наклонът на градиента в Северния централен регион (Ловеч). Най-високо е нивото на градиента в Североизточния регион (Варна), а най-ниско – това в София.



Фиг. 71

На фиг. 72 е показан градиента по математика за профилираните и не профилираните училища. Прави впечатление, че 10-тия процентил на SES скалата за профилираните училища е по-висок от този за не профилираните и за цялата страна (последните два са равни). Както може да се очаква, нивото на градиента за профилираните училища е много по-висок от тези на не профилираните. Наклонът му също е по-висок.



Фиг. 72

Таблица 21 съдържа данни за коефициентите на регресия за градиента за страната, отделните региони и двата типа училища.

Таблица 21. Коефициенти на регресия за градиента

	България		Видове училища			
			Не профилирани		Профилирани	
	M	S	M	S	M	S
Intercept	476.17 (1.20)	478.84 (1.37)	461.92 (1.27)	477.59 (1.41)	536.58 (3.66)	487.74 (6.12)
Slope	25.63 (1.36)	13.94 (1.54)	12.31 (1.46)	13.28 (1.62)	19.05 (377)	8.78 (6.30)

	Региони													
	1		2		3		4		5		6		7	
	M	S	M	S	M	S	M	S	M	S	M	S	M	S
Intercept	480.44 (3.26)	498.99 (3.46)	466.79 (4.21)	484.84 (4.93)	487.91 (2.81)	477.22 (3.26)	478.45 (3.34)	482.20 (3.73)	483.88 (4.19)	475.24 (4.48)	470.79 (2.18)	482.80 (2.44)	459.38 (3.56)	450.47 (4.20)
Slope	10.92 (3.76)	–0.41 (3.99)	14.76 (5.12)	9.14 (6.01)	25.67 (3.01)	19.58 (4.50)	31.51 (3.93)	12.63 (3.85)	30.93 (4.33)	10.52 (4.63)	18.01 (2.61)	15.10 (2.92)	45.14 (3.71)	28.32 (4.38)

Стандартната грешка е в скоби. M – Математика, S – Природни науки. 1-Северен централен (Ловеч), 2-Северозападен (Монтана), 3-Североизточен (Варна), 4-Югозападен (без София), 5-Югоизточен (Бургас), 6-Южен централен (Пловдив), 7-София.

С данните за България за TIMSS-2003 са направени анализи с няколко двустепенни HLM модела. По-долу ще систематизираме резултатите от тях.

С така наречения „нулев модел” (когато регресионните уравнения не съдържат независими променливи) може да се определи в каква част различията в ученическите постижения се дължат на различията в училищата на образователната система и в каква част те се дължат на различията на отделните индивиди. Данните са показани в таблица 22.

Таблица 22. Коефициенти на паралелкова корелация

		България	Региони						
			1	2	3	4	5	6	7
Паралелкова корелация	Математика	0.39	0.32	0.19	0.46	0.37	0.55	0.25	0.60
	Природни науки	0.37	0.31	0.27	0.49	0.44	0.40	0.20	0.45

1-Северен централен (Ловеч), 2-Северозападен (Монтана), 3-Североизточен (Варна), 4-Югозападен (без София), 5-Югоизточен (Бургас), 6-Южен централен (Пловдив), 7-София.

Коефициентът на паралелкова корелация в TIMSS-2003 за страната е 0,39 за математика и 0,37 за природни науки. Това означава, че около 39% от различията в постиженията на учениците по математика се дължат на различията в училищата и същото важи за около 37% за различията в постиженията по природни науки. Високата стойност на този коефициент показва, че ако анализите на резултатите за България се правят с традиционните статистически методи, ще се допусне значителна грешка. Затова трябва да се използват методи, които правилно отчитат, че извадката съществено се различава от проста случайна извадка.

Значителна е разликата в коефициента на паралелкова корелация между регионите. Той е много висок за София (голямо е значението на училището за постиженията на учениците) и е сравнително нисък за Северозападния регион (Монтана).

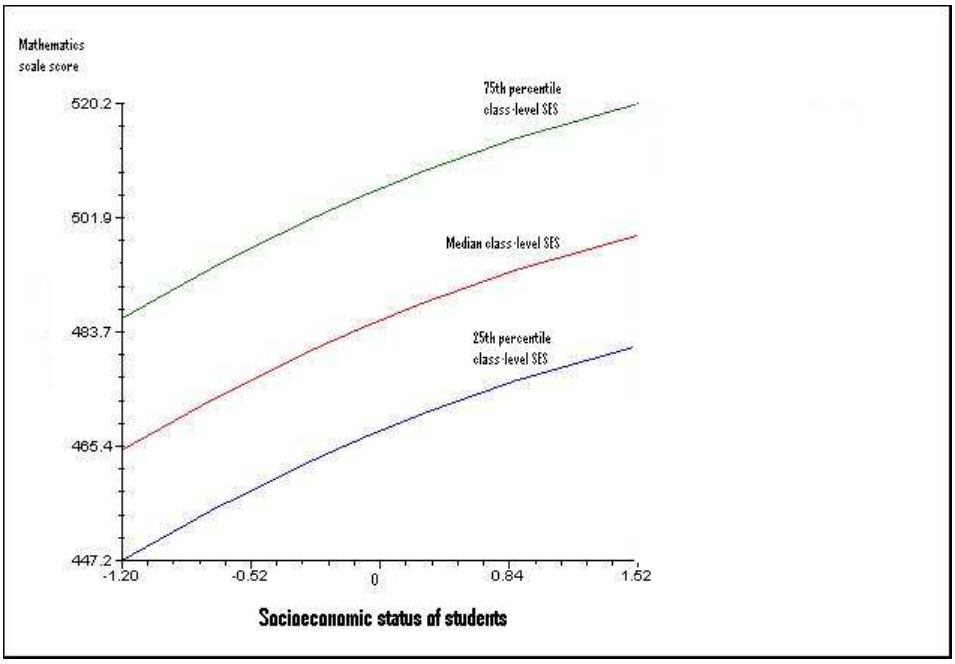
Резултатите показват, че няма статически значима зависимост между нивото и наклона на градиента. Това означава, че влиянието на SES статуса върху постиженията на група ученици не зависи от това дали групата (като цяло) има високи или ниски постижения.

Голямо е значението на средната стойност на степента на образование на родителите върху средното постижение на учениците в училището. Ако средната степен на образование на родителите за училище се увеличи с 1 единица по

ISCED скалата, очакваното повишение на постиженията по математика е с 46 единици.

Съществени за постиженията по математика са също така следните училищни фактори: големината на населеното място, „климата” в училище, ресурсите на училището за обучение.

Използван е и модел, в който независими променливи за характеристика на училището са: вида населено място (според брой жители), вида училище (профилирано или не) и региона, в който е училището (Bankov, K., D. Mikova, T. Smith, 2006). Резултатите показват, че учениците от профилираните училища, както и тези, които са от училищата с по-малък брой жители (по-малък от 3000 жители), имат по-високи резултати. Например, осмокласниците от профилираните училища имат с малко повече от едно стандартно отклонение по-висок резултат от тези от не профилираните училища. Учениците от „малките” населени места имат с 24% до 38% от стандартното отклонение по-високи постижения от тези от „големите” населени места, при положение, че вида училище и региона не се променят. Това е интересно, макар и странно явление, което заслужава допълнително изследване.



Фиг. 73

Има статистически значима разлика в постиженията по математика на учениците от Северния централен регион (Ловеч) и София (с около 33% от стандартното отклонение).

Както може да се очаква, ученическите постижения нарастват с нарастване на средната стойност на SES статуса за училището. Това означава, че постиженията са високи в училища с концентрация на ученици с висок SES статус (фиг. 73). Казано с други думи, учениците с нисък SES статус ще имат сериозни предимства за обучение, ако се учат в училища с висок среден SES статус.

Зависимостта на ученическите постижения от характеристиките на учителите (пол, степен на образование, основен предмет и др.) е слаба. Това означава, че за България не могат да се приложат изводи за мониторинг на образователната система, които включват характеристиките на учителите като съществен фактор за повишаване на ученическите постижения.

ДВАНАДЕСЕТА ГЛАВА
ОПРЕДЕЛЯНЕ НА СКАЛИ, НОРМИ И НИВА НА ПОСТИЖЕНИЯ

В четвърта глава обяснихме, че измерването е процес, при който на всеки един от дадено множество обекти или явления се съпоставя число по определено правило. При измерване с тестове, на всеки ученик се съпоставя число, като резултат от решението на задачите в теста. Това число по подходящ начин описва постиженията на ученика, които се измерват. Така, постиженията на групата ученици, с която е проведен теста, образуват множество от числа. За да се извършват смислени дейности с това множество, трябва по някакъв начин да се даде определен смисъл на числото 0 (началото на измерването), числото 1 (единичната мярка) и на някои операции, които могат да се извършват с числата от това множество. Казва се, че така се определя скалата на тестовите балове. На практика се използват най-различни скали. Тук са разгледани най-важните и най-разпространените от тях.

Тестовите норми служат, за да изразят баловете от числова скала по някакъв общоприет начин за оценяване на ученическите постижения. За Българската образователна система този „общоприет" начин е „поставянето на оценка" в така наречената „шестобална система" и използване на категориите „слаб", „среден", „добър", „мн. добър" и „отличен".

При отчитане на резултатите от тест за постижения основните акценти са предимно от количествено естество. Например, сериозно внимание се отделя, за да се сравнява представянето на отделни групи ученици; да се разпределят учениците по отношение на зададени норми или критерии и др. Този тип анализи са сравнително ясни, защото се основават на широко използвани статистически методи и процедури. Значителни трудности се наблюдават, обаче, при опити за качествени анализи на скалите за постижения. Например, не е лесно да се опише в контекста на предметната област на теста какви знания и умения имат учениците, чиито постижения са в определена точка от скалата, както и какви са техните основни пропуски.

Въпросите от качествено естество са важни за по-доброто разбиране на скалата на даден тест за постижения, както и за правилната интерпретация на резултатите от гледна точка на предметната област на теста. Тук се разглежда един метод за качествен анализ на скалата за постижения: *метода на опорните*

точки. Той дава възможност да се определят нива на постижения на учениците и да се опише в контекста на предметната област на теста какво могат да правят учениците във всяко едно от тези нива. Методът е разработен и приложен за първи път за нуждите на американската организация National Assessment of Educational Progress (NAEP) (Beaton, A. E. & Allen, N. L., 1992). Използван е и за анализ на скалите в международното изследване TIMSS (Martin, M. O., et al., 2004). В тази глава авторът дава математическо описание на модела и го прилага за българските резултати от TIMSS (Банков, К., 2002-А.).

След като са известни резултатите от някакво оценяване (тест), възникват въпроси, като например: Как да разпределим учениците в „нива на знания" от типа на „ученици с много ниски постижения", „ученици с ниски постижения", „ученици с високи постижения", „ученици с много високи постижения" (или подобни)? Ако все пак сме успели по „разумен" начин да направим такова разпределение, как да опишем какво могат (какво са постигнали) в смисъла на предметната област на теста, учениците от всяко ниво? Това са основните въпроси, на които може да се отговори с метода на опорните точки.

1. Скала на суровия тестов бал

Нека всяка тестова задача в някакъв тест се оценява с 1 точка при правилен отговор (решена задача) и с 0 точки, ако ученикът не е дал (посочил) правилния отговор (не е решил задачата). След като тестът е проведен, сборът от точките на всеки ученик се нарича *суров тестов бал*. Този сбор е цяло число между 0 и броя на задачите в теста. Смисълът на суровия тестов бал е лесно разбираем: за всеки ученик той е равен на броя на задачите от теста, които ученикът е решил правилно.

Суровият тестов бал за групата ученици, с която е проведен теста, е честотно разпределение на получените от тях точки. С това честотно разпределение се изчислява средната стойност на групата, дисперсията, стандартното отклонение, правят се графични представяния на данните и т.н. Със суровия тестов бал се анализират измерителните качества на задачите от теста, изчислява се надеждността на теста и други негови статистически характеристики.

Със суровия тестов бал може да се сравняват постиженията на ученици, правили един и същ тест – счита се, че този, който има по-висок суров тестов

бал „знае” повече (по тематиката на теста), защото е решил повече задачи. (Нужно е да се направят съответните тествания за статистическа значимост, които не са обект на този труд.)

Суровият тестов бал е измерване в интервална скала, което може да се трансформира линейно, т.е., ако имаме балове, означени с X, може да се премине балове, означени с Y по формулата $Y = aX + b$, където $a \neq 0$ и b са някакви числа. Да отбележим, че при такава трансформация между средните стойности на X и Y има следната зависимост: $\mu_Y = a\mu_X + b$. За дисперсията и за стандартното отклонение зависимостите са съответно $\sigma_Y^2 = a^2\sigma_X^2$ и $\sigma_Y = |a|\sigma_X$. Всяка линейна трансформация на суровия тестов бал не променя формата на честотното разпределение. В частност, наредбата на учениците според бала им се запазва, представянето на отделни подгрупи от тестираните ученици остава същото и т.н.

Суровият тестов бал от даден тест е удобно честотно разпределение за много цели, стига да се ограничим в рамките на един тест. Със суровия тестов бал, обаче, не може и не трябва да се сравняват резултати на ученици, правили различни тестове. Изобщо, ако трябва да се анализират резултати по два различни теста, не може да се използва суровия тестов бал. Това е така, защото, грубо казано, всеки тест е измерителен инструмент, който е оразмерен сам за себе си по отношение на суровия тестов бал. Ако искаме да правим сравнения между два различно оразмерени инструмента, трябва да преминем към някакво „универсално” оразмеряване („универсална” скала).

2. Стандартизирана скала и вторични скали

Естествена „универсална” скала е стандартизираното честотно разпределение на суровия бал. То се получава така, както се стандартизира всяко честотно разпределение.

Когато е даден суров тестов бал X, известни са числата μ – средната му стойност и σ – стандартното му отклонение. Стандартизиран тестов бал Z се изчислява по формулата $Z = \dfrac{X - \mu}{\sigma}$, т.е. от суровия тесто бал на всеки ученик се изважда получената средната стойност и се разделя на стандартното отклоне-

ние. Това е линейна трансформация на суровия тестов бал с коефициенти $a = \frac{1}{\sigma}$ и $b = -\frac{\mu}{\sigma}$. Стойностите на стандартизираните балове са не само цели числа. Теоретично, те може да са всички реални числа. Средната стойност на стандартизираните балове е 0, а стандартното им отклонение е 1.

Със стандартизирания бал могат да се сравняват резултати от различни тестове. Ето един пример. С група ученици са проведени два теста – един по български език и един по математика. Двата теста са с по 100 задачи с избираем отговор всеки. Иван решил 85 задачи от теста по български език и 62 задачи от теста по математика. Имат ли право родителите му да го упрекват, че Иван е по-лошо подготвен по математика?

От данните, с които разполагаме, може да се направи изводът, че суровият тестов бал на Иван по български език е 85, а по математика е 62. Понеже със суров тестов бал не може да се сравняване на резултати от различни тестове, няма как да се прави извод за това дали Иван е по-лошо подготвен по математика. С наличието на допълнителни данни може да се изчислят съответните стандартизирани балове. Тези данни са показани в Таблица 23.

Таблица 23. Данни от тестовете към примера

	Средна стойност	Станд. отклонение	Суров тестов бал на Иван
Български език	52	15	85
Математика	41	8,4	62

Изчисляваме стандартизирания бал на Иван за български език и математика и получаваме съответно $Z_{БЕ} = \frac{85-52}{15} = 2,2$ и $Z_{МАТ} = \frac{62-41}{8,4} = 2,5$. Както се вижда, стандартизираният бал по математика е малко по-висок от този по български език. Затова изводът, че Иван е по-лошо подготвен по математика, не е състоятелен.

От свойствата на нормалното разпределение следва, че около 68% от стандартизираните балове по даден тест са в интервала (–1; 1), около 95% от тях са в интервала (–2; 2) и около 99% са в интервала (–3; 3).

Стандартизираните балове имат някои недостатъци, които са предимно от психологическо естество. За разлика от суровия тестов бал, смисълът на стандартизирания бал не е така лесно разбираем. Част от стандартизираните балове са отрицателни числа, а поставянето на отрицателни балове на ученици не е прието.

Последният недостатък може да се избегне чрез използване на така наречените *вторични балове*. Всеки вторичен бал се получава с линейна трансформация на стандартизирания бал, т.е. $Y = aZ + b$, където $a \neq 0$ и b са някакви числа. Често се използва така наречения Т-бал, който се получава по формулата $T = 10Z + 50$. На практика, при него не е възможно да се получат отрицателни балове. В разгледания пример, Т-баловете на Иван са $T_{БЕ} = 10.2,2 + 50 = 72$ и $T_{МАТ} = 10.2,5 + 50 = 75$.

Използването на стандартизирани и вторични балове е смислено само когато разпределението на суровите тестови балове е близко до нормалното.

3. Процентна скала

Процентният бал на даден ученик е равен на процента от задачите в теста, които ученикът е решил правилно. Например, ако даден тест има 100 задачи и Мария правилно е решила 83 от тях, нейният процентен бал е 83. Ако даден тест има 40 задачи и Петя правилно е решила 31 от тях, нейният процентен бал е $\frac{31}{40} \cdot 100 = 77,5$. Стойностите на процентните балове са не само цели числа. Теоретично, те може да са всички реални числа между 0 и 100.

Смисълът на процентния бал е лесно разбираем. Докато суровият тестов бал дава абсолютния брой на решените от ученика задачи от теста, процентният бал дава относителния дял (частта) на решените от него задачи. Това дава възможност с процентния да се правят само *някои* сравнения между различни тестове.

Стандартизираната скала, вторичните скали и процентната скала се получава с линейна трансформация на суровия тестов бал. Затова всичките те дават едно и също разпределение баловете. Например: ако се направи хистограма на разпределението на баловете по всяка от тези скалите, формата на тези хистог-

рами ще е една и съща; наредбата на учениците по всеки от баловете ще е едно и също.

4. Процентилна скала

Процентилният бал на даден ученик е равен, грубо казано, на процента ученици, които имат суров тестов бал по-нисък от неговия. Като използваме това по-свободно тълкуване, да разгледаме един пример. Ако 100 ученици са правили даден тест и са подредени според получения суров тестов бал от най-малкия към най-големия, ученикът, който е на 53-то място има процентилен бал 52 (защото 52% от учениците имат по-нисък бал от неговия).

Позволихме си тази по-неточна интерпретация на процентилния бал, тъй като тя е лесно разбираема и съдържа изцяло смисъла на този бал. Изчисляването на процентилния бал за суров тестов бал i става с формулата $P_i = \frac{cf_{<i} + 0,5 f_i}{N} \cdot 100$, където $cf_{<i}$ е натрупаната честота на всички сурови тестови балове, които са по-малки от i, f_i е честотата на суровия тестов бал i и N е броят на всички ученици, с които е проведен теста.

Стойностите на процентилните балове са не само цели числа. Теоретично, те може да са всички реални числа между 0 и 100. Да отбележи, че не е възможно да се получи процентилен бал 100, защото дори ако ученикът с най-висок суров тестов бал е решил всички задачи от теста, той не може да има бал по-нисък от собствения си.

Процентилната скала определя позицията на всеки ученик от групата, правила теста, спрямо останалите ученици. Затова, процентилната скала е удобна, когато трябва да се прави класиране на учениците и особено, когато от това класиране се избират част от най-добре представилите се ученици според някаква квота.

Процентилната скала е *нелинейна* трансформация на суровия тестов бал. Тя запазва наредбата на учениците същата, както е и при скалата на суровия тестов бал. Това, което не се запазва при преминаване към процентилна скала е, че в различни интервали от скалата на суровия тестов бал нарастване с 1 точка води до различно по големина нарастване на процентилния бал. Това е така, защото обикновено стойностите на честотите на баловете в средата на скалата на

суровия тестов бал са по-големи, отколкото тези в двата й края (поради формата на нормалното разпределение – камбана). Следователно, нарастване с 1 точка в средата на скалата на суровия тестов бал дава по-голямо нарастване на процентилния бал, отколкото 1 точка нарастване в някой от двата края.

Тази особеност на процентилната скала трябва да се има пред вид, за да не се правят погрешни заключения, свързани с нея. Ето два типични примера.

- Не е смислено да се сравнява разлика в процентилни балове в различни интервали от скалата. Например, да предположим, че се оценява напредък в постиженията на група ученици за определен период от време с входен и изходен тест. Да предположим още, че всеки ученик е постигнал на изходния тест с 3 точки суров тестов бал повече, отколкото на входния тест, т.е. напредъкът на всички е един и същ. Ако се сравни напредъка на учениците в процентилен бал, ще се получи, че „най-силните" и „най-слабите" ученици показват по-малък напредък, отколкото тези със средни постижения.
- Сравняването на средните стойности на постиженията на отделни подгрупи ученици не е смислено да се прави с процентилен бал. За целта се използва суровия тестов бал или някои от неговите линейни трансформации.

5. Нормализирана стандартизирана скала

Понякога разпределението на суровия тестов бал не е близко до нормалното. Някои от анализите и изводите от баловете, обаче, до голяма степен зависят от предположението за „близко до нормалното" разпределение. Друго удобство да се работи с разпределение близко до нормалното е, че може да се оценява приблизителния процент на учениците в произволен интервал от разпределението.

Разпределение на баловете, което съществено се различава от нормалното, може да се нормализира, като се премине в така наречената нормализирана стандартизирана скала. Това става по следния начин. Най-напред се изчисляват процентилните балове за всеки тестов бал. След това се използва таблицата за нормално стандартизирано разпределение, от която се намират съответните стандартизирани стойности на нормалното разпределение, за които процентът на съответните лица е равен на изчислените процентилни балове.

Нормализираната стандартизирана скала запазва наредбата на учениците същата, както е и при скалата на суровия тестов бал. Тя, обаче, е *нелинейна* трансформация на суровия тестов бал. Разликата между разпределенията на стандартизираните балове и на нормализираните такива е толкова по-голяма, колкото е по-голямата разлика на суровия тестов бал от нормалното разпределение.

Използването на нормализирана стандартизирана скала трябва да става с особено внимание. Препоръчва се преминаването към нея да става от специалист, който е наясно с всички „подводни камъни”, които крие такава трансформация.

Таблица 24 представя данни от един тест с 30 задачи и различни скали за представяне на резултатите, получени от група ученици.

Таблица 24. Представяне на резултати от тест в различни скали

Суров тестов бал	Чес-тота	Натру-пана честота	Стандар-тизиран бал	Процентен бал	Процентилен бал	Нормализиран стандарти-зиран бал
11	2	2	–2,53	36,67	0,67	–2,48
12	1	3	–2,17	40,00	1,67	–2,13
13	6	9	–1,80	43,33	4,00	–1,75
14	5	14	–1,44	46,67	7,67	–1,43
15	12	26	–1,07	50,00	13,33	–1,11
16	17	43	–0,71	53,33	23,00	–0,74
17	21	64	–0,34	56,67	35,67	–0,37
18	28	92	0,02	60,00	52,00	0,05
19	19	111	0,39	63,33	67,67	0,46
20	15	126	0,75	66,67	79,00	0,81
21	10	136	1,12	70,00	87,33	1,14
22	5	141	1,48	73,33	92,33	1,43
23	3	144	1,85	76,67	95,00	1,64
24	4	148	2,21	80,00	97,33	1,93
25	2	150	2,58	83,33	99,33	2,48
$\mu = 17{,}94$ $\sigma = 2{,}74$	Общо 150		$Z = \frac{X - \mu}{\sigma}$	Брой задачи 30	$P_i = \frac{cf_{<i} + 0{,}5f_i}{N} \cdot 100$	

Ще обясним пресмятанията за последните два стълба от таблицата. За да изчислим процентилния бал за суров тестов бал 14, например, намираме натрупаната честота за тестов бал 13, която е 9; прибавяме половината от честотата на бал 14, което е 2,5; делим на броя на всички ученици, което е 150; умножаваме по 100, т.е. $P_{14} = \frac{cf_{<14} + 0{,}5f_{14}}{N} \cdot 100 = \frac{9 + 2{,}5}{150} = 7{,}67$.

За да изчислим нормализирания стандартизиран бал за суров бал 21, например, намираме съответния процентилен бал, който е 87,33; в таблицата за нормално стандартизирано разпределение намираме, че за съответната стойност на нормалното стандартизирано разпределение, чието лице е 0,8733, е 1,14. За да изчислим нормализирания стандартизиран бал за суров бал 15, намираме съответния процентилен бал, който е 13,33. Сега изчисляваме $100 - 13{,}33 = 86{,}67$. В таблицата за нормално стандартизирано разпределение намираме, че за съответната стойност на нормалното стандартизирано разпределение, чието лице е 0,8667, е 1,11; сменяме знака и получаваме бал –1,11.

6. Приравняване на балове от различни варианти на един и същ тест

Често се налага за един и същ тест да се направят два вариант, които да се проведат с различни ученици. Типичен пример е, когато искаме да разделим учениците на две групи, така че стоящите един до друг да нямат възможност да заимстват. Въпреки старанията двата варианта да се направят колкото е възможно по еднотипни (паралелни форми), винаги се появяват различия между тях. Това налага да се приравняват баловете, получени от единия вариант към баловете от другия вариант. Ако означим с *X* и *Y* статистическите редове на баловете по двата варианта, задача е да се намери функция *f*, която преобразува баловете *X* в балове $Y^* = f(X)$, които са „приравнени" към баловете *Y*. Така, баловете *X* могат да се сравняват с баловете *Y*.

Съществуват различни методи за такова приравняване. Тук няма да разглеждаме изчерпателно въпроса в неговата дълбочина. Ще обясним сравнително лесно разбираемия пример, когато тестваната група ученици по произволен начин е разделена на две подгрупи, като всяка подгрупа работи по един от двата варианта на теста. Предполага се още, че статистическите разпределения на баловете *X* и *Y* са еднотипни (т.е. хистограмите им имат една и съща форма).

Да означим с μ_X и μ_Y средните стойности на суровите тестови балове от двата варианта, а с σ_X и σ_Y съответно техните стандартни отклонения. Приравняването на баловете става в „универсалната" стандартизирана скала. За да има сравнимост между стандартизираните балове на двата варианта, трябва

да е изпълнено равенството $\frac{X-\mu_X}{\sigma_X}=\frac{Y-\mu_Y}{\sigma_Y}$. Като се реши това равенство относно *Y*, се намира нужната трансформация на баловете *X* в балове Y^*, а именно $Y^*=\frac{\sigma_Y}{\sigma_X}\left(X-\mu_X\right)+\mu_Y$. Така, баловете *X* от единия вариант са „еквивалентни" на баловете Y^*, които са приравнени към баловете *Y*.

Да разгледаме един пример. Нека $\mu_X=50$; $\sigma_X=10$; $\mu_Y=52$ и $\sigma_Y=11$. Тогава трансформацията на баловете *X* в балове Y^* е $Y^*=\frac{\sigma_Y}{\sigma_X}\left(X-\mu_X\right)+\mu_Y=1{,}1\left(X-50\right)+52$. Това означава, че суров тестов бал 45 по варианта *X* е еквивалентен на тестов бал $Y^*=1{,}1\left(45-50\right)+52=46{,}5$ по варианта *Y*.

7. Тестови норми

Резултатите от всяко измерване се анализират и се сравняват с определени норми (стандарти), въз основа на което могат да се вземат някакви решения. При нормативните оценявания резултатите на всеки ученик се сравняват с резултатите на останалите ученици от групата и в този смисъл нормите, които се използват, са относителни (за групата, която е правила теста), а не са универсални. Тук са разгледани два от най-популярните методи за определяне на норми: сигмалния метод и метода на процентилите.

Идеята на *сигмалния метод* за определяне на тестови норми е, че „средната" норма е около средната стойност μ на баловете на учениците, правили теста. За оценка на отклоненията от тази средна норма се използва стойността σ на стандартното отклонение на тези балове. Най-простият начин да сравним всеки получен тестов бал *X* със средната норма е да ползваме следните три категории (фиг. 74):

Фиг. 74. Сравняване на резултати със средната норма

- Ако $X<\mu-\sigma$, резултатът се оценява като нисък (под средното равнище).
- Ако $\mu-\sigma\le X\le\mu+\sigma$, резултатът се оценява като добър (на средното равнище).

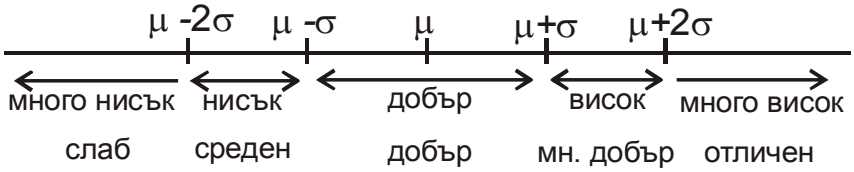
- Ако $X > \mu + \sigma$, резултатът се оценява като висок (над средното равнище).

Фактът, че така получените норми са относителни означава, че един и същ резултат върху един и същ тест може да се оцени по различен начин, в зависимост от общото постижения на групата, в която той е получен. В една по-силна група, даден резултат може да се оцени като нисък, а същият резултат в друга по-слаба група може да е добър, дори и висок.

Ако е необходима по-голяма диференциация, се използват следните пет категории (фиг. 75):

- Ако $X < \mu - 2\sigma$, резултатът се оценява като много нисък (много под средното равнище).
- Ако $\mu - 2\sigma \le X < \mu - \sigma$, резултатът се оценява като нисък (под средното равнище).
- Ако $\mu - \sigma \le X \le \mu + \sigma$, резултатът се оценява като добър (на средното равнище).
- Ако $\mu + \sigma < X \le \mu + 2\sigma$, резултатът се оценява като висок (над средното равнище).
- Ако $X > \mu + 2\sigma$, резултатът се оценява като много висок (много над средното равнище).

Тези пет категории може да се използват и за категориите в така наречената шестобална скала за оценяване, която се използва в българската образователна система (фиг. 75). Отново да подчертаем, че тези категории са относителни. В случая, това означава, че независимо от относителното ниво на групата ученици, с които е проведен теста, процентът на учениците в категориите „слаб”, „среден”, „добър”, „мн. добър” и „отличен” е винаги приблизително един и същ и той е съответно 2%, 14%, 68%, 14%, 2%.



Фиг. 75. Пет категории за сравняване на

Трябва да се отбележи още, че използването на сигмания метод дава смислени резултати само когато разпределението на суровия тестов бал е близко до нормалното. Процентите от предния абзац се отнасят за този случай.

*Методът на процентилит*е използва същата идея, но се различава по определяне на точките, които разделят категориите. При използване на три категории (нисък, добър и висок резултат) тези точки са стойностите на 27-тия и 73-тия процентил на статистическия ред от суровия тестов бал. Когато се използват пет категории, точките са стойностите на 5-тия, 25-тия, 75-тия и 95-тия процентил.

8. Трансформации в шестобална скала

Масовото използване на така наречената шестобална скала често налага да се трансформира суровия тестов бал (а понякога и други балове) в тази скала. С други думи, често се налага на учениците, с които е проведен теста, да се „поставят оценки" (понякога с точност до втория десетичен знак), а не само да се сравняват с определени норми. Има различни методи за такива трансформации, но всеки от тях има както определени предимства, така и някои недостатъци. Затова, препоръката е: *да се избягва трансформацията от тестов бал към шестобална (или друга оценъчна) скала винаги, когато това не е абсолютно необходимо*.

Тук са разгледани основните идеи на линейния и кумулативния метод за трансформиране към шестобална скала. Линейният метод е разгледан в контекста на: (1) структурата на теста; и (2) характеристиките на баловете.

Започваме с *линейния метод чрез използване на структурата на теста*. Три са основните стъпки при използването на този начин:

1) *Определя се долна граница от тестов бал за получаване на минималната положителна оценка „среден 3,00".* Тази граница трябва да е по-висока от теоретичната граница на налучкване. Например, ако имаме тест със задачи с избираем отговор с по 4 възможности за отговор, теоретично могат да се налучкат отговорите на 25% от задачите, т.е. теоретичната граница на налучкване е 25% от максималния тестов бал; следователно долната граница за получаване на минималната положителна оценка „среден 3,00" трябва да е по-голяма от 25% от максималния тестов бал. Ако искаме да сме „най-благоразположени" към тестваните ученици, може да поставим минималната положителна оценка „среден 3,00" за най-малката възможна стойност на бала, която е по-голяма от теоретичната граница на налучкване. Препоръчва се,

обаче, минималната положителна оценка „среден 3,00” да се поставя за малко по-високи стойности, например за теоретичната граница на налучкване плюс едно стандартно отклонение на статистическия ред от суровия тестов бал.

2) *Определя се горна граница от тестов бал за получаване на възможно най-високата оценка „отличен 6,00”.* Основното съображение за определяне на тази граница е психологическо. Допуска се, че дори ако тестван ученик знае и може да реши всяка задача от теста, той може да допусне грешки поради бързане, невнимание, моментно разсейване и други психологически причини. Затова границата за получаване на възможно най-високата оценка „отличен 6,00” е в границите на 95% – 98% от максималния тестов бал, т.е. допуска се 2% – 5% от задачите да бъдат „сгрешени” поради психологически причини, а не от незнание. Разбира се, при по-дълги тестове (по-продължително тестово време) процентът на решените задачи за възможно най-високата оценка „отличен 6,00” може да е по-малък, отколкото при по-къси тестове.

3) *Междинните тестови балове се преобразуват линейно, т.е. през равни интервали.* Това се прави по следния начин: Нека $a < b$ са тестови балове, като a е тестовият бал, за който се получава минималната положителна оценка A (в случая на така наречената шестобална скала $A = 3$, т.е. „среден 3,00”), а b е минималният тестов бал, за който се получава възможно най-високата оценка B (в случая на така наречената шестобална скала $B = 6$, т.е. „отличен 6,00”). Тестовите балове в интервала (a,b) се преобразуват като се построи линейна функция $y = kx + l$ с условия: при $x = a$, $y = A$ и при $x = b$, $y = B$.

Ще обясним метода с един пример. Да предположим, че имаме тест с 50 задачи с по 4 възможности за отговор. С налучкване се решават 25% от задачите, което е 12,5 задачи. Може да считаме, че 12 решени задачи са получени чрез налучкване и не показват наличие на знание по темата на теста, т.е. до 12 решени задачи оценката е „слаб 2,00”. Възможно най-ниският суров тестов бал за получаване на „среден 3,00” е 13. Както вече казахме, препоръчително е този бал да се завиши, например с едно стандартно отклонение на вариационния ред на резултатите от теста. Да предположим, че стандартното отклонение е 6,24. Тогава тестовия бал за получаване на „среден 3,00” трябва да е не по-малко от $12,5 + 6,24 = 18,74$. И така минималният суров тестов бал за получаване на „среден 3,00” е 19 точки. За минималния суров

тестов бал, за който се получава възможно най-високата оценка „отличен 6,00” приемаме 48 точки (приемаме, че 4% от задачите, т.е. 2 задачи, могат да се сбъркат по психологически причини). Търсим функция от вида $y = kx + l$ (y е оценката, а x е броя точки), така че при $x = 19$, $y = 3$ и при $x = 48$, $y = 6$. От системата уравнения $3 = 19k + l$ и $6 = 48k + l$ получаваме $a = \frac{3}{29}$ и $b = \frac{30}{29}$. Тогава функцията е $y = \frac{3}{29}x + \frac{30}{29}$. Таблица 25 дава превръщането на суровия тестов бал в шестобална оценка:

Таблица 25. Превръщане на суров тестов бал в шестобална оценка

Тестов бал	**Оценка**		**Тестов бал**	**Оценка**		**Тестов бал**	**Оценка**
до 18	2,00		28	3,93		38	4,97
19	3,00		29	4,03		39	5,07
20	3,10		30	4,14		40	5,17
21	3,21		31	4,24		41	5,28
22	3,31		32	4,34		42	5,38
23	3,41		33	4,45		43	5,48
24	3,52		34	4,55		44	5,59
25	3,62		35	4,66		45	5,69
26	3,72		36	4,76		46	5,79
27	3,83		37	4,86		47	5,90
						над 47	6,00

Едно от предимствата на този метод е, че таблицата за оценяване може да се обяви преди провеждане на теста. Разбира се, тук има уговорка. Ако искаме да използваме стандартното отклонение на вариационния ред като корекция за минималния бал, при който се получава на минималната положителна оценка „среден 3,00”, това число не е известно преди провеждането на теста. Ако имаме данни от подобна популация от предишни тествания, може да ползваме стандартното отклонение от последното такова тестване. Може да коригираме минималния бал, при който се получава на минималната положителна оценка „среден 3,00”, и по други съображения (например експертна оценка на

задачите и мнение на експертите колко най-малко задачи трябва да се решат, за да се получи минималната положителна оценка „среден 3,00”).

Два са основните недостатъци на този метод. Първо, той не е съобразен със статистическите характеристики на баловете, получени от учениците, с които е проведен теста. Второ, не е съгласуван с някои изследвания, които показват, че (в термините на суровия тестов бал) „цената на една точка не винаги е една и съща”, т.е. получаването на една точка става „по-лесно”, когато тестваният има „малък брой точки” и значително по-трудно, ако тестваният вече е събрал над половината от максималния брой точки. Това означава, че линейната функция, която се построява по този метод, не би трябвало да е линейна. Някои усъвършенствания на този метод вземат пред вид това.

Линейният метод чрез използване характеристиките на баловете предполага, че тестът вече е проведен и са известни средно аритметичната стойност μ и стандартното отклонение σ на получените балове. Категориите категориите „слаб”, „среден”, „добър”, „мн. добър” и „отличен” се определят така, както е обяснено в предния параграф (фиг. 75). Оценки с точност до десети или стотни могат да се получат като във всеки интервал се построява линейна функция, така, както е обяснено по-горе. Тук, обаче линейните функции са различни за различните интервали. Това в известна степен „компенсира” последния от недостатъците на предишния метод (това, че „цената на една точка не винаги е една и съща”).

Да разгледаме същия пример на тест от 50 задачи с избираем отговор с по 4 възможности за отговор. Нека средноаритметичната стойност на баловете е $\mu = 32,43$ и стандартното му отклонение е $\sigma = 6,24$. Тогава точките на деление са: $\mu - 2\sigma = 19,95$, $\mu - \sigma = 26,19$, $\mu + \sigma = 38,67$ и $\mu + 2\sigma = 44,91$. И така, до 19 точки включително е категорията „слаб”, от 20 до 26 точки включително е категорията „среден”, от 27 до 38 точки включително е категорията „добър”, от 39 до 44 точки включително е категорията „мн. добър”, от 45 точки нагоре е категорията „отличен”. Ако искаме да прецизираме оценките с точност до стотни, трябва да построим линейни функции за всеки от интервалите [20; 27], [28; 38], [39; 45] и [46; 50]. Краищата на интервалите съответстват на следните оценки: 20 съответства на 3,00; 27 съответства на 3,50; 39 съответства на 4,50; 45 съот-

ветства на 5,50 и 50 съответства на 6,00. Таблица 26 представя превръщането на суровия тестов бал в оценка и отдолу са написани съответните функции.

Таблица 26. Линейни функции за всеки от интервалите

Тестов бал	**Оценка**	**Тестов бал**	**Оценка**	**Тестов бал**	**Оценка**	**Тестов бал**	**Оценка**
до 19	2,00	28	3,58	39	4,50	46	5,60
20	3,00	29	3,67	40	4,67	47	5,70
21	3,07	30	3,75	41	4,83	48	5,80
22	3,14	31	3,83	42	5,00	49	5,90
23	3,21	32	3,92	43	5,17	50	6,00
24	3,29	33	4,00	44	5,33		
25	3,36	34	4,08	45	5,50		
26	3,43	35	4,17				
27	3,50	36	4,25				
		37	4,33				
		38	4,42				
Функция		Функция		Функция		Функция	
$y = \frac{1}{14}x + \frac{22}{14}$		$y = \frac{1}{12}x + \frac{5}{4}$		$y = \frac{1}{6}x - 2$		$y = \frac{1}{10}x + 1$	

Тук също може да се допусне, че изследваният ученик може да направи няколко грешки по психологически причини, както беше в случая на предишния начин. Тогава оценката 6,00 ще бъде давана за няколко точки под максималните (ако се придържаме към разглеждания пример, би могло да е за 48 точки и нагоре).

Предимство на този начин е, че той е по-добре съобразен със (статистическите) характеристики на получените балове. Като недостатък може да се счита, че таблицата за превръщане на баловете в оценки не може да се направи (публикува) преди провеждането на теста.

Този метод *може да се използва само когато разпределението на баловете е близко до нормалното.* В противен случай методът дава сериозни отклонения.

При *кумулативния метод* оценката P на ученика се получава с формулата $P = 2 + 4\frac{R}{M}$, където R е полученият бал на ученика, а M е максималният възможен тестов бал. За сравнение с другите методи, прилагаме таблица 27 за преминаване от суров тестов бал в оценка при този начин, като използваме разгледания вече тест от 50 задачи.

Таблица 27. Преминаване към шестобална оценка с използване на кумулативния метод

Тестов бал	**Оценка**		**Тестов бал**	**Оценка**		**Тестов бал**	**Оценка**
до 12	2,00		25	4,00		38	5,04
13	3,04		26	4,08		39	5,12
14	3,12		27	4,16		40	5,20
15	3,20		28	4,24		41	5,28
16	3,28		29	4,32		42	5,36
17	3,36		30	4,40		43	5,44
18	3,44		31	4,48		44	5,52
19	3,52		32	4,56		45	5,60
20	3,60		33	4,64		46	5,68
21	3,68		34	4,72		47	5,76
22	3,76		35	4,80		48	5,84
23	3,84		36	4,88		49	5,92
24	3,92		37	4,96		50	6,00

При този начин максималната оценка „отличен 6,00” се получава само за правилно решение на всички задачи от теста (т.е. не се разрешават грешки по психологически причини). Долната граница за получаване на минималната положителна оценка „среден 3,00” *винаги* е равна на 25% от максималния брой точки, независимо от трудността на теста и от получените тестови балове. За сравнение: ако възможностите за отговор са по-малко от 4, този метод поставя минималната положителна оценка „среден 3,00” под границата на налучкването; при 4 възможности за отговор тя е точно на границата на налучкването; при повече от 4 възможности за отговор тя е над границата на налучкването.

Предимство на този начин е, че той е приложим за всякакви тестове, в които резултатите са в брой точки (тестови балове).

Най-сериозният му недостатък е, че не се вземат пред вид никакви (статистически) характеристика (свойства) на теста (като структура) или на получените балове. Може да се каже, че методът е „механичен" в смисъл, че не се съобразява с предметната област, трудността на конкретния тест и особеностите на получените балове.

9. Описание на метода на опорните точки

Методът на опорните точки се използва за качествен анализ на знанията и уменията на учениците, чиито постижения са в определени точки от скалата. Той е приложим за всякакъв вид непрекъсната скала. Основните стъпки при изпълнение на метода са следните:

1. Избират се няколко точки $a_1 < a_2 < ... < a_n$ от скалата на постиженията. Тези точки се наричат *опорни*.
2. За всяко $i = 1,2,...,n$ се определя множеството A_i от ученици, чиито постижения по теста са равни на a_i.
3. За всяко $i = 1,2,...,n$ се определят задачите от теста, които „характеризират" опорните точки. Това са задачи, които са решени от „болшинството" ученици A_i. Условието е поне p% от учениците A_i да са решили съответната задача, където p е определено число. За $i > 1$ се поставя и условието задачата да не е решена от „много" ученици от групата A_{i-1}, т.е. по-малко от q% ученици от A_{i-1} да са решили задачата, където q е също определено число. За определените така задачи се казва, че са *асоциирани към опорната точка* a_i.
4. Нека B_i е множеството от задачи, асоциирани към опорната точка a_i. Екип от специалисти по предметната област на теста анализират множествата B_i, $i = 1,2,...,n$, с цел да опишат знанията и уменията на всяка от групите ученици A_i.

10. Избиране на опорни точки

Теоретично всяка точка от скалата може да бъде опорна. Това обаче не означава, че техният избор става произволно. В началото се определя метод за избиране. Той може да бъде:

- *Емпиричен* – точките се определят в процентили или в интервали от стандартното отклонение;
- *Експертен* – експертна група определя точките чрез така наречените „нива на постижения", които описва какво учениците би трябвало да знаят и могат в съответното ниво.

Например, в началото NAEP са използвали емпиричния метод изразен чрез смислени интервали от стандартната грешка. По-късно те са преминали към експертния метод чрез определяне на три нива на постижения: основно, средно, напреднало. Подобна е и историята за избор на базисни точки в TIMSS. В TIMSS-1999 се използват емпирично определените точки 25-ти, 50-ти, 75-ти и 90-ти процентил. В TIMSS-2003 година се преминава към фиксирани точки от скалата за постижения, а именно 400, 475, 550 и 625. (Скалата за ученическите постижения на изследването TIMSS е конструирана чрез Теорията на вероятностното моделиране (Item Response Theory – IRT) със средна стойност 500 и стандартно отклонение 100.) Опорните точки на TIMSS определят четири нива на постижения, които може да наречем съответно „*много ниски постижения*" (МНП), „*ниски постижения*" (НП), „*високи постижения*" (ВП) и „*много високи постижения*" (МВП). Нивото МНП характеризира постиженията на учениците от 25-тия процентил (за TIMSS-1999) или тези, които имат бал не по-висок от 400 (за TIMSS-2003). Нивото НП характеризира постиженията на учениците от 50-тия процентил (за TIMSS-1999) или тези, които имат бал между 400 и 475 (за TIMSS-2003). Нивото ВП характеризира постиженията на учениците от 75-тия процентил (за TIMSS-1999) или тези, които имат бал между 475 и 550 (за TIMSS-2003). Нивото МВП характеризира постиженията на учениците от 90-тия процентил (за TIMSS-1999) или тези, които имат бал по-висок от 625 (за TIMSS-2003).

Независимо от метода, по който става определянето на опорните точки, те трябва да са достатъчно много, за да дават възможност за смислено описание на скалата на постиженията, т.е. да обединят групите ученици, които имат раз-

лични постижения. От друга страна, опорните точки трябва да са на достатъчно голямо разстояние една от друга, за да разграничават учениците, чиито постижения са в две съседни точки. Практиката показва, че разумният брой на опорните точки е между 3 и 5.

11. Определяне на множеството ученици за всяка опорна точка

Обикновено броят на учениците, които влизат в някои групи A_i, така, както са дефинирани в точка 2 по-горе, е малък. За да получим достатъчно много ученици във всяка група, се постъпва по следния начин. Около всяка опорна точка a_i се определя околност $[a_i - \varepsilon, a_i + \varepsilon]$ и под A_i се разбира множеството от ученици, чиито постижения влизат в тази околност. Изборът на тези околности трябва да отговаря на следните две условия. Първо, те да са „големи", за да може всяка от тях да осигури достатъчно много ученици. Второ, те трябва да са разположени „далеч" една от друга, за да разграничават добре учениците от две различни околности по отношение на постиженията им.

Например, по времето, когато NAEP са ползвали емпиричния метод за определяне на опорни точки, околностите са избирани плюс и минус една четвърт от стандартното отклонение около съответната опорна точка.

В началото, когато TIMSS използваше емпиричния метод, се обмисляше вариант да се изберат околностите плюс и минус 25 (ε = 25) около всяка опорна точка. Оказа се, че такива интервали акумулират достатъчно много ученици, но не може да се направи добра разлика в постиженията на учениците от 75-тия и 90-тия проценти. Оптимално решение се получава като околностите се изберат плюс и минус 5 (ε = 5) около всяка опорна точка. Таблица 28 показва разпределението на интервалите за опорните точки на данните по математика за TIMSS-1999.

Таблица 28. Разпределение на учениците по нива за TIMSS-1999 (математика).

	МНП	НП	ВП	МВП
Опорна точка	25-ти проц.	50-ти проц.	75-ти проц.	90-ти проц.
Стойност	396	479	555	616
Интервал (ε = 5)	391 – 401	474 – 484	550 – 560	611 – 621
Брой ученици	3540	5690	5531	3703

След като преминава към фиксирани точки от скалата на постиженията, в TIMSS-2003 също се избират интервали с дължина плюс и минус 5 около всяка

опорна точка. Таблица 29 описва разпределението на интервалите за опорните точки на данните за математическата скала на TIMSS-2003. (Mullis, I. et al. 2004):

Таблица 29. Разпределение на учениците по нива за TIMSS-2003 (математика).

	МНП	НП	ВП	МВП
Опорна точка	400	475	550	625
Интервал (ε = 5)	395 – 405	470 – 480	545 – 555	620 – 630
Брой ученици	6372	8294	6955	3320

12. Определяне на задачите, асоциирани към опорните точки

Първият въпрос, който изниква тук, е свързан с определянето на процентите p и q. Читателят лесно ще съобрази как се влияе изборът на задачите асоциирани към дадена опорна точка с увеличаване или намаляване на всеки от тях. Да разгледаме примера с TIMSS като представителен. При него е избрано p = 65% и q = 50%. Това означава, че една задача се асоциира към МНП, ако поне 65% от учениците в групата A_1 са я решили. (За първата опорна точка процентът q не участва.) Една задача се асоциира към НП, ако поне 65% от учениците в групата A_2 са я решили и по-малко от 50% от учениците в групата A_1 са я решили. По аналогичен начин се определят и задачите, асоциирани към ВП и МВП.

Пример: За някаква задача в TIMSS са получени следните данни:

Ниво на постижение	МНП	НП	ВП	МВП
Процент вярно решили от групата A_i	14	43	72	86

Тогава тази задача се асоциира към нивото ВП.

Ако асоциираните задачи към някоя от опорна точка са малко, броят им може да се увеличи като се причислят и така наречените „почти асоциирани задачи”. Това са такива задачи, при които процентът p е „малко” занижен. Например, в TIMSS почти асоциирани са онези задачи, при които $p = 60\%$. Тогава, задачата с данни:

Ниво на постижение	МНП	НП	ВП	МВП
Процент вярно решили от групата A_i	47	63	79	93

е почти асоциирана към нивото НП.

Ако се окаже, че множеството от асоциираните и почти асоциираните задачи все още е малко, може да се пренебрегне и условието за процента q. Така получаваме множество от задачи, които са асоциирани само въз основа на p критерия. Например, задачи с данни:

Ниво на постижение	МНП	НП	ВП	МВП
Процент вярно решили от групата A_i	26	53	64	78

е асоциирана към нивото ВП само въз основа на p критерия.

Таблица 30 показва разпределението на броя математически задачи, асоциирани към опорните точки за TIMSS-1999.

Таблица 30. Брой задачи по математика в TIMSS-1999 по нива.

	Асоциирани	Почти асоциирани	Само по p критерия	***Общо***
МНП	4	2	0	6
НП	16	7	13	36
ВП	34	14	25	73
МВП	17	4	22	43
Общо	71	27	60	158

Таблица 31 дава подобна информация за TIMSS-2003 (Mullis, I. et al. 2004).

Таблица 31. Брой задачи по математика в TIMSS-2003 по нива.

	Асоциирани	Почти асоциирани	Само по p критерия	***Общо***
МНП	3	1	0	4
НП	25	5	10	40
ВП	46	10	19	75
МВП	41	5	17	63
Общо	115	21	46	182

От данните ясно се вижда, че включването на почти асоциирани задачи, както и на такива, които са асоциирани само въз основа на p критерия, значително увеличава броя на задачите, характеризиращи постиженията на учениците в отделните нива.

Фигурите от 66 до 81 представят някои типични математически задачи от TIMSS-2003 за всяко от нивата на постижения. За всяка задача е даден процентът на вярно решилите задачата за всяка от участващите държави и международното средно.

Фиг. 76

TIMSS 2003 Международна норма за **много ниски постижения** (400) по математика

Пример на задача, която ученици с много ниски постижения могат да решават*

8 Математика TIMSS 2003

Съдържателна област: Числа

Описание: Избиране на десетична дроб с две цифри след десетичната запетая, която е най-близо до дадено цяло число.

Кое от следните числа е най-близко до 10?

A) 0,10
B)* 9,99
C) 10,10
D) 10,90

	Държава	Процент Пълно решеие	
†	Холандия	97 (1.0)	h
	Швеция	96 (1.1)	h
	Естония	96 (1.2)	h
	Сингапур	95 (1.1)	h
1	Литва	95 (1.0)	h
	Белгия (фл.)	94 (1.4)	h
¿	Корея	94 (1.2)	h
	Малайзия	93 (1.4)	h
	Япония	92 (1.4)	h
1	Сърбия	91 (1.6)	h
	Норвегия	91 (1.3)	h
	Русия	91 (1.2)	h
	Латвия	90 (1.9)	h
	Словакия	90 (2.0)	h
	Италия	90 (1.9)	h
†	Хонг Конг	89 (1.6)	h
†	Шотландия	89 (2.0)	h
	Тайван	89 (1.5)	h
	Кипър	88 (2.0)	h
	Унгария	88 (2.0)	h
	Австралия	88 (1.8)	h
‡	САЩ	87 (1.1)	h
	Словения	87 (2.2)	h
	Нова Зеландия	86 (2.0)	h
	България	**85 2.7**	**h**
	Молдова	82 (2.5)	
2	Израел	81 (2.3)	
	Румъния	79 (2.5)	
2	Македония	78 (2.7)	
	Международно средно	77 (0.3)	
	Тунис	76 (2.3)	
1 ‡	Мароко	75 (3.1)	
1	Индонезия	74 (2.7)	
	Иран	69 (2.4)	i
	Чили	67 (1.9)	i
	Ливан	67 (2.7)	i
	Армения	66 (2.6)	i
	Йордания	55 (2.7)	i
	Палестина	50 (2.7)	i
	Бахрейн	49 (3.2)	i
	Египет	48 (2.5)	i
	Филипини	42 (2.8)	i
	Боцвана	40 (2.6)	i
	Саудитска Арабия	35 (2.6)	i
	Южна Африка	30 (2.7)	i
	Гана	24 (2.4)	i
¶	Англия	82 (2.5)	h
Допълнително участващи области			
	Район Баски, Испания	92 (2.0)	h
	Щат Индиана, САЩ	84 (3.2)	h
	Пр. Онтарио, Канада	91 (1.8)	h
	Пр. Квебек, Канада	91 (1.8)	h

h Средният бал на страната е статистически значимо по-голям от международното средно.

i Средният бал на страната е статистически значимо по-малк от международното средно.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

* Задачата е решене от голяма част от учениците в тази норма.
† Представителност на извадката е достигната само след добавяне на заместените училища.
‡ Почти е достигната представителност на извадката след добавяне на заместените училища.
¶ Не са покрити стандартите за представителност на извадката поради нисък процент участие на учениците.
1 Дефиницията на популацията на национално ниво не покрива изискванията на международната дефиниция.
2 Извадката е представителна за популация, която е по-малко от 90% от цялата национална популация.
¿ Корея е тествала същите ученици, но през есента на 2003 г., когато те са или в началото на следващата учебна година.
( ) Стандартното отклонение е показано в скоби. Резултатите са закръглени до най-близкото цяло число.

Фиг. 77

TIMSS 2003 Международна норма за **ниски постижения** (475) по математика
Пример на задача, която ученици с ниски постижения могат да решават*

8-ми клас TIMSS 2003 Математика

Съдържателна облат: Числа

Описание: Решаване на задача с думи чрез изваждане на десетични дроби с две цифри след десетичната запетая.

Ана пробягала дадено разстояние за 49,86 секунди.
Бистра пробягала същото разстояние за 52,30 секунди.
Колко повече време е бягала Бистра, отколкото Ани?

A)* 2,44 секунди
B) 2,54 секунди
C) 3,56 секунди
D) 3,76 секунди

	Държава	Процент Пълно решеие	
	Сингапур	88 (1,0)	h
¿	Корея	87 (1,1)	h
	Малайзия	81 (1,4)	h
†	Холандия	81 (2,0)	h
	Унгария	80 (1,9)	h
	Тайван	80 (1,6)	h
	Япония	78 (1,6)	h
	Русия	76 (1,8)	h
†	Хонг Конг	75 (1,6)	h
	Словакия	74 (2,1)	h
‡	САЩ	74 (1,7)	h
	Словения	73 (2,3)	h
	Естония	72 (1,8)	h
	Белгия (фл.)	71 (1,8)	h
†	Шотландия	71 (2,0)	h
	Молдова	69 (2,3)	h
1	Сърбия	68 (2,1)	h
	Латвия	67 (2,4)	h
	България	**66 (2,5)**	
1	Литва	65 (2,3)	
	Румъния	64 (2,4)	
	Тунис	63 (2,0)	
	Австралия	63 (2,4)	
	Швеция	63 (2,0)	
	Италия	62 (2,1)	
	Боцвана	61 (1,7)	
	Международно средно	61 (0,3)	
	Ливан	61 (2,3)	
	Армения	60 (2,2)	
2	Македония	59 (2,1)	
	Кипър	59 (1,8)	
	Египет	58 (1,7)	
2	Израел	58 (1,9)	
1	Индонезия	55 (2,0)	i
	Нова Зеландия	53 (2,4)	i
	Йордания	46 (2,2)	i
	Норвегия	46 (2,5)	i
	Филипини	45 (2,2)	i
1 ‡	Мароко	45 (2,6)	i
	Бахрейн	45 (2,0)	i
	Иран	44 (1,9)	i
	Чили	42 (1,8)	i
	Палестина	37 (1,7)	i
	Гана	32 (2,0)	i
	Южна Африка	29 (1,8)	i
	Саудитска Арабия	19 (2,3)	i
¶	Англия	54 (2,5)	i
Допълнително участващи области			
	Район Баски, Испания	64 (3,0)	
	Щат Индиана, САЩ	77 (2,2)	h
	Пр. Онтарио, Канада	73 (2,4)	h
	Пр. Квебек, Канада	76 (1,9)	h

Средният бал на страната е статистически значимо по-голям от международното средно. h

Средният бал на страната е статистически значимо по-малк от международното средно. i

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

* Задачата е решене от голяма част от учениците в тази норма.
† Представителност на извадката е достигната само след добавяне на заместените училища.
‡ Почти е достигната представителност на извадката след добавяне на заместените училища.
¶ Не са покрити стандартите за представителност на извадката поради нисък процент участие на учениците.
1 Дефиницията на популацията на национално ниво не покрива изискванията на международната дефиниция.
2 Извадката е представителна за популация, която е по-малко от 90% от цялата национална популация.
¿ Корея е тествала същите ученици, но през есента на 2003 г., когато те са или в началото на следващата учебна година.
( ) Стандартното отклонение е показано в скоби. Резултатите са закръглени до най-близкото цяло число.

Фиг. 78

TIMSS 2003 Международна норма за **ниски постижения** (475) по математика
Пример на задача, която ученици с ниски постижения могат да решават*

8-ми клас TIMSS 2003 Математика

Съдържателна област: Алгебра

Описание: Намиране на липсващ член в пропорция.

Ако $\frac{12}{n} = \frac{36}{21}$, то n е равно на:

A) 3
B)* 7
C) 36
D) 63

	Държава	Процент Пълно решеие	
	Сингапур	93 (0,7)	h
¿	Корея	89 (0,9)	h
†	Хонг Конг	88 (1,2)	h
	Белгия (фл.)	86 (1,4)	h
†	Холандия	85 (1,8)	h
	Малайзия	83 (1,5)	h
	Тайван	83 (1,5)	h
‡	САЩ	80 (1,1)	h
	Япония	79 (1,6)	h
	Унгария	79 (1,7)	h
†	Шотландия	79 (1,9)	h
	Австралия	76 (1,9)	h
	Словакия	74 (2,0)	h
	Словения	72 (2,3)	h
2	Израел	72 (2,0)	h
	Ливан	71 (2,6)	h
	Русия	71 (1,9)	h
	Естония	71 (2,2)	h
	Латвия	70 (2,1)	h
	Нова Зеландия	68 (2,3)	
	Швеция	66 (2,1)	
	Иран	66 (1,7)	
	Италия	65 (2,1)	
	Кипър	65 (1,8)	
	Международно средно	65 (0,3)	
	Тунис	64 (1,7)	
1	Литва	64 (2,1)	
1	Сърбия	63 (2,1)	
	Молдова	61 (2,5)	
	Румъния	61 (2,2)	i
	България	**59 (2,0)**	**i**
	Норвегия	59 (2,1)	i
1	Индонезия	58 (1,9)	i
	Египет	58 (2,2)	i
	Армения	54 (2,6)	i
1 ‡	Мароко	54 (3,0)	i
	Йордания	53 (1,9)	i
	Палестина	52 (1,6)	i
	Филипини	52 (2,1)	i
2	Македония	50 (2,3)	i
	Бахрейн	44 (2,2)	i
	Чили	44 (2,0)	i
	Боцвана	41 (1,7)	i
	Саудитска Арабия	30 (2,2)	i
	Гана	28 (1,6)	i
	Южна Африка	26 (1,5)	i
¶	Англия	74 (2,6)	h
Допълнително участващи области			
	Район Баски, Испания	77 (2,3)	h
	Щат Индиана, САЩ	83 (1,7)	h
	Пр. Онтарио, Канада	86 (1,8)	h
	Пр. Квебек, Канада	88 (1,4)	h

h — Средният бал на страната е статистически значимо по-голям от международното средно.

i — Средният бал на страната е статистически значимо по-малк от международното средно.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

* Задачата е решене от голяма част от учениците в тази норма.
† Представителност на извадката е достигната само след добавяне на заместените училища.
‡ Почти е достигната представителност на извадката след добавяне на заместените училища.
¶ Не са покрити стандартите за представителност на извадката поради нисък процент участие на учениците.
1 Дефиницията на популацията на национално ниво не покрива изискванията на международната дефиниция.
2 Извадката е представителна за популация, която е по-малко от 90% от цялата национална популация.
¿ Корея е тествала същите ученици, но през есента на 2003 г., когато те са или в началото на следващата учебна година.
( ) Стандартното отклонение е показано в скоби. Резултатите са закръглени до най-близкото цяло число.

Фиг. 79

TIMSS 2003 Международна норма за **високи постижения** (550) по математика
Пример на задача, която ученици с високи постижения могат да решават*

8-ми клас TIMSS 2003 Математика

Съдържателна област: Числа

Описание: Решаване на едностъпкова задача с думи, като се използва деление на цяло число с обикновена дроб.

Една лопатка загребва $\frac{1}{5}$ kg брашно. Колко лопатки са необходими за да се напълни торба от 6 kg брашно?

Отговор: ___ 6 : $\frac{1}{5}$ _____

6 . 5

30 лопатки

Показаните отговори илюстрират един начин за пълно решение на задачата.

Държава	Процент Пълно решеие	
Сингапур	79 (1,9)	h
† Хонг Конг	76 (1,8)	h
Тайван	75 (1,9)	h
† Холандия	74 (2,1)	h
¿ Корея	68 (1,5)	h
Япония	62 (1,8)	h
Белгия (фл.)	62 (2,2)	h
Швеция	60 (1,9)	h
Австралия	53 (2,6)	h
‡ САЩ	52 (1,7)	h
† Шотландия	51 (2,7)	h
Естония	51 (2,0)	h
Латвия	51 (2,7)	h
Унгария	51 (2,1)	h
Русия	49 (2,7)	h
2 Израел	48 (2,3)	h
Малайзия	47 (2,2)	h
Нова Зеландия	46 (3,2)	h
Словения	46 (2,1)	h
Армения	45 (2,2)	h
1 Литва	43 (2,3)	h
Словакия	43 (2,0)	h
Норвегия	39 (2,1)	
Румъния	39 (2,8)	
Международно средно	38 (0,3)	
1 Сърбия	38 (2,0)	
България	**38 (3,0)**	
Кипър	37 (1,8)	
Молдова	37 (2,7)	
Италия	34 (2,1)	i
1 Индонезия	26 (1,5)	i
2 Македония	22 (2,0)	i
Иран	20 (1,9)	i
Тунис	18 (1,4)	i
Египет	17 (1,4)	i
Йордания	16 (1,5)	i
Либия	15 (1,4)	i
Чили	13 (1,1)	i
Филипини	13 (1,3)	i
Бахрейн	11 (1,3)	i
Боцвана	11 (1,1)	i
Палестина	10 (1,2)	i
1 ‡ Мароко	8 (1,5)	i
Южна Африка	7 (1,3)	i
Саудитска Арабия	7 (1,9)	i
Гана	6 (1,0)	i
¶ Англия	50 (3,1)	h
Допълнително участващи области		
Район Баски, Испания	42 (2,5)	
Щат Индиана, САЩ	56 (4,0)	h
Пр. Онтарио, Канада	53 (2,2)	h
Пр. Квебек, Канада	61 (2,9)	h

Средният бал на страната е статистически значимо по-голям от международното средно. h

Средният бал на страната е статистически значимо по-малк от международното средно. i

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

* Задачата е решене от голяма част от учениците в тази норма.
† Представителност на извадката е достигната само след добавяне на заместените училища.
‡ Почти е достигната представителност на извадката след добавяне на заместените училища.
¶ Не са покрити стандартите за представителност на извадката поради нисък процент участие на учениците.
1 Дефиницията на популацията на национално ниво не покрива изискванията на международната дефиниция.
2 Извадката е представителна за популация, която е по-малко от 90% от цялата национална популация.
¿ Корея е тествала същите ученици, но през есента на 2003 г., когато те са или в началото на следващата учебна година.
( ) Стандартното отклонение е показано в скоби. Резултатите са закръглени до най-близкото цяло число.

Фиг. 80

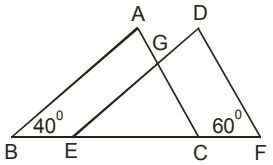
TIMSS 2003 Международна норма за **високи постижения** (550) по математика
Пример на задача, която ученици с високи постижения могат да решават*

8-ми клас TIMSS 2003 Математика

Съдържателна област: Геометрия

Описание: Използване на свойствата на еднакви триъгълници за намиране мярката на ъгъл.

На фигурата триъгълниците *ABC* и *DEF* са еднакви, като $BC = EF$.



Колко е мярката на ъгъл *EGC*?

A) 20^0
B) 40^0
C) 60^0
D)* 80^0
E) 100^0

Държава	Процент Пълно решеие	
¿ Корея	84 (1,4)	h
† Хонг Конг	81 (1,6)	h
Япония	80 (1,4)	h
Сингапур	79 (1,6)	h
Тайван	73 (1,9)	h
Естония	67 (2,0)	h
Белгия (фл.)	66 (1,7)	h
Латвия	63 (2,2)	h
България	**60 (2,6)**	**h**
2 Израел	57 (2,7)	h
Русия	55 (2,7)	h
Либия	55 (2,2)	h
† Шотландя	54 (2,7)	h
Словакия	54 (2,5)	h
1 Литва	51 (2,3)	h
Унгария	50 (2,4)	
Австралия	47 (2,1)	
Египет	47 (1,7)	
Малайзия	47 (2,4)	
Международно средно	46 (0,3)	
Армения	45 (2,4)	
Молдова	45 (3,0)	
Кипър	44 (2,2)	
† Холандия	44 (2,5)	
1 Сърбия	43 (1,9)	
Нова Зеландия	42 (3,6)	
Йордания	42 (1,8)	i
Итали	42 (2,3)	
Тунис	41 (1,6)	i
Бахрейн	41 (2,4)	i
Швеция	40 (2,1)	i
Палестина	39 (1,7)	i
Иран	37 (2,1)	i
Словения	37 (2,5)	i
‡ САЩ	36 (1,7)	i
2 Македония	33 (2,4)	i
Норвегия	32 (2,1)	i
1 Индонезия	31 (1,7)	i
1 ‡ Мароко	31 (2,2)	i
Чили	30 (1,8)	i
Саудитска Арабия	26 (2,5)	i
Южна Африка	21 (1,5)	i
Гана	20 (1,6)	i
Боцвана	20 (1,5)	i
Румъния	18 (1,7)	i
Филипини	15 (1,3)	i
¶ Англия	47 (2,8)	
Допълнително участващи области		
Район Баски, Испания	32 (2,5)	i
Щат Индиана, САЩ	30 (2,6)	i
Пр. Онтарио, Канада	50 (2,6)	
Пр. Квебек, Канада	69 (1,8)	h

Средният бал на страната е статистически значимо по-голям от международното средно. h

Средният бал на страната е статистически значимо по-малк от международното средно. i

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

* Задачата е решене от голяма част от учениците в тази норма.
† Представителност на извадката е достигната само след добавяне на заместените училища.
‡ Почти е достигната представителност на извадката след добавяне на заместените училища.
¶ Не са покрити стандартите за представителност на извадката поради нисък процент участие на учениците.
1 Дефиницията на популацията на национално ниво не покрива изискванията на международната дефиниция.
2 Извадката е представителна за популация, която е по-малко от 90% от цялата национална популация.
¿ Корея е тествала същите ученици, но през есента на 2003 г., когато те са или в началото на следващата учебна година.
( ) Стандартното отклонение е показано в скоби. Резултатите са закръглени до най-близкото цяло число.

Фиг. 81

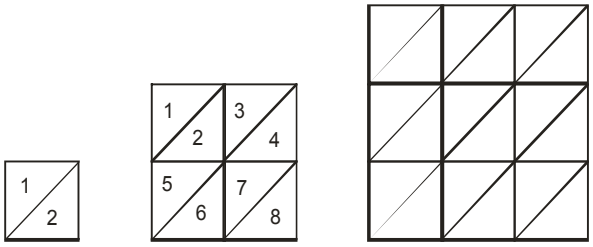
TIMSS 2003 Международна норма за **много високи постижения** (625) по математика
Пример на задача **(част С)**, която ученици с много високи постижения могат да решават*

8-ми клас TIMSS 2003 Математика

Съдържателна област: Алгебра

Описание: Част С–Обяснение за намиране на 50-тия член на редица при дадени първите няколко члена и двумерна геометрична интерпретация.

Трите фигури по-долу са разделени на еднакви малки триъгълници.



Фигура 1 Фигура 2 Фигура 3

А. Попълнете долната таблица. Най-напред напишете колко малки триъгълници има на Фигура 3. След това напишете колко малки триъгълници биха се получили на Фигура 4, ако редицата от фигури е продължена.

Фигура	Брой малки триъгълници
1	2
2	8
3	18
4	32

В. Редицата от фигури е продължена до Фигура 7. Колко малки триъгълници има на Фигура 7?

Отговор: ___98_____

$7^2.2=49.2=98$

С. Редицата от фигури е продължена до Фигура 50. Обяснете един начин да се намери броят на малките триъгълници на Фигура 50 без да е нужно да се начертае фигурата и да се преброят малките триъгълници.

$50^2.2=2500.2=5000$

Показаните отговори илюстрират един начин за пълно решение на задачата.

Държава	Процент Пълно решеие	
Тайван	49 (2,0)	h
¿ Корея	48 (1,8)	h
† Хонг Конг	45 (2,0)	h
Сингапур	44 (2,0)	h
Япония	44 (2,1)	h
† Холандия	36 (2,4)	h
Австралия	26 (2,7)	h
Унгария	24 (2,1)	h
† Шотландия	22 (2,2)	h
Белгия (фл.)	21 (1,3)	h
‡ САЩ	19 (1,5)	h
Швеция	17 (1,6)	
Нова Зеландия	16 (2,1)	
Естония	15 (1,3)	
Словакия	14 (1,5)	
Международно средно	14 (0,2)	
Италия	14 (1,5)	
Латвия	13 (1,5)	
Словения	13 (1,6)	
1 Сърбия	11 (1,2)	i
1 Литва	11 (1,3)	i
Румъния	11 (1,6)	i
Малайзия	10 (1,0)	i
2 Израел	10 (1,3)	i
Кипър	10 (1,1)	i
Норвегия	9 (1,3)	i
Русия	9 (1,2)	i
Армения	8 (1,2)	i
1 Индонезия	7 (0,9)	i
Чили	6 (0,8)	i
Йордания	5 (0,9)	i
Египет	5 (0,8)	i
Палестина	5 (0,7)	i
2 Македония	4 (0,9)	i
Филипини	4 (0,9)	i
България	**4 (0,8)**	
Бахрейн	4 (0,8)	i
Иран	3 (0,6)	i
1 ‡ Мароко	2 (0,8)	i
Боцвана	2 (0,5)	i
Южна Африка	1 (0,5)	i
Тунис	1 (0,3)	i
Ливан	1 (0,3)	i
Гана	1 (0,3)	i
Саудитска Арабия	0 (0,1)	i
Молдова	0 (0,1)	i
¶ Англия	20 (2,0)	h
Допълнително участващи области		
Район Баски, Испания	16 (2,0)	
Щат Индиана, САЩ	16 (1,9)	
Пр. Онтарио, Канада	26 (2,3)	h
Пр. Квебек, Канада	28 (2,7)	h

Средният бал на страната е статистически значимо по-голям от международното средно. h

Средният бал на страната е статистически значимо по-малк от международното средно. i

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

* Задачата е решене от голяма част от учениците в тази норма.
† Представителност на извадката е достигната само след добавяне на заместените училища.
‡ Почти е достигната представителност на извадката след добавяне на заместените училища.
¶ Не са покрити стандартите за представителност на извадката поради нисък процент участие на учениците.
1 Дефиницията на популацията на национално ниво не покрива изискванията на международната дефиниция.
2 Извадката е представителна за популация, която е по-малко от 90% от цялата национална популация.
¿ Корея е тествала същите ученици, но през есента на 2003 г., когато те са или в началото на следващата учебна година.
( ) Стандартното отклонение е показано в скоби. Резултатите са закръглени до най-близкото цяло число.

13. Описание на знанията и уменията на учениците за всяка опорна точка

Екип от специалисти преглеждат задачите, асоциирани към всяка опорна точка. Целта е да опишат знанията и уменията, които учениците трябва да имат, за да решат задачите от съответното множество B_i. Това описание в определен смисъл характеризира групата ученици от съответното ниво на постижения.

В таблица 32 са дадени кратките описания на учениците в опорните точки на международно ниво по математика за учениците от 8 клас в TIMSS-1999 (Mullis, I. et al. 2000, 55-90.) и в TIMSS-2003 (Mullis, I., at al. 2004, 55-100.)

Таблица 32. Описания на нивата на постижение по математика за TIMSS.

Ниво на постижение	**TIMSS-1999**	**TIMSS-2003**
МНП	Учениците могат да извършват основни действия с естествени числа.	Учениците имат някои основни математически познания. Те могат да извършват аритметични действия с цели числа без използване на калкулатор. Могат да определят десетична дроб с две цифри след десетичната запетая, която е най-близо до дадено цяло число. Могат да умножават десетични дроби с калкулатор. Познават основната терминология и могат да четат информация от прости графики.
НП	Учениците могат да прилагат основни математически знания в познати ситуации. Те събират и изваждат естествени числа и десетични дроби, за да решат едностъпкови задачи с думи; определят нагледно представени обикновени дроби по относителната големина на представянето; намират неизвестен член на пропорция; познават основните понятия свързани с пропорционалност и процент; използват основни свойства на геометричните фигури; разчитат и интерпретират данни, представени графично, с таблици или скали; разбират смисъла на прости алгебрични изрази.	Учениците могат да прилагат основни математически знания в прости ситуации. Те могат да събират, изваждат и умножават цели числа и десетични дроби, за да решават едностъпкови задачи с думи. Познават различни представяния на обикновените дроби и тяхната големина. Разбират прости алгебрични връзки и могат да решават линейни уравнения с едно неизвестно. Те разбират някои свойства на триъгълниците и имат основни геометрични представи, включително и за симетрия и ротация. Познават идеята за проста вероятност. Могат да четат и интерпретират данни от графики, таблици, карти и скали.

Ниво на постижение	**TIMSS-1999**	**TIMSS-2003**
ВП	Учениците могат да прилагат знанията и уменията си в относително по-трудни и разнородни ситуации. Те извършват пресмятания с дроби за решаване на задачи с думи; решават многостъпкови задачи с думи включително с използване на пропорционалност; решават задачи от елементарна вероятност; използват знания за геометричните фигури за решаване на задачи; извършват пресмятания с алгебрични изрази и решават (линейни) уравнения с едно неизвестно.	Учениците могат да прилагат знанията си в широк спектър от сравнително сложни задачи. Те могат да оперират с обикновени и десетични дроби за да решават многостъпкови задачи с думи, включвайки пропорционалност с цели числа. Учениците могат да пресмятат стойност на алгебричен израз, да решават системи линейни уравнения, да използват дадена формула за пресмятане на стойност на променлива. Те могат да намират лице и обем на прости геометрични фигури и да използват геометрични свойства на фигури за решаване на задачи. Могат да решават задачи за намиране на вероятност и да интерпретират данни от различни таблици и графики.
МВП	Учениците могат да организират информация, да правят обобщения и да обясняват решения на нерутинни задачи в нетипични ситуации. Те прилагат числови, геометрични и алгебрични зависимости при решаване на задачи; правят еквивалентни преобразувания на алгебрични изрази.	Учениците могат да организират информация, да правят обобщения, да решават нерутинни задачи и да правят заключения въз основа на статистически данни. Те могат да прилагат числови и алгебрични знания и зависимости за решаване на задачи. Могат да решават системи от линейни уравнения. Могат да прилагат знанията си от измерване и геометрия в сложни задачи. Учениците могат да интерпретират данни от различни таблици и графики, включително да правят интерполация и екстраполация.

14. Валидност на метода

Естествена критика на метода на опорните точки е използването на експертни оценки на различни места в прилагането му и особено в последната стъпка. Тези критики имат своите основания. Ясно е, че правилното определяне

на математическите знания и умения, заложени във всяка задача, са основата на крайния продукт от прилагането на метода.

През 1992 г. е направено изследване на валидността на метода върху данни на NAEP, като описанието на знанията и уменията на учениците по математика от всяко ниво е правено от две групи експерти независимо една от друга и описанията на двете групи са сравнени (Mullis, I.V.S. & Johnson, E.G., 1994, 893 – 907.). Авторите отбелязват, че описанията на двете групи си приличат и с точност до редакция описват идентично знанията и уменията на учениците в нивата на постижения.

Описанието на постиженията на учениците, което се получава чрез метода на опорните точки, е важно за популяризиране на резултатите от тест. На масовата публика те се представят във вид много близък до това, което групата експерти дава като характеристика на учениците в определените нива.

15. Как българските ученици покриват нивата на постижения?

В края на тази глава ще дадем информация за това как се изменя процентът ученици за България във всяко от нивата на постижения по математика в TIMSS с течение на времето. Това е показано на таблица 33, където се вижда изменението от TIMSS-1995, до TIMSS-1999 и след това до TIMSS-2003 (Банков, К. 2007-А.)

Таблица 33. Процент български ученици достигнали съответните нива на постижения по математика за TIMSS.

	TIMSS 1995	**TIMSS 1999**	**TIMSS 2003**
Много високи постижения	19	17	9
Високи постижения	40	32	19
Ниски постижения	69	67	51
Много ниски постижения	90	90	82

Намаляването на процента ученици във всяко ниво е следствие от драстичния спад на общите постижения. Тук ще коментирам само последния стълб. Той показва, че през 2003 година по математика 18% от българските ученици не са достигнали до нивото с много ниски постижения. Това означава, че 18% от учениците в VIII клас в България на могат да правят онова, което е записано за нивото на много ниски постижения.

ЧЕТВЪРТА ЧАСТ
ИЗСЛЕДВАНЕ НА ПОДГОТОВКАТА НА УЧИТЕЛИ ПО МАТЕМАТИКА

През последните 10-15 години световната научна общност проявява засилен интерес към знанията, които трябва да имат учителите по математика, за да могат успешно да преподават този предмет. Правят се много теоретични и емпирични изследвания. Оказва се, че въпросът какви математически знания са необходими за преподаване на математиката е твърде сложен и до сега не е изяснен.

Един от начините да се получи информация по този въпрос е да се организират широко мащабни оценъчно-диагностични изследвания в областта на подготовката на учители по математика, както и за оценка на знанията на действащи учители по математика. Такива изследвания хвърлят светлина на въпроси като: Какви знания се дават на бъдещите учители по математика? Какви са различията в това отношение между различните държави? Какви знания имат действащите учители по математика? Какви възможности за обучение имат бъдещите учители по математика? Какви възможности за допълнителна квалификация имат действащите учители по математика? Какво е съдържанието на програмите и курсовете за подготовка и квалификация на учители по математика? и др. Провеждането на такива изследвания е сравнително нова дейност. Въпреки че в национален мащаб подобни изследвания са правени (Ball & Bass, 2000 и Ball, Lubienski & Mewborn, 2001), единици са международните такива. На автора са известни MT21 (Mathematics Teaching in the 21st Century), TEDS-M (Teacher Education and Development Study in Mathematics), TALIS (Teaching and Learning International Survey), от които първите две изследват подготовка на учители по математика, а последното – на действащи учители.

Основната информация за тази част от дисертационния труд са участията на автора в изследванията MT21 и TEDS-M, както и получените резултати от тях (Банков 2007-Б, Банков 2008, Банков 2010, Schmidt et al. 2007, Schmidt et al. 2011).

ТРИНАДЕСЕТА ГЛАВА
УЧИТЕЛИТЕ ПО МАТЕМАТИКА И ТЯХНАТА ПОДГОТОВКА

Световната научна общност, която се занимава с математическо образование, отбелязва следните етапи в развитие на изследванията в тази област през последния половин век. През годините 1960-1980 най-голямо внимание се отделя на *учебните планове* (curriculum development). Следващите 20 години, от 1980 до 2000, вниманието се пренася върху изследвания на *постиженията на учениците* (student assessment). От 2000 година насам се забелязва засилен интерес към *учителите по математика и тяхната подготовка* (teacher preparation). Това е съвсем логично. В търсене на обяснение за различията в ученическите постижения изследователите са установили, че един от основните фактори за тяхното правилно разбиране е учителският колектив, който провежда обучението.

Когато говорим за влиянието на учителите върху ученическите постижения, без съмнение, най-основната учителска характеристика е подготовката, която има даден учител, за да преподава предмета си. Следователно, естествено е да очакваме засилен интерес към изследвания на подготовката на учители. Такива изследвания са база за оценка на влиянието на институционалните и организационни политики за обучение на учители върху качеството и дълбочината на знанията, които придобиват бъдещите учители.

1. Необходимост от познаване на знанията на учителите по математика

С всяко появяване на данни от изследване на знанията на учениците по математика у нас нараства безпокойството от математическата подготовка на българските ученици. Резултатите от TIMSS показват непрекъснат спад на ученическите постижения по математика в VIII клас в България. Подобни резултати се отчитат на националните изпити по математика за прием в профилирани училища след завършен VII клас, както и на кандидат студентските изпити. Вече не е тайна, че образованието в България, и в частност обучението по математика, е в дълбока криза.

През последното десетилетие подобни безпокойства се забелязват и в други държави. Изследователите правят много опити, за да установят причините за спада на ученическите постижения. Без съмнение, тези причини може да

са различни в различните места по света. Но има някои, които се оказват общи: организацията на училището и условията за обучение, неподходящи учебни програми и учебни пособия, и, не на последно място, неадекватна подготовка на учителите по математика. Тук ще се спрем на последната от тези причини.

Неотдавнашно изследване на OECD за подбор, обучение и задържане на учителите в професията показва, че подготовката на учители се различава много в участващите 25 държави (OECD, 2005). В някои от тях студентите започват обучението си за учители веднага след завършване на средното образование, а в други – след като те имат известен опит като „преподаващи" или са учили известен брой години във висше училище. В някои държави подготовката на бъдещи учители става само в университетите, а в други – в тяхната подготовка са намесени и други институции (политехнически институти, колежи и др.). На някои места подготовката на учители става чрез едновременно обучение по всички предмети (concurrent program), а на други – обучението по педагогика и методика на преподаването се извършва в така наречената „втора фаза", след като най-напред студентите имат завършено университетско образование (бакалавърска степен) по някой предмет (consecutive program). Продължителността също варира между 3 и 8 години. Навсякъде подготовката на учители съдържа обучение по: (1) предмета, който се очаква да преподават; (2) обща и специална педагогика; (3) психология на обучението; (4) практическа дейност. Разбира се, акцентите са различни в различните държави.

Въпреки отбелязаните различия, съществува консенсус относно факта, че определен вид знания са особено важни за учителите. В частност, за учителите по математика, единодушно се изтъква, че учители, които имат добра математическа подготовка, по-добре се справят с предизвикателствата на професията. Единодушно е мнението, че знанията по дидактика на математиката също са основни за учителите по математика. Също така без съмнение е важността на подготовката по обща педагогика.

Липсата на универсална общоприета дефиниция за „знание, необходимо за учител по математика" се компенсира с разработването на общи концепции за това, на което трябва да се учат бъдещите учители по математика. Като синтез на тези разработки е публикацията на Шулман (Shulman, 1987), в която ясно се очертава необходимостта от трайни и задълбочени знания в три направления: *математика, дидактика на математиката и обща педагогика*.

Някои автори (Putnam & Borko, 2000) застъпват тезата, че теоретичните знания на учителя, колкото и да са важни, не са единствено решаващи. Нужно е още учителят да има така наречените „общи знания за обучението". Те са свързани със знания и разбиране на контекста, в който се извършва обучението, и с възможностите учителят да прилага придобитите знания в практическа обучаема ситуация. Тук е мястото да се цитира ясната позиция на Ball, Lubienski & Mewborn (2001), че: „Онова, което има значение, в края на краищата, са не само курсовете, които учителят е прослушал, и знанията, които има, но също до колко и как той може да приложи математическите си знания в процеса на своята работа." Трябва да се отбележи обаче, че голяма част от програмите за обучение на учители по математика са основани на разбирането, че те трябва да дадат солидна теоретична подготовка, оставяйки онова, което тук нарекохме „общи знания за обучението" да бъде придобито по-късно с течение на учителската практика.

Според някои изследвания, които сравняват как работят учителите с продължителен стаж (да ги наречем за краткост „експерти") и тези, които са в началото на учителската си кариера (да ги наречен за краткост „начинаещи") се забелязва следната разлика: „Учителите експерти имат добре развита система от знания за учебните планове, организацията на класната дейност и учениците, което им позволява да прилагат общите си знания във всеки частен случай. Там, където начинаещите обръщат внимание на повърхностни белези или конкретни обекти, експертите използват запасите си от знания, организирани в обобщени понятия, свързани с обучението..." (Munby, Russel & Martin, 2001). Програмите за обучение на учители разчитат, че при наличието на богати теоретични знания, практическите умения могат да се придобият в първите години от практикуване на професията и тогава комбинацията от двете ще доведе до експертното ниво на знания. Няма обаче достатъчно факти, които да покажат, че този модел работи добре, за да „научи" един учител по математика как по-ефективно да преподава предмета си.

Това, че качеството на работата на учителя по математика зависи от математическите знания, които той има, не е изненадващ. Не е изненадващо също, че в много държави, включително и в България, има учители по математика, чиято математическа подготовка е далеч под необходимата. За съжаление, обаче не е толкова ясно какво лекарство трябва да се приложи срещу тази болест.

В литературата могат да се намерят различни предложения. Някои предлагат просто задължително да се изучава повече математика в специалностите за подготовка на учители. Други предлагат по-практичен подход като в специалностите за подготовка на учители се акцентира повече на изучаване на математика, която се преподава в училище, изучаване и работа с учебните програми по математика, анализиране на ученически твърдения и др. подобни дейности и знания, свързани директно с работата на учителя в класната стая.

В основата на всички тези предложения стои въпросът какви математически знания са необходими за професията „учител по математика". Нужни ли са знания по всички математически курсове, които се преподават на студентите-математици в математическите факултети (напр. абстрактна алгебра, функционален анализ, комплексен анализ, диференциални уравнения)? Или пък задълбочени знания само по онази математика, която се преподава в училище е достатъчна? С други думи, какви са онези професионални математически знания, които трябва да се преподават на бъдещите учители по математика?

В литературата са описани някои изследвания за влиянието на математическите знания на учителите върху постиженията по математика на техните ученици. От резултатите, обаче не може да се прецени същността на математическите знания на учителите, които имат положително влияние. Нещо повече, дори обемът им също не е определящ. Например Begle, 1979, установява, че изучаването на повече теоретична абстрактна (висша) математика (напр. абстрактна алгебра, функционален анализ, комплексен анализ, диференциални уравнения) от учителите води по повишаване на постиженията по математика на съответните ученици в едва 10% от случаите, а ефектът е отрицателен в 8% от случаите. Предположението, че разковничето е в изучаването на повече „дидактика на математиката" (курсове, насочени към това как да се преподава математика на ученици) също няма достатъчно емпирични доказателства. Много малко са изследванията, насочени към въпроса за математическите програми (чисто теоретично математически, с училищно практическа насоченост, или едновременно и двете), които са подходящи за обучение на учители по математика (Wilson and Berne, 1999).

И така, под „математически знания, необходими за преподаване на математиката" разбираме съвкупността от математическите знания, идеи, умения, методики, технологии, практически опит и др., които са необходими, за да може

един учител по математика да се справи успешно с предизвикателствата на професията си. Усилията на много учени са насочени към конкретизиране на това понятие.

2. Професионални изисквания към учителя по математика

Един от начините да се опишат математическите знания, необходими за преподаване на математиката, е да се наблюдава дейността на учителя по математика и да се отчита какви знания са необходими в ежедневната му работа. В този параграф се прави подобен (макар и далеч не пълен) анализ. Основните въпроси са: „Какви математически знания са необходими на учителя в процеса на обучение на учениците?”, „Кога и как се използват математически знания в процеса на обучение?”, „Как математическите знания на учителя се преплитат с друг вид знания в процеса на обучение?”.

Трудно могат да се опишат накратко професионалните изисквания към учителя по математика. Най-общо казано, той трябва да помогне на учениците да развият математически способности и да осмислят математиката като система от човешки идеи, които да прилагат в различни ситуации. За да стане това, учителят трябва да научи учениците да разбират и използват факти, понятия, принципи и процедури, сътворени от други хора – натрупаните в математиката вековни знания. Той трябва да ги научи още да „правят математика”, да създават и развиват математическата си интуиция, да откриват и да аргументират идеите си. Това означава, че добрият учител по математика трябва не само добре да разбира математическите понятия и процедурите, но и да има познания за математиката: като например какво означава „да се изгради математическа теория”, кои аргументи могат да се приемат за валидни (според възрастта на ученика, например). С други думи, необходими са математически знания извън чисто теоретичните такива. Може да кажем, че това са знания по дидактика на математиката.

За да покажем колко сложна система са професионалните изисквания към учителя по математика ще разгледаме един пример. Учителят (У) е завършил изучаването на обикновени дроби в пети клас. Той поставя на учениците задача да покажат дали е вярно твърдението „ако вземем обикновена дроб, на която числителят е с 1 по-малък от знаменателя, може да се намери правилна

дроб, която е по-голяма от нея”. След като учениците изпробват няколко конкретни примера, учителят се обръща към тях:

У: „Успя ли някой да намери обикновена дроб, на която числителят е с 1 по-малък от знаменателя и за която няма правилна дроб по-голяма от нея?”

Ученик А: „Аз пробвах много примери и винаги успявах да намеря правилна дроб по-голяма от първоначалната. Затова си мисля, че твърдението е вярно.”

Ученик Б: „Аз също мисля, че твърдението е вярно. Но как може да съм сигурен? Има безбройно много обикновени дроби, на които числителят е с 1 по-малък от знаменателя. Не може да проверим за всички тях.”

Ученик В: „Да, но колкото повече примери имаме, които подкрепят твърдението и понеже нямаме пример, който не го подкрепя, толкова повече може да сме сигурни, че то е вярно.”

У (към ученик В): „Покажи ни някои от твоите примери.”

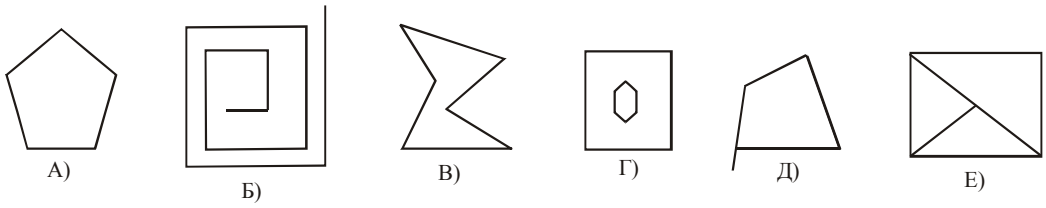
Ученик В: „$\frac{2}{3} < \frac{3}{4}$, или $\frac{7}{8} < \frac{8}{9}$, или $\frac{15}{16} < \frac{16}{17}$. Всяко неравенство се проверява като дробите се приведат към общ знаменател.”

Ученик Г: „Няма нужда да привеждаме към общ знаменател. Например $\frac{15}{16}$ означава да разделим нещо на 16 части и да вземем 15 от тях; значи сме оставили една част от шестнадесетте. От друга страна, $\frac{16}{17}$ означава да разделим същото нещо на 17 части и да вземем 16 от тях; значи сме оставили една част от седемнадесетте. Но една част от нещо разделено на 17 части е по-малко от една част от същото нещо разделено на 16 части. Значи, във втория случай сме взели повече.”

Каква математика е нужно да знае учителят в описаната ситуация? Без да се впускаме в подробности е ясно, че преди всичко учителят трябва да може да прецени ученическите изказвания. Учениците А, Б и В дават различни доводи за това, че твърдението „изглежда е вярно”. Учителят трябва да прецени валидността на всеки от тях, да установи математическата идея в тях и да прецени как може да помогне на учениците да разберат несъстоятелността им. Освен това, учителят трябва да прецени (най-напред за себе си) до колко аргументацията на ученик Г може да се приеме за убедителна и да обясни това на класа. Като се вземе пред вид, че петокласниците не са особено силни в използването на математическа символика, може да се предположи, че формално математическо доказателство на твърдението е извън силите им. Какви разсъждения тогава могат да се приемат за „достатъчно силни” от математическа гледна точка и все пак

достъпни за петокласниците? Изпълняват ли тези условия аргументите на ученик Г? Може ли това да се приеме за доказателство?

Ето още един пример, който показва спецификата на ежедневните математико-педагогически ситуации, в които може да попадне учителят по математика. Урокът е за многоъгълници в шести клас. След кратко обяснение за това какво е многоъгълник, учителят поставя на учениците задачата да начертаят многоъгълници. В тетрадките той вижда начертани следните фигури (Фиг. 82).



Фиг. 82

Учителят трябва да обясни кои от тези фигури са многоъгълници и кои не са. Естествено, най-напред учителят си припомня математическото определение за многоъгълник: „*Проста затворена начупена равнинна линия се нарича многоъгълник.*" Може ли това определение да му свърши работа? Бързо съобразява, че не може. Петокласниците не знаят смисъла на голяма част от думите в определението („проста", „затворена", „начупена") и следователно не могат да го използват. Нужно е да се намери „определение", което да бъде разбираемо за тях и, разбира се, да бъде математически коректно. Например: „*Многоъгълник се нарича редица от отсечки в равнината като всяка следваща започва от точката, където завършва предишната и последната отсечка свършва в точката, от която започва първата. Освен крайните точки, общи за две последователни отсечки, никои две отсечките нямат други общи точки.*"

Този пример показва, че дейността на учителя по математика изисква специфичен усет към използване на математическия език и абстракция. От една страна учителят трябва да е наясно какво знаят неговите ученици (включително и познания за учебните програми), за да не излиза извън тези рамки, а от друга страна, трябва да може да прецени дали дадена дефиниция (твърдение и др.) е математически коректно. Като казва на ученици от първи клас „вие не може да изваждате по-голямо число от по-малко" учителят трябва да е уверен, че казва нещо прагматично за случая, но то не е математически факт. Друг хубав пример в това отношение е въвеждането на ирационалните числа в училище. Безнадеж-

дно е да си мислим, че някой ученик ще разбере точното математическо определение за ирационално число, така, както се прави в университетските курсове по математика. Но далеч по-голям шанс имаме да бъдем разбрани от учениците, ако кажем, че „*ирационално число е всяка безкрайна непериодична десетична дроб*”. Последното е математически коректно и разбираемо от учениците в съответния клас.

Така се приближихме до друго професионално изискване към учителя по математика: да умее да представя сложни математически идеи във вид, в който те могат да бъдат представени на учениците. Една от силните страни на математиката е възможността идеите да се „компресират” в силно абстрактен вид. Така, използвайки подходящи символи за представяне на математическите абстракции, математиците изучават съответните обекти и структури, въвеждайки нови идеи в същия абстрактен вид. В работата си учителят трябва да „разопакова” онова, което математиците са компресирали. Например, когато учениците се запознават с обикновени дроби, това не става като се започва с реални числа, нито дори с рационални числа. Те започват с някои „ежедневни” ситуации, като например нужда за означаване на части от цяло, разделяне на количества (цели числа) на няколко равни части и т.н. Учителят не може разбираемо да обясни понятието „обикновена дроб”, ако работи с абстрактното понятие за реално число или дори с формалната дефиниция за рационално число. Ето още няколко примера в това отношение. Как да обясним формулата за лице на кръг, например, на ученици, които не знаят понятието определен интеграл? Нужно е по подходящ начин да се представи математическата идея с последователните приближения с правоъгълници. Аналогичен е и проблемът с изучаването на нормалното разпределение в часовете по вероятности и статистика. Изискванията към учителя са не само да разбира математическите идеи, но и да може да ги обясни на учениците по разбираем за тях начин.

Тези примери показват, че математическите знания, необходими за преподаване на математиката, включват не само познаване на самата математика, но и съществени елементи от дидактика на математиката – неща, които не са така очевидни и ясно изразени в съдържанието на учебниците по математика.

От гледна точка на обучението по математика, математическата наука може да се разглежда като „съвкупност от последователни и логически свързани идеи” (Good, Grouws & Ebmeier, 1983) и „да се знае математика” означава да се

разбират тези връзки. Една от целите на обучението е учениците да запомнят онова, което учат, но предположението е, че те по-лесно ще го запомнят, ако го разберат. Затова голямо внимание се отделя на този елемент от обучението по математика. От гледна точка на дидактиката на математиката, обаче освен изучаването и познаването на математическите връзки, важно е още да се разбере какво означава „да се изгражда математическа теория", „да се дават валидни аргументи", „да се оценяват аргументи, изказани от други", „да се осъзнае значението на математиката зад границите на ежедневието" и др.

Ако внимателно се анализира работата на учителя по математика по начин, по който направихме това с примерите по-горе, ще се очертаят следните важни дейности на учителите, които изискват специфични математически знания:

- Конструиране на математически коректни обяснения, които са разбираеми и използваеми от учениците.
- Използване на математически коректни и разбираеми дефиниции.
- Представяне на математическите идеи по разбираем начин чрез използване на нагледност, графични модели, символи и др.
- Интерпретиране и преценка (от математическа и педагогическа гледна точка) на ученически въпроси, решения, задачи, идеи (както предвидими, така и неочаквани).
- Адекватно реагиране на куриозни математически изказвания от ученици.
- Преценка на качеството (от математическа и педагогическа гледна точка) на учебници и учебни помагала.
- Поставяне на математически въпроси и задачи, които стимулират учениците.
- Оценяване на математическите постижения на учениците и използването на резултатите за педагогически цели.

Изхождайки от тези професионални изисквания към учителя по математика може да се мисли кои са онези математическите знания, необходими за преподаване на математиката, т.е. каква трябва да е онази математика, която е най-подходяща за обучение на учители по математика.

ЧЕТИРИНАДЕСЕТА ГЛАВА
ИЗМЕРВАНЕ НА МАТЕМАТИЧЕСКИТЕ ЗНАНИЯ НЕОБХОДИМИ ЗА ПРЕПОДАВАНЕ НА МАТЕМАТИКАТА

Тази глава е посветена на изследванията MT21 и TEDS-M. Разгледа се по-подробно онази част от рамките им, която е свързана с концепциите за оценяване на знанията по математика, необходими за преподаване на математиката. Авторът на този труд е работил в тези две изследвания от самото им начало. Така той естествено е един от авторите на тези концепции, на рамките и тестовите спецификации, както и на не малка част от задачите за измерване на постиженията на бъдещите учители по математика.

1. Методологична основа на изследванията MT21 и TEDS-M

Изследванията MT21 (Mathematics teaching in the 21st century) и TEDS-M (Teacher Education and Development Study – Mathematics) фокусират върху подготовката на учители по математика за началната и прогимназиалната степен на средното училище. Разработването на рамките на MT21 и на TEDS-M е направено въз основа на разбирането, че знанията на бъдещите учители по математика трябва да се изграждат в три основни направления: *математика, дидактика на математиката, обща педагогика* (Shulman, 1987). Това означава, че за измерване на техните знания трябва да се конструират три основни скали – по една за всяко от тези направления. Тук ще разглеждаме само скалите по математика и дидактика на математиката и няма да се занимаваме със скалата по обща педагогика.

Направлението *математика* включва онази част от „университетската математика", която обикновено се преподава на студентите-математици във факултетите по математика, както и „училищна математика" (т.е. математика, която се преподава в училище) като математическа дисциплина, без да се изучават методите за преподаване. За класификация на задачите по математика в MT21 е използвана следната структура: Алгебра, Данни и вероятности, Функции, Геометрия, Числа. Това означава, че по математика са конструирани 5 под-скали – по една за всяка от изброените области. Структурата на съответните под-скали по математика в TEDS-M е нарочно взета от TIMSS-2003 и тя е близка до тази

на MT21 (Алгебра и Функции са обединени), а именно: Числа, Геометрия, Алгебра, Данни. Таблица 34представя по-подробно съдържанието на тези области.

Таблица 34. Области по математика за MT21

Числа	• Естествени числа • Обикновени и десетични дроби • Числови изрази • Числови зависимости • Цели числа • Отношение, пропорция и процент • Ирационални числа • Комплексни числа • Теория на числата
Геометрия	• Прави и ъгли • Двумерни и тримерни геометрични фигури • Измерване в геометрията • Координатна геометрия • Геометрични преобразувания
Алгебра	• Алгебрични изрази • Линейна и абстрактна алгебра
Функции	• Зависимости • Уравнения, формули и функции • Математически анализ
Данни и вероятности	• Организиране и представяне на данни • Четене и тълкуване на данни • Вероятност

Структурирането на математическите области за MT21 и за TEDS-M е основано на задълбочени изследвания на учебните програми по математика в училище в над 40 държави, направени в рамките на изследването TIMSS (Schmidt et al., 1997), както и от изучаване на училищната математика, която учителите трябва да преподават, в държавите, участващи в TEDS-M. Това изследване е направено от Maria-Teresa Tatto и Кирил Банков (не е публикувано до момента) чрез съвременна методика за анализ на учебна документация. Тя дава възможност да се сравняват разнородни по форма и по език документи, свързани с учебните програми. Документите се „разделят” на малки части (блокове) и всяка от тях се кодира с код, който описва съдържанието по отношение на предварително фиксиран списък от математически теми. Фигура 83 представя обобщени резултати от тези изследвания. Както се вижда, преобладаващите теми, които учителите би трябвало да преподават в училище, могат да се групират в казаните по-горе математически области на MT21 и TEDS-M.

Тема / Клас	1	2	3	4	5	6	7	8	9	10	11	12
Цели числа и тяхното значение	●	●	●	●	●	●						
Действия с цели числа	●	●	●	●	●							
Единици за измерване	●	●	●	●	●	●	●					
Обикновени дроби		●	●	●	●	●	●					
Уравнения и формули				●	●	●	●	●	●	●	●	●
Представяне и анализ на данни		●	●	●	●	●	●	●	●		●	●
Основни понятия в равнинната геометрия			●	●	●	●	●	●	●			
Многоъгълник, окръжност и кръг		●	●	●	●	●	●	●	●	●	●	
Измерване: обиколка, лице и обем			●	●	●	●	●	●	●		●	
Закръгляване					●	●	●					
Приближено смятане				●		●						
Свойства на операциите с цели числа		●	●	●	●	●	●					
Оценка на количества и размери		●	●	●	●							
Десетични дроби			●	●	●	●						
Връзка м/у обикновени и десетични дроби				●	●							
Свойства на обикновени и десетични дроби					●	●	●					
Процент					●	●						
Понятието пропорционалност						●	●	●	●			
Свойства на пропорционалността					●	●	●		●			
Едномерни и двумерни координати		●			●	●	●	●		●	●	
Геометрични преобразувания		●				●	●	●	●			
Цели числа. Свойства						●	●	●		●		●
Теория на числата						●	●	●	●			
Степени и корени							●	●	●	●		
Степени. Порядък на величина												
Оценка на грешката при измерване		●	●	●	●	●						
Построения с линийка и пергел								●				
Стереометрия		●	●	●	●	●	●	●			●	●
Еднаквост и подобност					●			●	●			
Рационални числа и техните свойства												
Закономерности, релации и функции								●	●	●	●	●
Тригонометрия								●	●	●		
Реални числа. Подмножества. Свойства										●		
Обосновка и доказателство							●	●	●	●	●	●
Абстрактност в математиката												
Комплексни числа. Свойства												
Вероятност						●	●	●	●		●	●
Безкрайни процеси											●	●
Растене и намаляване												
Вектори											●	
Комбинаторика												

Темата се изучава в поне 3 държави от TEDS ●

Темата се изучава в поне половината "най-добре" представили се държави в TIMSS

Фиг. 83. Преобладаващи теми в училищните учебни програми

Името на направлението *дидактика на математиката* не е унифицирано в международен план. Някъде то се нарича *методика на математиката*, на други места *педагогика на математиката*. Въпреки различните имена, съдържанието е еднакво. Тук се включват знания, които са специфични за цялостния процес на обучение по математика в училище. Като се имат пред вид дейностите на учителя, могат да се разграничат три основни области за оценяване на тези знания, представени в Таблица 35.

Таблица 35. Области по дидактика на математиката

Знания за учебните планове	• Дефиниране на подходящи цели на обучението по математика • Познаване на различни формати за оценяване на учениците • Разбиране на връзките между темите в учебните планове и подбиране на подходяща тематична последователност • Откриване на основните идеи в учебните планове • Познаване на плановете за обучение по математика
Знания за преподаване на математика	• Планиране и подбиране на подходящи дейности • Подбиране на формат за оценяване на учениците • Предсказване на типични ученически отговори, включително и типични грешки • Планиране на подходящи методи за обучение по математика • Свързване на дидактическите методи със структурата на урока • Подбиране на различни начини за решаване на математическа задача • Планиране на уроците по математика
Контакт „ученик-учител”	• Анализиране и оценяване на ученически решения и аргументация • Анализиране смисъла на ученически въпроси • Диагностициране на типични ученически отговори, включително и типични грешки • Обяснение и представяне на математически понятия и процедури • Задаване на подходящи въпроси • Адекватно реагиране на неочаквани математически ситуации • Обезпечаване на подходяща обратна връзка

Едната от тях са общите знания за целите на обучението по математика, учебните планове и разбиране на основните връзки между темите в учебните планове. Тази област може да наречем *знания за учебните планове*. Втората област е свързана с дейността на учителя преди влизането му в класната стая (подготовка за урока), а именно подбиране на подходящи материали и дейности за обучение, определяне на различни подходи за обучение по определени теми от учебните планове, предвиждане на евентуални ученически отговори, включително и неправилно разбиране на материала и типични грешки. Тази област наричаме *знания за преподаване на математиката*. Третата област фокусира

върху знания, свързани с адекватното поведение на учителя в часовете по математика, а именно обяснения на понятия и процедури, анализиране на ученически въпроси и отговори, осигуряване на подходяща обратна връзка и др. подобни. Тази област може условно да наречем *контакт „ученик-учител”*.

Задачите за постижения по математика и по дидактика на математиката в MT21 са класифицирани в 4 познавателни равнища: Познаване на понятия и използване на алгоритми, Решаване на задачи, Аргументиране и доказване, Моделиране на задачи от реалността. За задачите в TEDS-M са използвани трите познавателни равнища от TIMSS-2003: Познаване, Приложение, Аргументиране.

Освен разглежданите скали за постижения, в MT21 и TEDS-M са конструирани скали за „възможности за обучение” (какви курсове и други форми на обучение се предлагат на бъдещите учители по математика) и за „отношение към математиката” (разбирания за същността на математиката, за целите на обучението по математика, как учениците изучават математика, как се обучават ученици с големи различия в знанията по математика и др.) Описанието на тези скали и резултатите по тях не са обект на тази работа.

При съставянето на задачите за постижения по математика и дидактика на математиката трябваше да се търсят отговори на някои въпроси от методологично и технологично естество. Първо, в двете основни скали има много подскали (общо 8 за MT21 и 7 за TEDS-M). За да се получат надеждни резултати по всяка от тях, нужно е да има „достатъчно много” задачи към всяка една подскала. Това води до използване на тест с много задачи. От друга страна, общото тестово време не може да бъде дълго. Беше прието, че повече от 90 минути работа върху въпросник ще доведе до отказване на студентите – бъдещи учители да работят върху него. Като се вземе пред вид необходимостта от време за въпроси, за другите скали, за знанията по математика и дидактика на математиката остават около 45 минути. Разбира се, използването на вероятностното моделиране (вж. четвърта глава) дава възможност целият тест да се раздели на части и отделни групи бъдещи учители да работят върху различни части. Този метод на скалиране беше възприет, но поради наличието на курсове с малък брой студенти в някои от участващите държави, се оказа, че броят на тестовите книжки не може да е голям.

2. Съставяне на задачи

Един от сериозните проблеми на оценяването в MT21 и TEDS-M е осигуряването на „достатъчно добро" участие на бъдещите учители. Тестът е предназначен за възрастни хора. Тяхното отношение към оценяването им е резервирано. Освен това, известно е, че студентите не са редовни посетители на учебните занятия. Всичко това може да доведе до получаване на нисък процент отговори от някои институции. За частично преодоляване на този ефект трябваше да се помисли за стимулирането на бъдещите учители за активното им участието в експеримента. Един от начините за стимулиране е да се подготвят „интересни" тестови задачи.

Така съставянето на задачи поставя следните предизвикателства пред авторите: (1) задачите да измерват знанията по математика и дидактика на математиката, които са фиксирани в тестовите спецификации; (2) задачите да дават възможност да се конструират скали и под-скали (т.е. те да могат да се „точкуват"); (3) задачите да са привлекателни за бъдещите учители по математика. Един от начините да се обединят тези изисквания е да се поставят математически задачи в подходяща педагогическа ситуация. Идеята е да се измерят знания по математика и по дидактика на математиката чрез задаване на дидактически ситуации. Такива задачи са подходящи поради това, че чрез тях бъдещите учители по математика се поставят в обстановка, която е възможно най-близка до реалната работа в училище. Предполага се, че това би трябвало да им е интересно, защото е свързано с тяхната бъдеща професия. Освен това такива задачи биха могли да хвърлят светлина върху въпроса, който разисквахме в предишната глава относно знанията по математика, необходими за преподаване на математиката.

Така се появява друг проблем – проблемът с класификацията на задачите. Към коя от двете скали (математика или дидактика на математиката) трябва да се причисли задача, която поставя математически проблем в дидактическа ситуация? Консултациите, които бяха проведени с водещи психометрици, доведоха до благоприятно решение – възможно е една задача да бъде причислена едновременно към двете скали. От съдържателна гледна точка това е допълнително преимущество, не само защото общият брой на задачите можеше да бъде намален, понеже някои от задачите са причислени към двете скали, но и

защото с такива задачи още по-тясно може да се оцени връзката между чисто математическите знания и знанията, необходими за обучението по математика.

По-долу са представени задачи от МТ21, които са предложени от автора на този труд. За всяка задача са коментирани отнасянето й към спецификациите на теста; конструктите, които тя измерва; връзката на задачата със знанията, необходими за преподаване на математика. Освен това за всяка задача е даден процентът на българските бъдещи учители по математика, които са отговорили по отделните възможности за отговор (Банков, 2007-Б). Най-напред са разгледани задачите към скалата по математика, след това – онези, които са класифицирани по двете скали (математика и дидактика на математиката) и на края – задачите от скалата по дидактика на математиката.

Започвам с важната за българското училище тема „уравнения”.

Пример 1. Ако $a > 0$, колко различни реални решения има уравнението $x^2 + x - a = 0$?

*Отбележете **едно** квадратче.*

1	4,9
2	46,3*
Безбройно много.....	2,4
Броят на реалните корени зависи от стойността на *a*.	37,8

Без отговор: 8,5

В тази задача, както и във всички останали, в последната колонка вместо квадратчета за отбелязване на отговора (както е в оригиналния тест), са поставени процентите на българските бъдещи учители по математика, отговорили по съответните възможности за отговор. Правилният отговор е отбелязан със *.

Задачата от Пример 1 е от скалата по математика от областта *Функции*. Тази задача не поставя педагогическа ситуация, но е свързана с преподаването на важни знания относно квадратния тричлен – положението му относно абсцисната ос в зависимост от свободния член. Разбирането на това е ключов момент в работата на учителя по математика гимназиалните класове.

Представянето на българските бъдещи учители по математика върху тази задача не е добро. По-малко от половината от тях са дали правилен отговор и повече от една-трета считат, че броят на реалните корени зависи от стойността на *a*. Това говори за недобро разбиране на онази математика, която се оценяват тази задача.

Пример 2. За всяко от следващите уравнения, отбележете най-малкото множество от числа, на които принадлежат *всичките* му решения.

*Отбележете **едно квадратче** на всеки **ред**.*

	Естествени числа	Цели числа	Рационални числа	Реални числа	Комплексни числа	Без отговор
$x+2=3$	70,7*	9,8	3,7	4,9	1,2	**9,8**
$x^2-4=0$ ...	7,3	59,8*	8,5	8,5	1,2	**1,2**
$x+3=2$	12,2	52,4*	13,4	9,8	0,0	**12,2**
$x^2-2=0$ ...	2,4	4,9	22,0	42,7*	13,4	**14,6**
$2x-3=0$ ...	1,2	4,9	61,0*	14,6	7,3	**11,0**
$x^2+2=0$ ...	2,4	4,9	3,7	2,4	70,7*	**15,9**

Последната колонка в тази задача (както и в някои от следващите задачи) липсва в оригиналния тест. Тя е добавена, за да покаже процента на неотговорилите по всяка от опциите.

Задачата е от скалата по математика. Класифицирана е към областта *Алгебра*. Същността на задачата не е да се решат уравненията, а да се проверят знанията за множествата от числа. Тези знания са основни за учителя по математика. В училище се изучават последователно всички числа (само комплексните числа са в програмата за профилирана подготовка). Ето защо учителите трябва добре да разбират разширението на понятието число, като се започне от естествените числа и се стигне до реалните (дори и до комплексните). Задачата не поставя дидактическа ситуация и затова не е причислена към скалата по дидактика.

Резултатите на българските бъдещи учители по математика са сравнително добри. Може би трябва да се обърне внимание на четвъртото уравнение. Дали наистина някои не знаят, че $\sqrt{2}$ е ирационално число или причината за грешните отговори (рационални числа, комплексни числа) е в нещо друго? Висок е процентът на неотговорилите за последните уравнения. Това може да се дължи на отегчение от „дългата” задача, но може да означава неразбиране на понятието за някой вид числа.

Пример 3. Кое равенство е еквивалентно на $\log_a b = c$?

*Отбележете **едно** квадратче.*

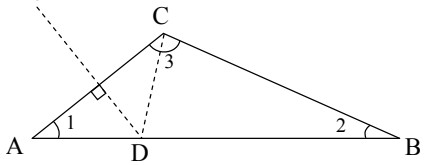
1. $a^b = c$	5,1
2. $a^c = b$	55,7*
3. $b^a = c$	6,3
4. $c^a = b$	21,5

Без отговор: 11,4

Задачата от пример 3 е класифицирана към областта *Алгебра*. В задачата няма дидактически момент и тя не е към скалата по дидактика на математиката. Задачата проверява знание за определението на логаритъм – нещо, което учителите трябва да преподават в училище. Следователно това е знание, необходимо за директно прилагане в обучението по математика.

Само малко над половината от бъдещите учители по математика в България са се справили успешно с определението за логаритъм. Това наистина буди безпокойство.

Пример 4. В ΔABC, симетралата AC пресича *AB* в точка *D*, така че *AD* = *DB*. Кое от следните твърдения е вярно?



*Отбележете **едно** квадратче на **всеки ред.***

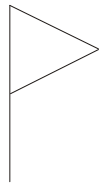
	Вярно	Невярно	Не съм сигурен	Без отговор
Δ *ADC* е равнобедрен	43,9*	17,1	8,5	30,5
$\sphericalangle CDB = \sphericalangle ACB$	9,8	45,1*	11,0	34,1
D е центърът на описаната окръжност за Δ *ABC*	36,6*	15,9	14,6	32,9
$\frac{AD}{DB} = \frac{AC}{BC}$	23,2	25,6*	14,6	36,6
$\sphericalangle 1 + \sphericalangle 2 = \sphericalangle 3$	30,5*	23,2	12,2	34,1
$\sphericalangle 3 = 90^0$	37,8*	14,6	13,4	34,1

Задачата от пример 4 е класифицирана към областта *Геометрия* на скалата по математика. Тя няма дидактически елемент.

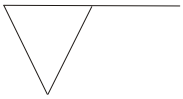
Тази задача проверява способностите на бъдещите учители да решават задачи подобни на онези, които се срещат в учебниците по математика за прогимназията (всички без четвъртото твърдение се изучава в 7 клас в България). Не се изискват допълнителни знания извън онези от задължителния учебен материал. Всеки учител би трябвало да може да се справя с подобен вид задачи, тъй като те се решават в клас.

Резултатите за България показват, че правилните отговори варират между 30% и 45%. Като се има пред вид, че за всяко твърдение възможностите за отговор са три („Вярно”, „Невярно” и „Не съм сигурен”), правилните отговори са малко над теоретичната граница на налучкване. Около една-трета от учениците не са отговорили на твърденията, което най-вероятно означава, че не осъзнават смисъла на задачата. Всичко това поставя сериозно въпроса за подготовката на бъдещите учители по математика в България да решават задачи от учебниците по математика.

Пример 5. Коя от следните трансформации може да преобразува флага от позиция A в позиция B?



A



B

Отбележете едно квадратче.

Осева симетрия	2,5
Транслация	3,8
Последователно прилагане на две осеви симетрии с различни оси	12,7
Последователно прилагане на ротация около точка и осева симетрия	70,9*

Без отговор: 10,1

Задачата от пример 5 е от скалата по математика, област *Геометрия*. Тя проверява знания от училищния курс за геометрични преобразувания. Въпреки че задачата не е от онези, които могат да се намерят в учебниците, тя проверява основни знания, свързани със сравнително не сложни свойства на геометричните преобразувания и техните приложения. Не се поставя дидактическа ситуация и затова задачата не е към областта дидактика на математиката.

Резултатите на българските бъдещи учители по математика са сравнително добри. Една от вероятните причини е, че задачата е от материал, който е добре застъпен в Училищния курс по геометрия в университетите, които подготвят учители по математика.

Следващата задача (Пример 6) е също от областта математика, под-скала Данни и вероятности. Тя няма дидактически момент и не е към областта дидактика на математиката. Задачата оценява тълкуване на статистически данни и използване на информацията от тях за получаване на отговори на „типични въпроси", използващи статистически данни. Редно е да отбележим, че такива знания са слабо застъпени в училищната математика в България. В обучението на бъдещите учители по математика у нас обикновено има курс „Вероятности и статистика", но той предимно разглежда математическата теория и малко се застъпва на приложната част. От друга страна, такъв тип знания се считат необходими за учителите по математика (особено в някои страни на Западна Европа и в САЩ), понеже в тяхната училищна математика те трябва да преподават такъв материал. У нас вече също се засилва обучението по статистика в училище. Въпреки че още няма съществени стъпки в това отношение, много вероятно е в бъдеще задачи като тази от пример 6 да намерят място в нашите учебни планове по математика. От тази гледна точка задачата измерва съществени за учителската професия знания.

Като се има пред вид казаното по-горе, не е чудно, че процентът на бъдещите учители по математика, които не са отговорили на тази задача, е голям (от 50% до 75%). Също така не трябва да се учудваме на малкия процент правилни отговори. Разбира се, това не означава, че нещата трябва да останат такива, каквито са. Нужно е да се помисли за съответно обучение в програмите за подготовка на учители по математика.

Пример 6. На всички ученици от осми клас в едно училище е даден тест, който се оценява по 100-бална система. Следващата таблица показва резултатите.

Брой на точките	***Брой на учениците***
Под 31	26
31-40	18
41-50	16
51-60	20
61-70	21
71-80	19
81-90	11
91-100	5
	Общо: 136

Директорът на училището разделя учениците на четири нива:

Ниво 3 – тези, които имат поне 81 точки
Ниво 2 – тези, които имат между 51 и 81 точки
Ниво 1 – тези, които имат между 31 и 51 точки
Ниво 0 – тези, които имат под 31 точки

Вие имате три задачи.

Задача 1: Покажете, че броят на учениците в Ниво 2 е по-голям от броя на учениците във всяко едно от останалите нива.
Задача 2: Покажете, че училището отговаря на стандарта, според който "в ниво 0 трябва да има по-малко от 20% от всички ученици."
Задача 3: Покажете, че онези 20% ученици, които са с най-високи постижения, идват от повече от едно ниво.

Кое от изброеното по-долу е *най-подходящо* за *всяка* една от горепосочените задачи?

*Отбележете **само едно** квадратче за **всяка задача**.*

	Задача 1	Задача 2	Задача 3	Без отговор
Бих направил хистограма за всички 8 реда на таблицата.	12,7	1,3	11,4	74,7
Бих направил хистограма за четирите, определени от директора, нива.....	27,8*	2,5	7,6	62,0
Бих посочил, че $\frac{26}{136}$ е по-малко от 20%.	0	34,2*	7,6	58,2
Бих посочил, че *Ниво 0* съдържа по-малко ученици, отколкото всяко друго ниво.....	2,5	20,3	1,3	75,9
Ще изчисля процента на учениците на всяко ниво	13,9*	7,6	16,5	49,4
Ще изчисля процента на учениците на *Ниво 3*	2,5	1,3	21,5*	74,7

Задачата от следващия пример е от скалата по математика, област *Функции*. Проверява познаването на дефинициите за непрекъснатост на функция в точка от дефиниционната област. Това е основно математическо знание. Счита

се, че един учител по математика, дори и да не му се налага да преподава за непрекъснати функции, трябва да знае какво е това непрекъсната функция. В този смисъл задачата проверява нещо важно. В задачата няма педагогическа ситуация и затова тя не е причислена към скалата по дидактика на математиката.

Пример 7. Кое от изброените твърдения е правилна дефиниция за непрекъсната функция *f* в точка x_0 от дефиниционната област на *f* ?

*Отбележете **едно** квадратче на **всеки ред**.*

	Да	Не	Не съм сигурен	Без отговор
За всяко $\varepsilon > 0$ съществува такова $\delta > 0$, че за всяко x, за което $\lvert x - x_0 \rvert < \delta$, е изпълнено $\lvert f(x) - f(x_0) \rvert < \varepsilon$	68,3*	8,5	12,2	9,8
За всяко $\delta > 0$ съществува такова $\varepsilon > 0$, че за всяко x, за което $\lvert x - x_0 \rvert < \delta$, е изпълнено $\lvert f(x) - f(x_0) \rvert < \varepsilon$	17,1	46,3*	14,6	2,0
Ако $\lvert x - x_0 \rvert < \delta$, тогава $\lvert f(x) - f(x_0) \rvert < \varepsilon$	13,4	45,1*	20,7	20,7
$\lim\limits_{x \to \infty} f(x) = f(x_0)$	25,6	34,1*	14,6	25,6
За всяка редица $x_1, x_2, \ldots, x_n, \ldots$ клоняща към x_0, (x_n е от дефиниционната област на *f*), редицата $f(x_1), f(x_2), \ldots, f(x_n), \ldots$ клони към $f(x_0)$	46,3*	17,1	17,1	19,5
Съществува редица $f(x_1), f(x_2), \ldots, f(x_n), \ldots$ клоняща към $f(x_0)$	6,5	48,8*	25,6	17,1

Бъдещите учители по математика в България имат сериозна подготовка по математически анализ. Тези дефиниции се обясняват подробно. Очакванията са, че студентите са ги разбрали и могат да ги използват. Резултатите показват, обаче, че само около половината са дали правилни отговори. Голям е процентът на онези, които са отговорили „не съм сигурен” или въобще не са дали отговор. Изводът е, че разбирането на дефинициите за непрекъсната функция не е на високо ниво.

Разпространено е мнението, че бъдещите учителите по математика трябва да учат и да разбират не само математиката, която ще преподават, но и не малка част от математиката, която традиционно се преподава на студентите в математическите факултети. Следващият пример показва една задача, чиято цел

е да оцени математически знания от така наречената „университетска математика”.

Пример 8. Реалната функция *f* се нарича линейна, ако $f(x) = px + q$, където *p* и *q* са дадени реални числа. Трансформацията *F* се нарича линейна, ако $F(x+y) = F(x) + F(y)$ и $F(ax) = aF(x)$ за всяко реално число *a*.

Кое от следващите твърдения описва връзката между линейна функция и линейна трансформация?

*Отбележете **едно** квадратче.*

Няма връзка между тях.	7.6
Линейна трансформация в множеството на реалните числа е линейна функция.	11.4
Линейна функция е линейна трансформация в множеството на реалните числа.	10.1*
Линейните функции и линейните трансформации съвпадат в множеството на реалните числа.	12.7

Без отговор: 58,2

Задачата от пример 8 е класифицирана към скалата по математика, област *Функции*. Тя не поставя дидактическа ситуация и затова не е към скалата дидактика на математиката.

Големият процент на неотговорилите и сравнително близките стойности на процентите по възможностите за отговор недвусмислено показват, че бъдещите учители в България не разбират връзката между линейна функция и линейна трансформация. Този факт вероятно не е много съществен от гледна точка на училищната математика. От друга страна, обаче той показва, че не се разбира нещо основно в така наречената университетска математика, което говори за недостатъчна математическа култура. Съвсем друг е въпросът дали искаме или не нашите учители по математика да притежават такава.

На пръв поглед задачата от следващия пример също оценява знания, които не се преподават директно в училище. Тази задача, обаче проверява разбирането на много важното понятие „релация на еквивалентност”. То се счита за основно в училищната математика и върху неговото разбиране се работи много в курса „Методика на математиката”.

Пример 9. Релация, която е рефлексивна, симетрична и транзитивна, се нарича релация на еквивалентност.

Кое от посоченото по-долу можете да използвате като пример за релация на еквивалентност?

*Отбележете **едно** квадратче на **всеки ред.***

	Да	Не	Не съм сигурен	Без отговор
Релацията „равенство" (=) между реални числа..	5,0*	6,8	68,3	19,9
Релацията „неравенство" между реални числа ...	10,6	57,1*	7,5	24,8
Релацията „подобност" в множеството на всички триъгълници	17,4*	25,5	7,5	24,8
Релацията „роден на същата дата като" в множеството от всички хора.	23,6*	19,9	31,1	25,5
Релацията „има общ множител с" в множеството на естествените числа	30,4	25,5*	14,9	29,2
Релацията „дава един и същ остатък след делене на 7" в множеството на целите числа	26,7*	24,2	21,7	27,3
Релацията „равно на същата десетична дроб" в множеството на всички обикновени дроби	29,8*	16,1	27,3	26,7

Задачата е от скалата по математика. Тя не е причислена към никоя област, тъй като оценява разбиране на универсално понятие. Резултатите на българските бъдещи учители по математика показват неразбиране на понятието *релация на еквивалентност*. Само около една-четвърт са дали правилен отговор и друга една-четвърт въобще не са отговорили. Особено фрапиращо е положението с отговорите на първата релация (равенство). Трудно може да се намери нормално обяснение за получения резултат.

Следващите примери показват задачи, причислени едновременно към двете направления: математика и дидактика на математиката.

В пример 10 се поставя математическа задача в педагогическа ситуация. Математическата област е *Функции*, а педагогическата – *Контакт „ученик-учител"*.

Математическата част на задачата е свързана с решаването на показателни уравнения, които (чрез заместване) се свеждат до квадратни. При това, е възможно някои от корените на квадратното уравнение да не са в областта от стойности на функцията, която заместваме. Разбирането на този факт е много важно за преподаването на математика в гимназиалните класове. От гледна

точка на дидактика на математиката, бъдещият учител се поставя в положение да диагностицира типични ученически отговори, включително и типични грешки. Трябва да се разбере каква точно е грешката на Питър, за да може тя да се обясни и дискутира с учениците. От учителя се иска не само да знае и разбира свеждането на показателни уравнения към квадратни, но и да може да разгадае и обясни типични грешки, които се допускат в използването на този метод.

Пример 10. На учениците е дадена следната задача:

Намерете броя на реалните корени на уравнението
$9^x - 3^x - 6 = 0$.

Питър означава $y = 3^x$ и получава уравнението $y^2 - y - 6 = 0$, което има 2 различни корена. Той заключава, че даденото уравнение също има 2 различни реални корена.
Кое от следните твърдения е вярно за решението и аргументацията на Питър?

*Отбележете **едно** квадратче.*

Заключението и аргументацията на Питър са правилни.	10
Първоначалният подход на Питър към задачата (заместването с $y = 3^x$) не е правилен.	4,4
Питър не разлага правилно лявата и дясната страна на уравнението.	1,3
Квадратното уравнение $y^2 - y - 6 = 0$ няма 2 различни корена..	2,5
Питър не взема предвид областта от стойности и дефиниционната област на функцията $y = 3^x$	50,6*

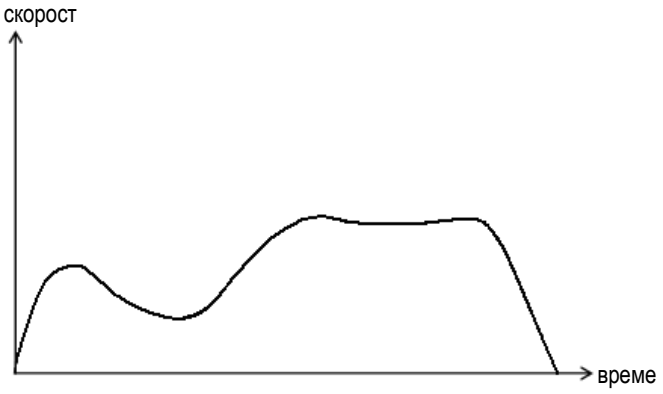
Без отговор: 30,3

Както се вижда, само половината от бъдещите учители по математика могат да правят тези неща. Безпокойство буди фактът, че 30% от тях не са отговорили на задачата. Това означава, че те не са наясно какво точно се случва в тази ситуация. Други 10% намират, че в разсъжденията на Питър всичко е наред. На фона на засиленото изучаване на математически анализ във факултетите по математика, където се обучават тези студенти, подобен род грешки са неоправдани.

Задачата от пример 11 е от областта *Функции* към скалата по математика. Проверяват се знания за разчитане на графика на функция. Това се изучава пе-

риодично в няколко класа в училище и поради това може да считаме, че е основно знание.

Пример 11. Графиката изразява зависимост между скоростта и времето на колоездач.



Ако помолите учениците си в прогимназията да съчинят разказ, илюстриран от графиката, кой от посочените по-долу ще приемете за подходящ?

*Отбележете **едно** квадратче на **всеки ред**.*

	Да	Не	Не съм сигурен	Без отговор
Колоездачът се е изкачил по някакъв хълм, а след това по второ възвишение, по-високо от първото.	23,2	39,0*	9,8	28,0
Колоездачът е тръгнал по равен път, изкачил се е на хълм, слязъл е от него, карал е по равно и е спрял.	19,5*	47,6	6,1	26,8
Колоездачът се е движил известно време, след което е спрял, после отново е тръгнал, спрял е за втори път и е слязъл от велосипеда.	4,9	47,6*	19,5	28,0
Колоездачът се е движил по път, а графиката представлява извивките на пътя.	29,3	34,1*	8,5	28,0

Понеже в задачата има педагогическа ситуация, тя е причислена и към скалата по дидактика на математиката, *Контакт „ученик-учител”*. Дейността на учителя в случая е като анализира ученическите отговори, да установи правилността на всеки от тях. Разликата от задачата от пример 10 е, че тук учителят не търси грешка в решението, а трябва да анализира дали съответният „разказ” може да се илюстрира с дадената графика. Общото с пример 10 е концепцията

за адекватна реакция на учителя към ученически отговори на поставени въпроси.

Постиженията на българските бъдещи учители по математика върху тази задача не са добри. Освен големия процент неотговорили, прави впечатление, че малко са установили като правилен единственият „разказ", който може да се илюстрира с дадената графика. Буди недоумение сравнително високият процент отговори „Да" на първия и четвъртия „разказ". Всичко това говори, че бъдещите учители по математика имат трудности в разчитането на графики на функции (от математическа гледна точка) и се нуждаят от допълнителни упражнения в анализ на ученически отговори. Това е посланието от тази задача към програмите за обучение на учители по математика.

Задачата от пример 12 е причислена към областта *Числа* към скалата по математика и към областта *Контакт „Ученик-учител"* към скалата по дидактика на математиката (диагностициране на типични ученически отговори). Задачата би могла да се причисли и към областта Данни и вероятности, но решението тя да е към областта Числа беше взето след като се прецени, че разбирането на факта, че последната цифра на квадрата на дадено число зависи само от последната цифра на числото, е по-важно за решението на задачата.

Както може да се види от втората част на задачата (която всъщност е задача от дидактика на математиката), двата възможни правилни отговора са свързани именно с разбирането на появата на последната цифра при повдигане на квадрат и с разбирането на понятието вероятност. Правилните отговори на дидактическата част на задачата може да се дадат само ако бъдещите учители добре са разбрали двата важни математически елемента на задачата, а именно разбирането на това как се получава последната цифра на квадрата на целите числа и понятието за вероятност.

Както се вижда от резултатите на българските учители по математика, само малко повече от една трета от тях са посочили правилния отговор на първата част. Резултатът по втората част е още по-обезкуражаващ. Постоянен процент (28%) въобще не са решавали задачата. Тези резултати говорят за поне две неща: (1) за българските бъдещи учители по математика е проблем да разберат двата математически елемента на задачата; (2) сериозен проблем за тях е да разберат трудностите, които биха имали учениците при решаване на тази задача.

Пример 12. На ученици е дадена следната задача.

Джон избира произволно естествено число, повдига го на квадрат и взема последната цифра от квадрата му. Каква е вероятността това число да е 1?

Ето отговорите на три ученички:

Лиза	Моника	Силвия
Общият брой на всички цифри е 10. Всяка от тях има еднаква вероятност да бъде последната цифра. Следователно, отговорът е $1/10$ =10%.	Последната цифра на квадрата зависи само от квадрата на последната цифра на избраното число. Последните цифри на квадратите на първите 10 естествени числа са 1, 4, 9, 6, 5, 6, 9, 4, 1, 0. Тъй като в тази редица има две единици, отговорът е $2/10$ =20%.	Тази вероятност не може да бъде определена, тъй като съществуват безкрайно много естествени числа и не можем да проверим всички възможности.

Кой от трите отговора смятате за *най-подходящ*?

*Отбележете **едно** квадратче.*

Лиза.....	12,2
Моника.....	37,8*
Силвия.....	22,0

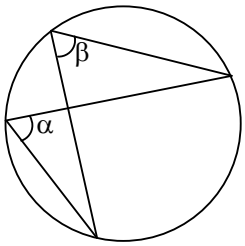
Без отговор: 28,0

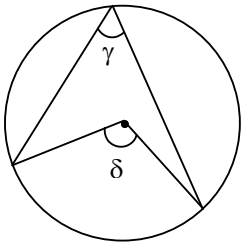
Кое от следващите предположения **най-добре** описва проблемите на другите две ученички?

*Отбележете **до 2** квадратчета.*

	Да	Без отговор
не разбира какво се случва, когато се повдига на квадрат.....	18,3*	28,0
не разбира какво е естествено число.....	32,9	28,0
не взема предвид, че последната цифра вече е определена да е 1.....	30,5	28,0
намира задачата твърде трудна и се отказва.....	18,8	28,0
не разглежда подходящите естествени числа.....	24,4	28,0
не разбира понятието вероятност.	26,8*	28,8

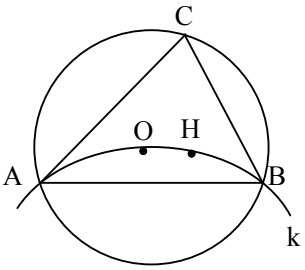
Пример 13. Обяснявате на ученици в прогимназията, че на фигурата по-долу $\alpha = \beta$ и $\delta = 2\gamma$.





Ученици трябва да решат следната задача:

Нека O и H са центърът на описаната окръжности ортоцентърът на ΔABC и нека A, B, O, H лежат на окръжността k. Намерете големината на $\sphericalangle ACB$.



Кое от следните упътвания намирате за най-подходящо?

*Отбележете **едно** квадратче.*

Докажете, че AB е диаметър на k	4,9
Докажете, че триъгълникът ABC е равнобедрен.	14,6
Изразете $\sphericalangle AOB$ и $\sphericalangle AHB$ посредством $\sphericalangle ACB$	58,5*
Опитайте се да построите такъв триъгълник ABC	6,1

Без отговор: 15,9

Задачата от пример 13 е причислена към областта *Геометрия* от скалата по математика и към *Знания за преподаване на математиката* (планиране на уроците по математика) от скалата по дидактика на математиката. Това е също добър пример на математическа задача в дидактическа ситуация. Тази ситуация може да се случи непредвидено по време на учебен час, но старателният учител обикновено планира подходящи упътвания към задачите, които има намерение да решава, по време, когато си подготвя урока. В тази задача не е задължително да се реши математическата част, за да се отговори на дидактическата. Достатъчно е отговарящият да е малко или много запознат с геометричната конструк-

ция и да има определен усет да отхвърли неправилните отговори (първите два са неправилни от математическа гледна точка). Понеже все пак има и дидактическа задача, трябва да прецени кое от останалите две упътвания е по-подходящо.

Това е една от задачите, на която българските бъдещи учители по математика са показали сравнително добри постижения. Предположенията са, че геометричната конструкция е сравнително добре позната (две от забележителните точки в триъгълника лежат на една окръжност с два от върховете му). Така, за тях остава сравнително лека дидактическа задача – да посочат кое от последните две упътвания е по-подходящо. Това, в известен смисъл, също е математическа задача, защото: (1) търсеният ъгъл ACB се среща в третото упътване; (2) в конкретния случай построяването на триъгълника ABC надали ще доведе до намиране на търсения ъгъл.

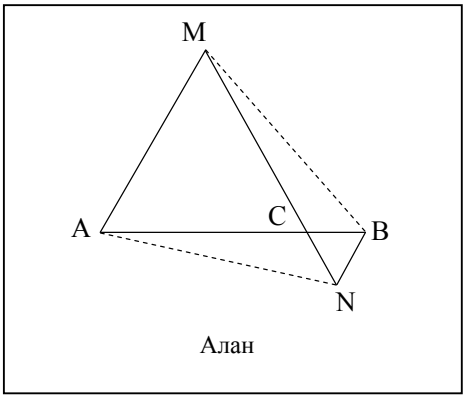
Задачата от пример 14 е класифицирана към областта *Геометрия* на математическата скала и към *Контакт „Ученик-учител”* на скалата по дидактика на математиката. Дидактическата ситуация е анализиране и оценяване на ученически решения. В случая трябва да се прецени дали даден чертеж отговаря на условието на задачата или не. Българските бъдещи учители по математика сравнително добре са се справили с това – около две-трети са отговорили правилно.

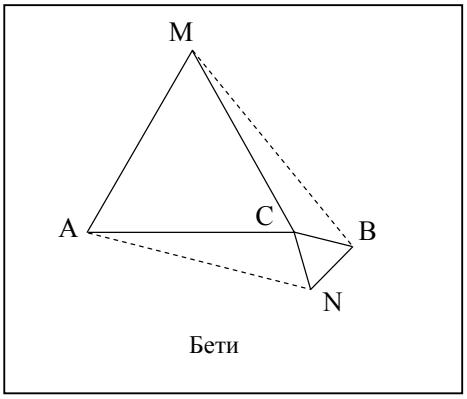
Повече математика има във втората част на задачата. Въпреки че това е сравнително известна математическа ситуация (например за кандидатстващите в профилирани училища след завършен 7 клас), резултатите на българските бъдещи учители по математика не са добри – по-малко от половината са дали правилен отговор. Забележете, че не се иска решение на задачата.

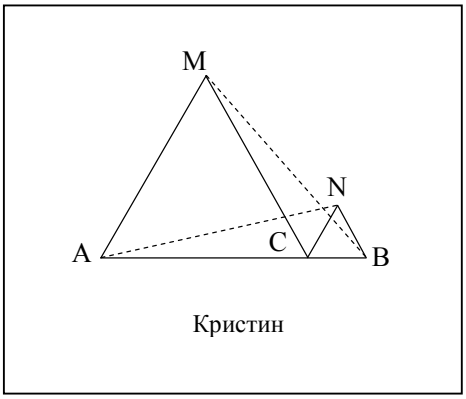
Пример 14. На учениците е дадена следната задача:

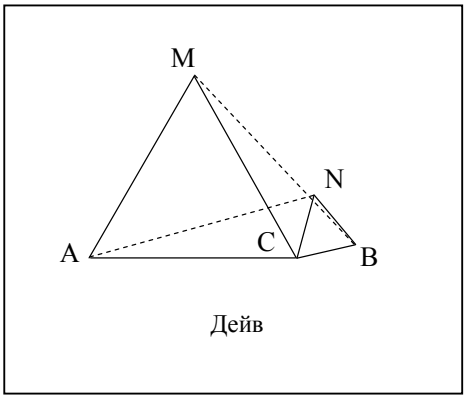
Нека *C* е точка от отсечката *AB*. Начертани са равностранните триъгълници *ACM* и *BCN*. Вярно ли е, че *AN*=*BM*?

Четирима ученика нарисували следните чертежи:









1) Чертежът на кой ученик отговаря на дадените условия?

*Отбележете **едно** квадратче.*

A) Само на Алън 17,1
B) Само на Бети 0
C) Само на Кристин 0
D) Само на Дейв 0
E) Само на Алън и Бети 0
F) Само на Алън и Кристин 64,6*
G) Само на Бети и Дейв 1,2
H) Всичките четири чертежа 8,5

Без отговор: 8,5

*Отбележете **едно** квадратче*

	Винаги	Понякога	Не съм сигурен	Без отговор
2) Вярно ли е, че *AN* = *BM*?.....	47,6*	29,3	14,6	8,5

Пример 15. Хората, изглежда, имат различни подходи към решаване на задачи с деление на дроби, като, например, задачата $1\frac{3}{4} \div \frac{1}{2} =$.

За да направят това по-разбрано за децата, много учители се опитват да го свържат със ситуации от реалния свят или със задачи с думи. За всяка една от задачите с думи преценете поотделно дали тя е адекватна или не?

*Отбележете **едно** квадратче на **всеки** ред.*

	Адекватна	Неадекватна	Без отговор
Двама души имат една цяла торта и ¾ от друга торта. Как да разделят тортите по равно на всеки?.	63,4	29,3*	7,3
Знаем, че лицето на един правоъгълник е равно на произведението от неговата дължина и широчината му. Нека лицето на една правоъгълна дъска е $1\frac{3}{4}$ кв.м., а нейната широчина е $\frac{1}{2}$ м. Каква е нейната дължина?	48,8*	40,2	11,0
Вчера карах велосипед от град А до град В. За $1\frac{3}{4}$ часа изминах $\frac{1}{2}$ от пътя си. За колко време съм изминал целия път?.....	47,6*	43,9	8,5
Имате едно цяло и три четвърти литра течност в кана. Искате да я разделите наполовина, за да може всеки от вас двамата да изпие половината. Колко литра е изпил всеки от вас?.	47,6	45,1*	7,3

Задачата от пример 15 е класифицирана само към скалата по дидактика, област *Знания за преподаване на математика* (свързване на дидактическите методи със структурата на урока). Задачата не е свързана със скалата по математика, защото основното в нея е дидактическата ситуация – подбор на подходящ метод за представяне на дадено математическо действие.

Неадекватните първа и четвърта опция са свързани с често срещаната грешка: да се раздели на 2, вместо исканото – да се раздели на $\frac{1}{2}$. Първата опция представя ситуация, в която е по-трудно да се установи това, и, както показват резултатите на българските участници, голяма част от тях са сбъркали отговора. Интересното е, че във втората, третата и четвъртата опция процентите на

отговорите правилно и неправилно са почти едни и същи. Това е сериозна причина за безпокойство. В задачата се иска само да се моделира действието деление. Вярно, че това е „най-сложното” от четирите аритметични действия, но бъдещите учители по математика би трябвало да могат да различат коя ситуация кое действие моделира.

Пример 16. На ученици е дадена следната задача в един тест:

Чифт обувки се продавал през първите две седмици за цена от 75 лв. След това цената била увеличена с 10%. Поради рязкото намаление на продажбите, новата цена била намалена с 10%. Петър си помислил: „Сега цената на обувките е същата, както в началото.” Какво мислите Вие по този въпрос? Обосновете отговора си.

Ето отговорите на двама ученика:

Ученик 1: Това е все едно като 20 + 10 = 30 и след това 30 – 10 = 20 … получаваме същото, понеже цената е намалена със същото количество, с което е била увеличена преди това.

Ученик 2: Петър е прав, понеже най-напред обувките са стрували 75 лв., след това са стрували със 7,50 лв. повече и после – със 7,50 лв. по-малко.

1. Какво е заблуждението на ученик 1?

*Отбележете **едно** квадратче*

Той смята процента (10%) за константна величина, която не зависи от цената.	77,2*
Той смята, че е все едно дали първо намаляваш цената и после я увеличаваш или обратно.	10,1
Той би трябвало да напише 75 + 10 = 85 и после 85 – 10 = 75.....	3,8

Без отговор: 8,9

2. Какво е заблуждението на ученик 2?

*Отбележете **едно** квадратче*

Той не знае, че началната цена няма да бъде променена, без значение дали първо намаляваш цената и после я увеличаваш или обратно.	8,9
Той не разбира, че промяната в цената зависи от базовата стойност, въз основа на която се изчислява процента.	74,7*
Той не знае, че можеш да прибавяш проценти като фиксирани числа.....	2,5

Без отговор: 13,9

Задачата от пример 16 е от областта *Контакт „Ученик-учител”* (диагностициране на типични грешки) от скалата по дидактика на математиката. Задачата не е класифицирана към скалата по математика, защото математическата част в нея е сравнително елементарна и добре известна. Акцентът в задачата е върху дидактическата ситуация – да се открие заблуждението на всеки ученик.

Резултатите на българските бъдещи учители по математика върху тази задача са сравнително добри. Темата „процент” се изучава подробно в училище. Това става във възраст, когато учениците не са загубили желанието си за учене (както става в горните класове на обучени). Поради това знанията са трайни и това оказва своето положително въздействие. Неправилните отговори на задачата са такива, които обикновено се дискутират в училище. Това също е от значение за намиране на правилните решения.

Пример 17. Промяна в учебната програма премества изучаването на квадратен корен от програмата по математика за средното училище в гимназиалната програма.

Кое от изброените твърдения посочва на какво учениците в средното училище все още могат да бъдат учени, вследствие на тази промяна?

*Отбележете **едно** квадратче на **всеки ред**.*

	Винаги	Понякога	Никога	Не съм сигурен	Без отговор
Те все още могат да бъдат учени как да решават линейни уравнения. ...	58,2*	13,9	1,3	8,2	17,7
Те все още могат да бъдат учени как да решават квадратни уравнения, използвайки формулата	31,6	25,3	20,3*	2,5	20,3
Те все още могат да бъдат учени как да използват стандартните тъждества за квадратен корен от изрази.....	16,5	27,8	16,5*	17,7	21,5
Те все още могат да бъдат учени как да използват Питагоровата теорема за решаване на задачи, включващи намиране на хипотенузата на правоъгълен триъгълник.	36,4	32,9*	7,6	3,8	20,3
При тях все още може да се въведе множеството на ирационалните числа.	16,5*	20,3	26,6	17,7	19,0
Те все още могат да бъдат учени как определят еднакви триъгълници.	58,2*	13,9	2,5	6,3	19,0

Задачата от пример 17 е от скалата по дидактика на математиката, област *Знания за учебните планове* (разбиране на връзките между темите в учебните планове и подбиране на подходяща тематична последователност). В нея не се поставя математически проблем и затова тя не е към скалата по математика.

Дидактическата ситуация изисква да се прецени как ще се наруши последователността на някои теми, ако „липсва" понятието за квадратен корен. Някои от предложените теми могат да се изучават без това понятие, други не могат, а трети – могат да се изучават частично.

Отговорите на българските бъдещи учители по математика са „странни". С изключение на първата и последната тема, на която почти 60% от участниците са дали правилен отговор, останалите отговори показват сериозно неразбиране на мястото на понятието „квадратен корен". Будят недоумение отговорите по третата тема, в която ясно е казано „квадратен корен от изрази". Трудно е да си представим формулата за решаване на квадратно уравнение без наличието на квадратен корен. Използване на Питагоровата теорема и въвеждане на ирационалните числа поставят по-трудна ситуация, но споменатите втора и трета тема не трябва да затрудняват бъдещите учители по математика.

Представянето на горните задачи има следните цели:

- Да покаже какви знания и умения се ценят в международен план в обучението на учители по математика. Това са знания за предмета математика, от една страна, но не теоретичните знания, необходими за така наречения „чист математик", а математически знания, представени в и свързани с дидактически ситуации.
- Да анализира накратко математическата същина и дидактическите постановки в международно признати задачи за изследване на подготовката на учители по математика.
- В по-конкретен план, да представи част от задачите в изследването MT21. Това е важно за правилното тълкуване на резултатите изследването.

В заключение може да се каже, че при търсенето на отговора на въпроса „Кои са онези математически знания, които са необходими за преподаване на математика?" може да се обърнем към задачите, с които се правят изследвания на подготовката на учители по математика. Те в известна степен отразяват постигнатото до момента разбиране на проблема.

ПЕТНАДЕСЕТА ГЛАВА
НЯКОИ РЕЗУЛТАТИ ОТ MT21, КАСАЕЩИ ПОДГОТОВКАТА НА УЧИТЕЛИ ПО МАТЕМАТИКАТА В БЪЛГАРИЯ

В тази глава се разглеждат някои от резултатите от международното изследване на обучението на учители по математика MT21 (Банков, 2010). Кратко описание на това изследване има във втора глава. Да напомним само, че в изследването участват шест държави: България, Германия, Корея, Мексико, САЩ и Тайван.

Най-напред ще разгледаме някои международни резултати (Банков, 2008), т.е. ще направим сравнение на подготовката на учители по математика в България с тази в другите пет държави, участващи в изследването. Тези резултати (Schmidt et al., 2007) бяха обявени на света на официална пресконференция във Вашингтон на 11 декември 2007. В словото си пред журналистите ръководителят на MT21 (William Schmidt) обобщи резултатите в математическата част на теста така: „Постиженията на бъдещите учители от Тайван и Корея по всички пет математически области (Алгебра, Функции, Геометрия, Числа, Данни и вероятности) на TM21 са най-високи. Постиженията на САЩ са по-ниски от тези на Германия и, в някои от областите, по-ниски от тези на България.” Така по отношение на математическите знания на бъдещите учители по математика България беше поставена след Германия и сравнима със САЩ. Този факт е в пълно съответствие с действителните резултати от MT21.

След това ще разгледаме някои специфични за България резултати: представянето по отделните скали, разликата в постиженията на студенти от първи и четвърти курс, разлика между участващите институти и др.

В края на тази глава се правят някои изводи, касаещи подготовката на учители по математика в България. Продължава да се търси отговор на въпроса какви математически знания са необходими, за да се преподава математика в училище.

1. Международни резултати и мястото на България

Основното събиране на данни за MT21 стана през март-април 2006 г. Таблица 36 представя големината на извадката от институти и бъдещи учители по математика за всяка от участващите държави.

Таблица 36. Големина на извадката в MT21.

Държава	Брой институти	Брой бъдещи учители
България	3	161
Германия	Първи – 4*; Втори – 2*	848
Корея	4	210
Мексико	6	358
САЩ	12	382
Тайван	5	668
Общо	**34**	**2627**

* В Германия обучението на учители става в две институции. Най-напред те се обучават по съответния предмет (бакалавърска степен по математика) в университет („първи" институт) и след това постъпват в („втори") институт за обучение на учители (обикновено 2 години).

Извадката за изследването е целева (purposeful). Тя е правена на ниво „институт". Във всеки „институт" са обхванати всички бъдещи учители по математика, които са в първата и последната си година на обучение. Всички държави са избрали в извадката „основни/важни институти", в които се обучават бъдещи учители по математика. Може би „най-слабо представителна" е извадката на САЩ. В тази държава, обаче има около 1300 институти, които подготвят учители по математика и за представителност може да се говори, ако се изберат много повече институти. Въпреки това, тези 12 института в САЩ са избрани по такъв начин, че да обхванат най-важните програми и да дадат престава за основните различия в подготовката на учители по математика.

Както е известно, в България подготовката на учители по математика за средното училище се извършва от факултетите по математика в някои университетите. В основното изследване на България за MT21 участваха 3 такива университета с общо 78 студенти от първи курс и 83 студенти от четвърти курс. В момента на изследването в тези три университета се подготвяха общо над 50% от завършващите специалността учител по математика (и информатика) за средното училище в страната. Броят на изследваните завършващи студенти (83) беше около 40% от завършващите за страната.

В MT21 бяха използвани следните въпросници: (1) *Въпросник за института*, в който се събираха данни за условията на обучение на бъдещите учители, като основно внимание се отделя на програмите за обучени и на четените курсовете; (2) *Въпросник за преподавателите*, в който се събираше информация за спецификата на курсовете, които те четат, както и за отношението на преподавателите към математиката и нейното преподаване; (3) *Въпросник за бъдещия учител по математика*, в който освен проверка на знанията, съдържаше още

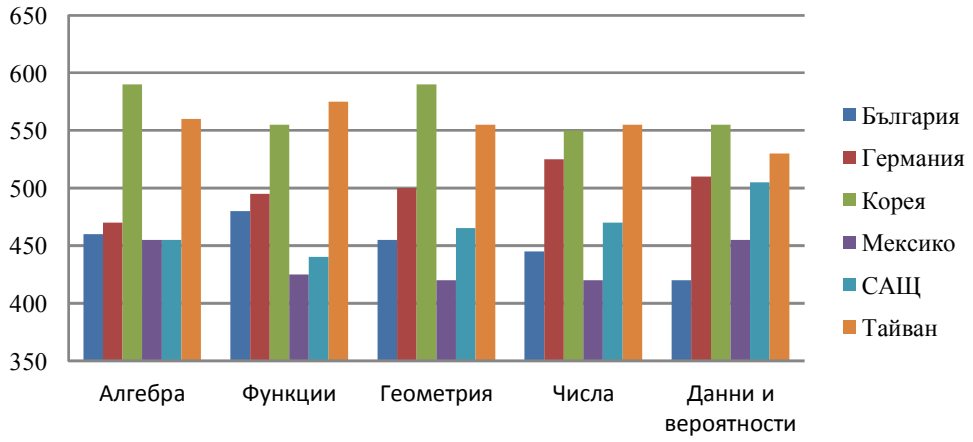
общи (демографски) въпроси, въпроси, свързани с възможностите за обучени и отношението на студентите към математиката и нейното преподаване.

Постижения по математика и дидактика на математиката. В областта на математиката са построени пет скали, а именно Алгебра, Функции, Геометрия, Числа, Данни и вероятности. В теста за постижения към всяка от тях има съответно по 36, 14, 11, 19, 12 едноточкови „задачи”. Кавичките са поставени, защото някои от задачите на MT21 се състоят от подусловия и се точкуват с повече от една точка. Примерни задачи могат да се видят в предишната глава. В областта на дидактика на математиката са построени три скали, а именно Знания за учебните планове, Знания за преподаване на математиката, Контакт „ученик-учител”. По-подробно описание на съдържанието им е дадено също в предишната глава. В теста за постижения за всяка от тях има съответно 18, 18, 23 „задачи”.

Използвани са две тестови книжки за бъдещите учители с непълен свързан тестов дизайн. Всеки бъдещ учител е работил само върху една от тях. След това е приложен едно-параметричния вероятностен модел (модел на Раш). Всяка от осемте скали е оразмерена така, че има средна стойност 500 и стандартно отклонение 100.

Диаграмите на фигура 84 показват разпределението на средните стойности на постиженията на участващите държави по петте математически скали.

Фиг. 84. Средни постижения по математическите скали



Разликите в средните постижения на държавите са статистически значими ($p < 0,0001$). Очевидно е, че резултатите на Тайван и Корея са от една-четвърт до едно стандартно отклонение над останалите държави. Германските

резултати са около средните. Мексико е доста под средното по всички скали. България и САЩ варират по постижения в различните скали.

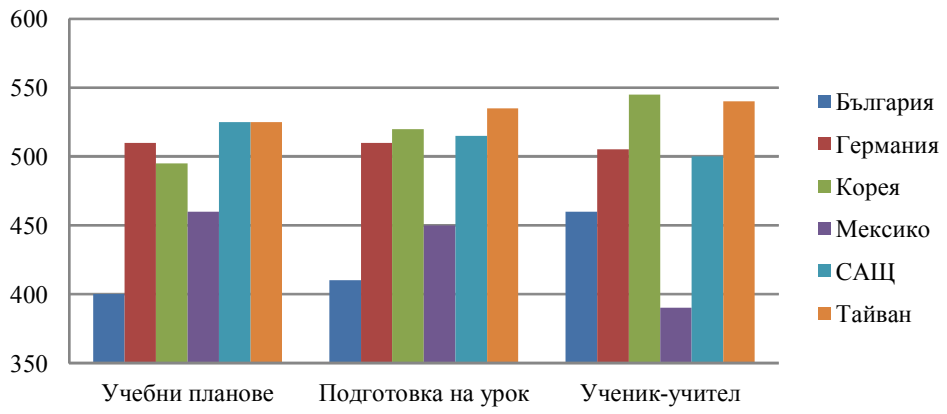
Изучаването на учебните планове на международно ниво, направено в рамките на TIMSS (Schmidt et al., 1997), показва, че основно в прогимназията се изучава алгебра и геометрия. Обикновено по това време в училище приключва придобиването на предимно изчислителни умения в аритметиката и се преминава към по-абстрактно изучаване на математиката. Като се има пред вид възловото място на алгебрата и геометрията в прогимназията, особено важно е бъдещите учители по математика да имат солидни постижения в тези области. Онова, което виждаме за бъдещите учители по математика в България в тези две области, не е радостно – постиженията са около половин стандартно отклонение под средното и с над едно стандартно отклонение под Тайван и Корея. Обърнете внимание на скалата по Геометрия. През последното десетилетие у нас много се говори, че България е едно от малкото места в света, където геометрията в средното училище все още не е пренебрегната. Според официалната документация това може и да е така, но според знания, то със сигурност не е. Резултатите от MT21 показват, че дори в САЩ бъдещите учители по математика са се представили по-добре по геометрия от Българските. А разпространено мнение е, че в средното училище на САЩ геометрия не се изучава. Трябва да се обърне внимание и на факта, че геометричните задачи в MT21 са близки до задачите, които се решават в българското училище. Някои примери на такива задачи могат да се видят в предишната глава.

Безпокойство будят и изключително слабите резултати на нашите бъдещи учители по математика в скалата Числа. Въпреки че това не е основна тема в прогимназиалните и в гимназиалните класове, където се очаква да преподават тестираните бъдещи учители, знанията за числата са основа на математическите знания. А без здрава основа трудно може да се очаква сериозно преподаване. Слабите резултати по скалата Данни и вероятности е поредното доказателство, че в българското образование по математика това е слаба (дори пренебрегвана) област – факт, за който доста се говори в последните години, но нищо сериозно не е направено. Единствената математическа скала, в която българските бъдещи учители са показали сравнително по-добри резултати (макар и малко под средните), е тази по Функции. Вероятно солидното обучение по ма-

тематически анализ в началните курсове на университетите е оказало своето влияние в тази област.

Диаграмата на фигура 85 показва разпределението на средните стойности на постиженията на участващите държави по трите скали по дидактика на математиката.

Фиг. 85. Средни постижения по скалите по дидактика на математиката



Както по математика, така и тук разликите в средните постижения на държавите са статистически значими ($p < 0,0001$). Но те са по-малки, отколкото в математическите скали. Подреждането на държавите малко се различава от онова, което наблюдавахме при математическите скали. Например, в скалата Знания за учебните планове най-добре са се представили Тайван и САЩ. В останалите две скали по дидактика на математиката подреждането е почти както в математическите скали. С тази разлика, че САЩ е на доста по-добра позиция.

Българското представяне в областта дидактика на математиката е много слабо. България е на последно място в скалите Знания за учебните планове и Знания за преподаване на математиката и на предпоследно място по Контакт „ученик-учител”. Разликите с другите държави са големи. Очевидно е, че курсовете по дидактика (методика) на математиката, които се четат в българските факултети за подготовка на учители трябва да се преосмислят и осъвременят.

Съдържание на прослушаните курсове. Едно от обясненията за разликите в постиженията на бъдещите учители по математика е в различните подготовки, които те получават. В частност, предполага се, че съществено значение за

знанията им оказват курсовете, които те слушат в програмата за подготовка на учители, както и времето, което се отделя за тях.

В основата на изследването MT21 е разбирането, че знанията на бъдещите учители по математика трябва да се изграждат в три основни направления: *математика, дидактика на математиката и обща педагогика* (Shulman, 1987). Всяко от тези направления, макар в различна степен и тежест, е застъпено в обучението на учители от всички държави, участващи в MT21 и в TEDS-M. Таблица 37 показва средния брой *астрономически* часа, в които се изучават съответните направления в отделните държави. Данните са от учебната година 2005-2006, когато е проведено изследването MT221.

Таблица 37. Среден брой астрономически часа, в които се изучават съответните направления в обучението на учители.

	Математика	Дидактика на математиката	Обща педагогика	Втора специализация	Друго	Общо часове
България	1037	77	98	483	116	1811
Германия	588	234	516	822		2160
Корея	741	53	210		722	1726
Мексико	896*		2688			3584
САЩ	518	114	505	139	471	1747
Тайван	1286	165	225		634	2310

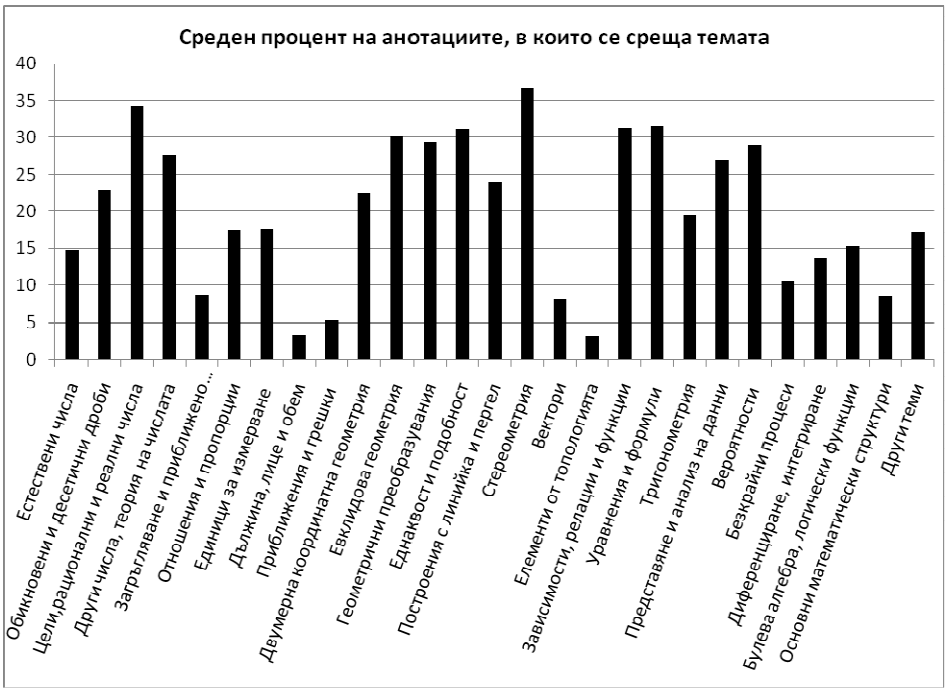
* Тук са включени и часовете по дидактика на математиката.

Две неща правят впечатление за България от тази таблица. Първо, сравнително големият брой часове, предвиден за изучаване на математика (на второ място след Тайван, която е сред държавите с най-високи постижения по математика в TIMSS). Второ, изключително малкият брой часове по обща педагогика. Последното наблюдение се отнася в известна степен и за часовете по дидактика на математиката.

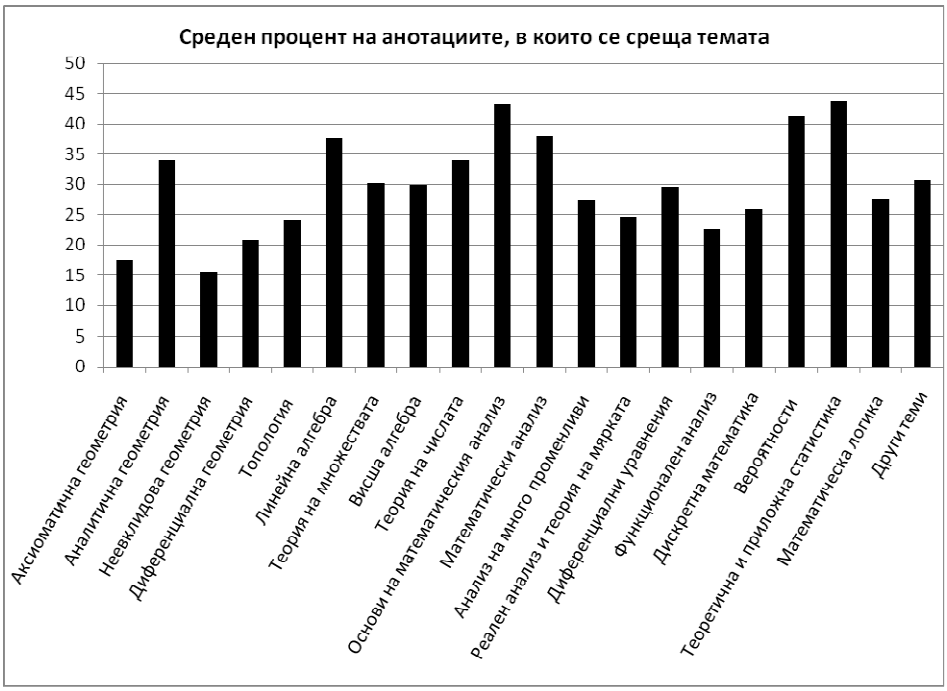
Допълнение към тази картина дава едно задълбочено изследване на учебните планове и преподавателските анотации на курсовете за подготовка на учители по математика, направено от Maria-Teresa Tatto и Кирил Банков (не е публикувано до момента) в рамките на изследването TEDS-M. В него се предполага, че е важно да се получи информация за обучението на бъдещите учители по така наречената „училищна математика”, т.е. математиката, която те би трябвало да преподават в училище. Затова, резултатите са систематизирани в четири направления – училищна математика, математика (включва се матема-

тика, която обикновено е над нивото на тази, изучавана в средното училище, и е типична за университетите), дидактика на математиката и обща педагогика.

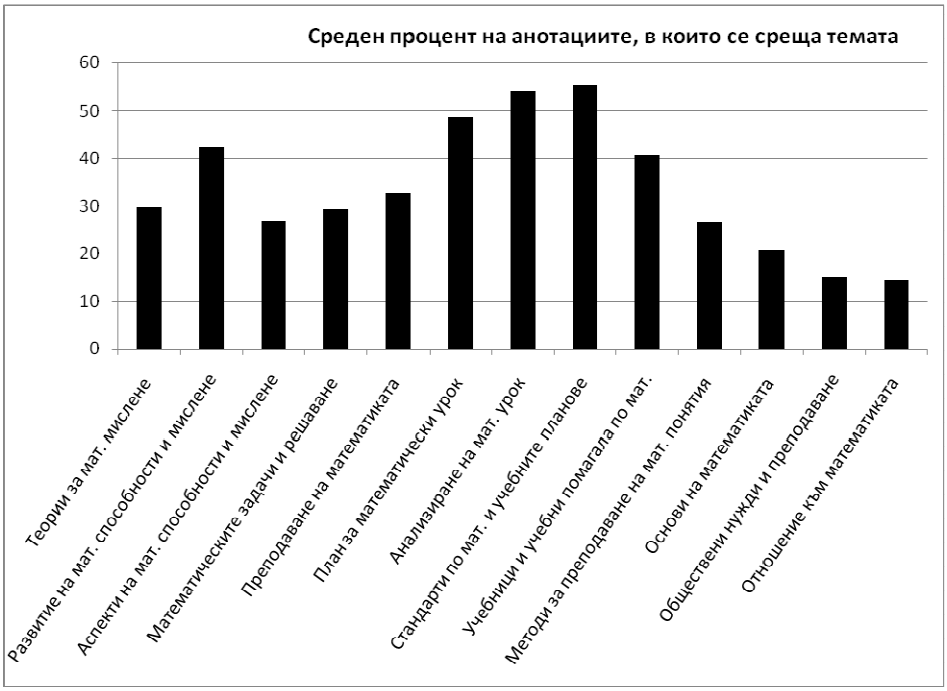
Диаграмите на фигури 86 до 89 представят средния процент на преподавателските анотации, в които се среща съответната тематика. Данните са от 8 държави (общо 264 изследвани анотации), участвали в TEDS-M.



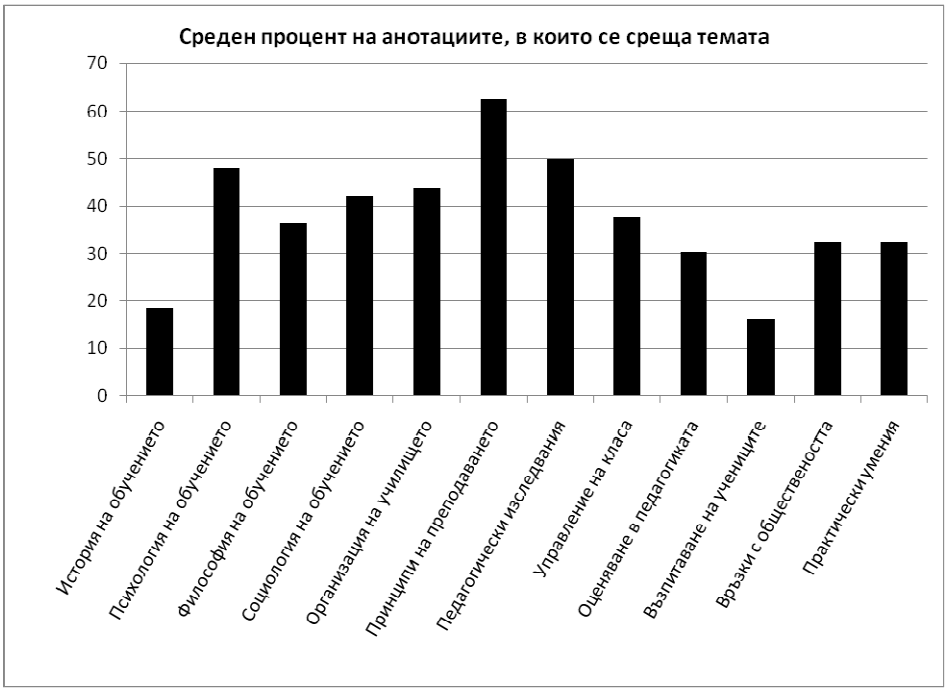
Фигура 86. Училищна математика



Фигура 87. Математика



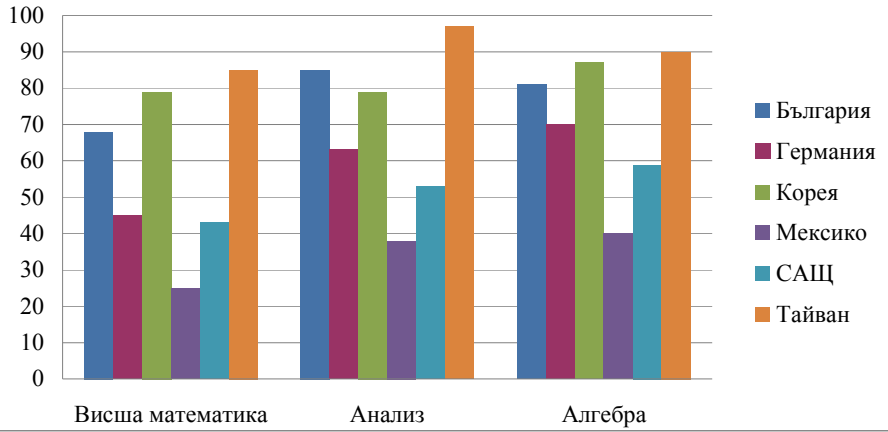
Фигура 88. Дидактика на математиката



Фигура 89. Обща педагогика

В рамките на изследването MT21 студентите трябваше да дадат информация за тематиката по математика, която те са изучавали в прослушаните курсове. Във въпросника бяха написани теми, които обикновено се включват в университетското обучение по математика. Бъдещите учители трябваше да отбележат кои от тях са изучавали, независимо дали като отделен курс или като част от курс.

Фиг. 90. Среден процент изучавана тематика



Диаграмата на фиг. 90 показва средния процент на темите по математика изучавани в отделните държави. Висша математика включва теми от университетските курсове, които обикновено се преподават на студентите-математици, като например абстрактна алгебра, диференциално и интегрално смятане (на една и много променливи), комплексен анализ, функционален анализ, диференциална геометрия, топология и др. Отделно е даден процентът на изучаваната тематика по анализ (диференциално и интегрално смятане, диференциални уравнения, комплексен анализ и др.), както и алгебра (линейна алгебра, теория на групите, полетата, пръстените, висша алгебра и др.). Причината е предположението, че анализ и алгебра са тясно свързани с преподаването на математика в средното училище.

Разликите в средния процент на изучаваната тематика между отделните държави са статистически значими ($p < 0,0001$). Както може да се очаква, бъдещите учители в Тайван и Корея изучават голям процент (между 79% и 86%) от тематиката по висша математика. Особено впечатляващо е, че в Тайван те изучават практически всичко (средно 96%) от тематиката по анализ. На трето място по среден процент изучавани теми по математика се нарежда България. В това

няма нищо странно, защото нашите бъдещи учители по математика се подготвят във факултетите по математика и информатика към съответните университети и те задължително изучават голяма част от математическите курсове, предназначени за студентите по математика. Трябва да отбележим, обаче, че разликата между България и двете азиатски държави в това отношение е значителна.

В процеса на разработване на въпросниците се оказа, че подобен метод за получаване на информация за курсовете, свързани с преподаването в училище, не може да се приложи. Причината е, че поради голямото разнообразие в обучението в тези области не е възможно да се напише списък с теми по „училищна математика", които „обикновено се преподават в университетите". Затова бъдещите учители бяха помолени да класифицират възможностите, които те са имали да бъдат въвлечени в теоретични или практически занятия по определени дейности. Използвана е шест степенна скала: от „никак" (0) до „в голяма степен" (5).

Дейностите са групирани в три категории: Училищна математика (математика, която учителите се предполага, че ще трябва да преподават), Математическо образование (визират се не толкова теоретични основи, а повече практически дейности, свързани с обучението, като например: „преподаване на учениците да използват алгоритми", „обучение на учениците да правят математически доказателства", „представяне на математически теми по подходящ за ученическата възраст начин" и др.) и Обща педагогика (избрани са не теоретични основи, а на практически аспекти, свързани с обучението на ученици).

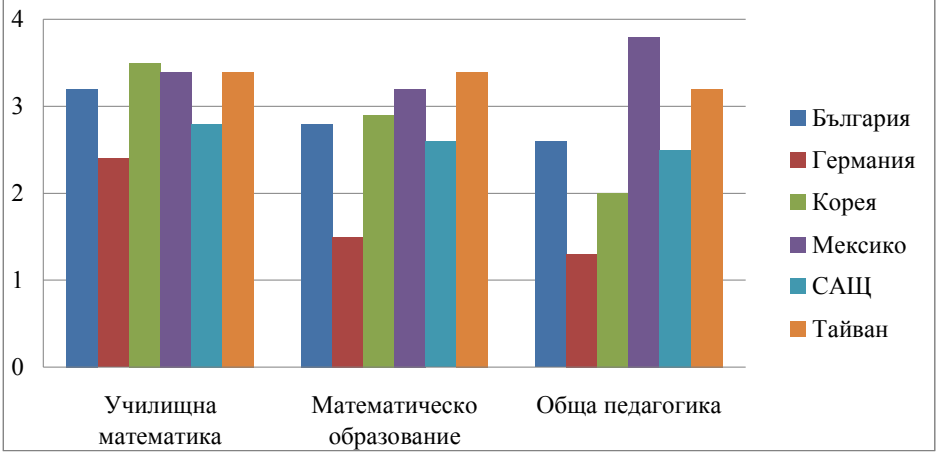
Диаграмата на фиг. 91 показва средния процент на класификация по въпросите от трите области за всяка държава.

Съществена разлика от предишните диаграми е, че в тази област бъдещите учители на Мексико имат значително по-добри възможности за обучение. Може да се каже, че обучението на бъдещи учители по математика в Мексико е изключително насочено към практически знания за преподаването в училище.

В Училищна математика с почти равни възможности за обучение са Корея, Тайван, Мексико и България. По-слабо обучение тук имат САЩ и Германия. В практическите аспекти на математическото образование доминират Тайван и Мексико, следвани от България, САЩ и Корея. Тази област е най-слабо застъпена в Германия. В областта на Обща педагогика от шестте държави най-

големи възможности за обучение предлага Мексико, следвано от Тайван. После са България и САЩ. Отново в Германия тази област е най-слабо застъпена.

Фиг. 91. Класификация на обучението в преподаване

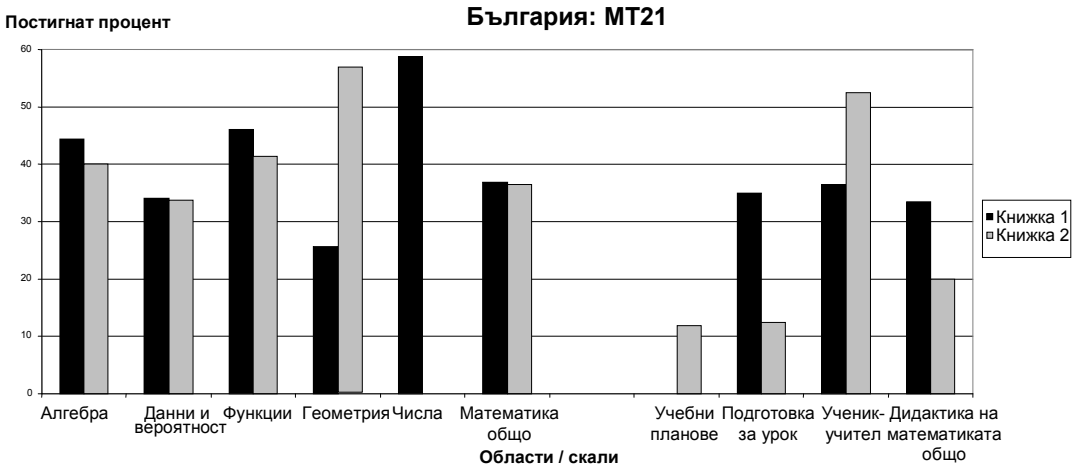


Онова, което прави впечатление, е, че в двете азиатски държави Тайван и Корея бъдещите учители по математика получават много сериозна подготовка, както по математика, така и по дидактика на математиката, а също и по знания и умения за преподаване в училище. Това обяснява (донякъде) високите резултати на тези държави в теста за постижения.

2. Някои резултати за представянето на България

Най-напред ще направим сравнение на представянето на българските учители по математика по двете тестови книжки. Преди това, обаче е редно да се даде едно пояснение. Резултатите от предния параграф са представени в единна скала за всички държави, защото само така може да се прави сравнение между тях. Диаграмите, които следват по-долу, показват средния процент правилни отговори за отделните скали за България. Това означава, че няма единна скала нито за сравнение с другите държави, нито за отделните математически области.

Диаграмата от фиг. 92 показва българското представяне по отделните книжки и по всяка от под-скалите в тях.



Фиг. 92. Среден процент за всяка от двете книжки на МТ21

Това, което прави впечатление от диаграмата е, че в книжка 1 няма под-скала Числа. Причината е, че в тази книжка няма задача от тази подобласт. По същата причина в книжка 2 няма под-скала Знания за учебните планове. Ясно се вижда, че представянето на българските бъдещи учители по математика в това изследване не е добро. Най-високият среден процент на постиженията по отделните области е едва под 60%.

Според замисъла на МТ21, всяка от участващите държави трябваше да избере по около равен брой студенти, от началната и последната година на обучението си за учители по математика. Така, за България, например, извадката се състои от 78 студенти от 1-ви курс и 83 студенти от 4-ти курс.

Читателят вероятно предполага, че сравнително слабите резултати за България, се дължат на факта, че данните обединяват студенти от първи и четвърти курс и може би нещата ще изглеждат по-добри, ако резултатите се разделят по курсове. За съжаление това не е така. Таблица 38 показва резултатите в скалата по математика, разделени по курсове.

Таблица 38. Средна стойност на процент верни отговори по математика

	Книжка 1	Книжка 2
1-ви курс	41,73	41,58
4-ти курс	39,95	38,38

Същият вид резултати за скалата по дидактика на математиката са представени в таблица 39.

Таблица 39. Средна стойност на процент верни отговори по дидактика на математиката

	Книжка 1	Книжка 2
1-ви курс	49,58	48,78
4-ти курс	59,66	66,79

Оказва се, че и в двете скали няма статистически значима разлика между резултатите на първокурсниците и тези на четвъртокурсниците. Това е причината при разглеждане на резултатите по задачи, което беше направено в предишната глава, те да се представят заедно за всички участници – разделянето по курсове не води до съществено различни резултати при първокурсниците и при четвъртокурсниците.

Както вече беше казано, в България изследването MT21 се проведе в три институти (математически факултета), които подготвят учители по математика, които по-долу са представени с кодове[1] съответно 1, 2 и 3. Оказа се, че в резултатите на тези институти има известна разлика. Таблица 40 показва резултатите на трите института по процент верни отговори в скалата по математика по двете книжки.

Таблица 40. Средна стойност на процент верни отговори по институти – математика

	Книжка 1	Книжка 2
Институт 1	53,61	53,45
Институт 2	32,97	31,91
Институт 3	28,92	34,49

Направеният t-тест за статистическа значимост показва, че има статистически значима разлика (p<0,05) между първия институт (най-добри резултати) и всеки от другите два, а няма такава между втория и третия институт.

Таблица 41 показва подобен резултат в скалата по дидактика на математиката.

[1] Изследването MT21 се проведе, при условие че имената на тези институции няма да бъдат съобщавани, освен с изрично тяхно съгласие.

Таблица 41. Средна стойност на процент верни отговори по институти – дидактика на математиката

	Книжка 1	Книжка 2
Институт 1	46,11	58,88
Институт 2	30,01	32,04
Институт 3	28,96	46,60

При дидактика на математиката има статистически значима разлика (p<0,05) между втория институт (най-слаби резултати) и всеки от останалите два, а няма такава между първия и третия институт.

Единственият институт, в който има статистически значима разлика между резултатите на първокурсниците и четвъртокурсниците, е първият. Тази разлика се наблюдава както по математика, така и по дидактика на математиката и е по-голяма във втората тестова книжка.

3. Някои изводи от представените резултати

Изводите са направени главно в сравнителен аспект, за почвайки от мястото на България сред участващите шест държави в MT21.

Обучението на учители по математика в шестте участващи държави съдържа трите основни направления: математика, дидактика на математиката и педагогика. Тежестта на всяко от тях, обаче е различна. В Тайван, например, и трите направления са силно застъпени. България и Корея имат сравнително засилено изучаване на математика, но не чак толкова се набляга на дидактика на математиката и педагогика. Мексико е пример за държава с особено силно обучение по педагогика и дидактика на математиката, но много слабо застъпена математика. В Германия средно по сила е обучението по математика и слабо по дидактика и педагогика, докато в САЩ обучението по математика е слабо, а по дидактика и педагогика – средно по сила.

Заслужава внимание едно наблюдение, което е особено важно за България. Фактите показват, че нашите бъдещи учители по математика имат засилено обучение по математически дисциплини. Въпреки това, постиженията им по математическите скали на теста не са добри. В сравнение с Германия, например,

българските бъдещи учители по математика изучават повече математика по време на обучението си (около 68% от изучаваните теми за България спрямо около 45% за Германия). Но резултатите на Германия по математическите скали са по-добри (особено в областите Числа и Данни и вероятности) от тези на България.

Този факт повдига въпроса дали знанията по математика на учителите влияят върху това те да бъдат „добри учители". Понятието „добър учител" е нещо комплексно и до сега в литературата никой не е успял да му даде определение или да го опише подробно. Предположението, че за учителите по математика важи максимата „Колкото повече математика изучават, толкова по-добри учители стават" не е вярно (Begle, 1979). Много изследователи са търсили обяснение на този факт. Най-достоверното от тях изглежда е, че от значение е не само колко математика се изучава, но и как тя се преподава на бъдещите учители. С други думи, важно е бъдещите учители по математика да изучават математика, но нейното съдържание и начина на поднасяне трябва да е съобразен с особеностите на учителската професия. Образно казано, както има „математика за инженери" или „математика за икономисти", така може да има и „математика за учители по математика". Ръководно начало при изучаване на математиката е да се набляга на дидактическите принципи и идеи, да се прави връзка със знанията, които трябва да се поднесат на учениците в училище, с начина на преподаване, как съответните математически знания и идеи могат да се сведат по подходящ начин на ученици в определена възраст, как самите ученици възприемат математическите идеи, и т.н.

Подготовката на българските учители по математика става във факултетите по математика и информатика в университетите. Това има сериозни предимства, защото математика на тях се преподава от професионални математици в количество, сравнимо с онова, което учат студентите в математическите специалности. Тази математика, обаче е високо теоретична и не е свързана с дейността на учителя. Може би математическите курсове за учители би трябвало да са по-тясно свързани с особеностите на учителската професия.

Заслужава да се кажат и няколко думи за подготовката на бъдещите учители по математика в областта на дидактика на математиката. Като се разгледат примерите от тази област, представени в предишната глава, става ясно, че те поставят математически въпрос в подходяща дидактическа ситуация. Това е се-

риозен принос на изследванията MT21 и TEDS-M. Дидактическите знания са важни за учителя. Но, както в случая с математическите знания, и тук става дума не само за теоретични знания. В обучението на учители трябва да залегне нещо много по-близо до учителската практика – обсъждане на математически проблеми в подходящи дидактически ситуации. Така бъдещите учители по математика могат да получат знания и опит за адекватна реакция в реална учебна обстановка.

Като обобщение на казаното, може да заключим, че подготовката на учителите по математика има значение за знанията на учениците. Двете азиатски държави Тайван и Корея имат изключително сериозна подготовка на учителите по математика и постиженията на техните ученици на международни изследвания по математика като TIMSS и PISA-2003 например, са измежду най-високите в света. България и САЩ са по средата на класирането на държавите в TIMSS-2003, както и Германия в PISA-2003.

Редно е да се коментира и изключително обезпокоителният факт, че като цяло за България няма значим разлика в постиженията на бъдещите учители по математика в началото и в края на тяхното обучение. Такава разлика (в полза на резултатите на студентите от четвърти курс) се наблюдава само в един от трите участващи български института.

Представените резултати повдигат още много въпроси, като например: „Какво научават студентите – бъдещи учители в университетите?”, „Ако въобще научават нещо, какво е то, щом като не допринася за професионалната им подготовка?”, „Има ли нужда от обучението, което там им се предлага?”, „Как да бъдат обучавани те, така че наистина да има полза от обучението им за учители?”, и др.

ЗАКЛЮЧЕНИЕ

Основата за участието и провеждането на широко мащабни (международни и национални) ОДИ е разбирането, че изводите, базирани на тях, дават много добра възможност да се отчетат и разберат причините за наблюдаваните различия, условията и факторите, при които се провежда учебния процес в отделните страни/региони, както и връзката между различни аспекти на обучението. Получените по този начин данни могат да се използват многопосочно, но най-вече в областта на развитието на националната образователна политика, както и при предприемане на определени изменения в образователната система като цяло или в нейни отделни елементи. Международните изследвания имат неоценима полза за повишаване на методологичното равнище на изследванията във всяка отделна държава, за подобряване на диалога между представителите на държавната администрация, учените и учителите, а така също и между родителите и училищните власти. На всички заинтересовани лица се предоставя надеждна и обективна информация както за равнището на постиженията на учениците във всяка отделна държава, така и в сравнителен план между различни ученици в различни страни, обучавани по сравними учебни планове и програми.

Тази информация, поднесена по подходящ начин, подлежи на анализи. Понякога обясненията на фактите не са очевидни и е нужно време, търпение и настойчивост, за да се „сглоби картината" на цялата популация. Ще подкрепя това с един пример.

Постиженията на българските ученици в TIMSS показват стабилна тенденция към понижаване. Това беше наблюдавано в периода от TIMSS-1995 до TIMSS-1999 и после пак от TIMSS-1999 до TIMSS-2003. Разбира се, търсеха се обяснения за този факт, които са по-дълбоки от емоционалното „това е следствие от разпада в българското образование". В резултат на съвместната работа на Kelvin Gregory и на автора на този труд (Kelvin, G., K. Bankov, 2006) бяха установени някои фактори, които обясняват спада на постиженията на българските ученици по математика в периода от TIMSS-1995 до TIMSS-1999. Ето някои от тях.

Да обърнем внимание, че, както беше коментирано в четвърта глава, избраният от TIMSS вероятностен модел не е задоволителен за анализа на българс-

ките данни. Това означава, че всяка интерпретация на данните трябва да се прави със засилено внимание.

Ако внимателно се погледнат извадките от TIMSS-1995 и TIMSS-1999, прави впечатление, че те не са „еквивалентни", както по структура, така и по възраст (въпреки че и в двете години се сравняват ученици от 8 клас). Освен това, през 1995 година се наблюдава доста висок процент на неучастващи училища и неотговорили ученици, което може да се счита като фактор за изкуствено завишаване на резултатите. С други думи, в сравняването на постиженията на българските ученици за TIMSS-1995 и TIMSS-1999 има доста *голяма извадкова грешка*.

В цитираното изследване на Gregory и Банков се прави анализ на постиженията по математика на учениците от профилирани и не профилирани паралелки по всяка една от осемте тестови книжки на TIMSS-1995 и TIMSS-1999. Установено е, че: 1) Учениците от не профилираните паралелки имат малко по-слаби постижения през 1999 година по две от тестовите книжки и статистически неразличими постижения от тези през 1995 година по останалите шест книжки. 2) Учениците от профилираните паралелки имат статистически значимо по-ниски постижения през 1999 по всички тестови книжки. Всъщност, учениците от профилираните паралелки имат много по-добри постижения от тези в не профилираните и през двете изследвания TIMSS-1995 и TIMSS-1999, но през 1999 година разликата между постиженията на учениците от двата вида паралелки вече не е така голяма както през 1995 година.

В направените анализи на учебните планове и представянето на българските ученици по отделните области на математическите задачи от TIMSS, Gregory и Банков установяват, че спадът на ученическите постижения от 1995 до 1999 година може отчасти да се обясни и с различията в прилагането на учебните планове в различните училища. Освен това, българските ученици изучават голяма част от математическата тематика на TIMSS преди 8 клас (класът, в който се събират данните за TIMSS). Освен фактора „забравяне", тук играе роля и това, че много от учениците от профилираните паралелки са изучавали материала в друго училище. След постъпването си в профилираната паралелка, те започват да изучават математика на по-абстрактно ниво и не се „връщат" към математиката, която обичайно присъства в тестовете на TIMSS.

Фройдентал (Freudental, H., 1995) казва, че постижението на една държава на международно тестово изследване е резултат от съответствието между тестовия инструментариум и математическите планове в съответната държава. В случая на TIMSS, няма добро съответствие между тези два елемента за България. Въпреки че голяма част от материала, който влиза в математическия тест на TIMSS, е изучаван от българските ученици, това се е случило от 1 до 3 години по-рано от момента на провеждането на теста. С други думи в измерването има и *голяма измервателна грешка*.

Казаното само показва дълбочината на проблема с анализа и тълкуването на резултатите от широко мащабни ОДИ. Това обаче е още един аргумент в полза на богата и разнообразна информация, която дават тези оценявания.

ЛИТЕРАТУРА

Банков, К. (1997). Възрастови групи и извадки. В *Иновации в образованието и науката*. Научноизследователски институт по образование, 6, стр. 7-10.

Банков, К. (1998-А). Конструиране на извадки за международни и национални педагогически изследвания. *Математика и математическо образование - двадесет и седма пролетна конференция на СМБ*, 1998.

Банков, К. (1998-Б). Национална проверка по математика за ученици от IV клас – методи, технология и резултати. *Математика и информатика*, 3.

Банков, К. (1998-В). Резултати от националната проверка по математика за ученици от 8. клас. *Математика и информатика*, 5.

Банков, К. (1999). Подготовка на изпити на тестова основа за прием на ученици след завършен седми клас. *Математика и математическо образование – двадесет и осма пролетна конференция на СМБ*.

Банков, К. (2002-А). Качествена интерпретация на резултатите от тестове за постижения. *Математика и математическо образование – тридесет и първа пролетна конференция на СМБ*.

Банков, К. (2002-Б) Вероятностно моделиране за измерване на ученическите постижения. *Математика и информатика*, 4.

Банков, К. (2003.) Измерване на училищната ефективност по математика. *Математика и математическо образование – тридесет и втора пролетна конференция на СМБ*.

Банков, К. (2006) Подготвя ли българското училище математически грамотни хора? (Един урок от TIMSS-1993). *Математика и информатика*, 2.

Банков, К (2007-А). Спад на ученическите постижения по математика и природни науки в България. *Образование*, бр. 1.

Банков, К. (2007-Б). Ефективна ли е подготовката на учители по математика в България? *Математика и информатика,* 6.

Банков, К. (2008). Сравнително изследване на подготовката на учители по математика в шест държави. *Математика и информатика,* 1.

Банков, К. (2010). Проблеми в обучението на учителите по математика. *Математика и математическо образование – тридесет и девета пролетна конференция на СМБ*,

Банков, К., Витанов, Т. (1999). Едно изследване на постиженията на учениците по математика в 7. клас. *Математика и информатика*, 3.

Банков, К., Витанов, Т. (2000). Какво оценяваме в обучението по математика? *Математика и математическо образование – двадесет и девета пролетна конференция на СМБ*.

Банков, К., Витанов, Т. (2001). Някои резултати от националния изпит-тест – 2000. *Математика и математическо образование – тридесета пролетна конференция на СМБ*.

Банков, К., Витанов, Т. (2010). Външното оценяване по математика, състояние и перспективи. Дискусия на 39-тата пролетна конференция на СМБ. *Математика и математическо образование – тридесет и девета пролетна конференция на СМБ*.

Георгиева, Н., Банков, К. (1997). Измерване на постиженията на учениците по математика в края на IV и VIII клас. *Стратегии на образователната и научната политика*, 4.

Георгиева, Н., Банков, К., Тодорова, П. (1998). Национална проверка по математика в IV клас – технологичен модел и резултати. *Начално образование*, 6-7, 1998.

Стоименова, Е. (2000). *Измерителни качества на тестове*. НБУ, С.

Adams, R.J., & Wu, M.L. (Eds.) (2002) *PISA 2000 technical report.* Paris: OECD Publications.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics* (pp. 83-104). Westport, CT: Ablex.

Ball, D., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teacher's mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th Ed.). Washington, DC: American Educational Research Association.

Bankov, K., Mikova, D., Smith, T. (2006). Assessing Between-School Variation in Educational Resources and Mathematics and Science Achievement in Bulgaria, "*Prospects*", vol. XXXVI, No. 4, December 2006, UNESCO, Springer Netherlands. Online publication: March 31, 2007, http://www.springerlink.com/content/6057m07163123p4h/.

Bankov, K., O'Sullivan, B. Manual for the use and reporting of national and international assessment data in Bulgaria, SPAN Consultants, 2008.

Beaton, A. E. & Allen, N. L. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics*, 17, 191 – 204.

Begle, E. G. (1979).Critical Variables in Mathematics Education: Findings from of a Survey from the Empirical Literature. Washington D.C. Mathematical Association of America and National Council of Teachers of Mathematics.

Blomeke, S. (2007). The Impact of Global Tendencies on the German Teacher Education System. In *Reforming Teaching Globally*, edited by M. T. Tatto, Oxford Studies in Comparative Education.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives*, Handbook I, The Cognitive Domain, Longmans, London.

Campbell, Kelly, Mullis, Martin, & Sainsbury (2001). Framework and Specifications for PIRLS Assessment 2001, Boston College.

Crocker, L. & Algina, J. (1986). *Introduction to Classical and Modern Test Theory,* Harcourt Brace Javanovich College Publishers.

Freudental, H. (1995). Pupil's achievements internationally compared – IEA. *Educational Studies in Mathematics*, 6.

Gonzalez, E.J., Galia, J., & Li, I. (2004). Scaling methods and procedures for TIMSS 2003 mathematics and science scales. In M.O. Martin, I.V.S. Mullis, & S.J. Chrostowski (eds.), *TIMSS 2003 Technical Report*, Chestnut Hill, MA: Boston College.

Good, T., Grouws, D., & Ebmeier, H. (1983). *Active mathematics teaching*. New York: Longman.

Haladyna, T. (1999), *Developing and Validating Multiple-choice Test Items*, Lawrence Erlbaum Associates, Inc. Mahwah, NJ, USA.

Hambelton, R., Swaminathan, H., Rogers, H. (1991). *Fundamentals of item response theory*, SAGE Publications.

Johnson, E. G. and Rust, K.F. (1992). Population references and variance estimation for NAEP data. *Journal of Educational Statistics*, 17, 175-190.

Kellaghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education*. Paris: UNESCO: International Institute for Educational Planning.

Kelvin, G., Bankov, K.. (2006). Exploring the Change in Bulgarian Eight-grade Mathematics Performance from TIMSS 1995 to TIMSS 1999. In *Contexts of Learning Mathematics and Science*, edited by S. Howie and T. Plomp, Routledge.

Key competences for lifelong learning in Europe (2005). http://ec.europa.eu/education/policies/2010/doc/keyrec_en.pdf.

Kish, L. (1965). *Survey sampling.* New York: John Wiley & Sons.

Martin, M., Mullis, I., Gregory, K., Hoyle, C., Ce Shen (2000). Effective schools in science and mathematics, IEA, Boston College.

Martin, M., Mullis, I., Chrostowski, St. (2004). TIMSS 2003 Technical Report, IEA, 275 – 285.

Mislevy, R.J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56*, 177–196.

Mislevy, R., Beaton, A., Kaplan, B. & Sheehan, K. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, v. 29, No. 2, 131–161.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression*. Reading, MA: Addison-Wesley.

Mullis, I., Martin, M., Gonzalez, E., Gregory, K., Garden, R., O'Connor, K., Chrostowski, St., Smith, T. (2000). TIMSS 1999 International Mathematics Report, Boston College.

Mullis, I., Martin, M., Gonzalez, E., Chrostwski, St. (2004). TIMSS 2003 International Mathematics Report, IEA.

Mullis, I.V.S. & Johnson, E.G. (1994). The NAEP scale anchoring process for the 1992 mathematics assessment. In E.G. Johnson & J.E. Carlson (Eds.) *The NAEP 1992 technical report*. Washington, DC: National Center for Educational Statistics, 893 – 907.

Munby, H., Russel, T., & Martin, A. K. (2001). Teacher knowledge and how it develops. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed.). Washington, DC: AERA.

OECD (Organization for Economic Cooperation and Development) (2005). *Teachers Matter Attracting, Developing and Retaining Effective Teachers*, c. 180pp.

Popham W.J. (2001). *The Truth about Testing: An educator's call to action*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

Raminez, F.O., Luo, X., Schofer, E., & Meyer, J.W. (2006). Student achievement and national economic growth. *American Journal of Education, 113,* 1-29.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Danish Institute for Educational Research.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models,* second edition. Sage.

Raudenbush, S.W., & Willms, J.D. (1995). The estimation of school effects. *Journal of Educational and Behavior Statistics*, 20 (4), 307-335.

Reddy, V. (2005). Cross-national achievement studies: Learning from South Africa's participation in the Trends in International Mathematics and Science Study. *Compare, 35,* 63-77.

Rubin, D.B. (1987). *Multiple imputations for non-response in surveys*. New York: Wiley.

Rust, K. (1985). Variance Estimation for Complex Estimators in Sample Surveys. *Journal of Official Statistics*, 1, 381-397.

Rust, K. (1986). Efficient Replicated Variance Estimation. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, pp. 81-87.

Schmidt, W.H., Raizen, S.A, Britton, E.D., Bianchi, L.J. & Wolfe, R.G. (1997). Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics. Dordrecht: Kluwer Academic Publishers.

Schmidt, W., Tatto, T., Bankov, K., Blomeke, S., Cedillo, T., Cogan, L., Han, Sh., Hoang, R., Hsieh, F., Paine, L., Santillan, M. & Schwille, J. (2007). The Preparation Gap: Teacher Education for Middle School Mathematics in Six Countries (MT21 Report), MSU.

Schmidt, W., Blomeke, S., Tatto, T. (2011). *Teacher Education Matters. A study of Middle School Mathematics Teacher Preparation in Six Countries*. Teachers College, Columbia University, NY and London.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57 (1), 1-22.

Snijders TAB & Bokster, RJ. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks: CA: Sage.

Putnam, R. & Borko, H. (2000). What Do New Views of Knowledge and Thinking Have to Say About Research on Teacher Learning? *Educational Researcher*, 29(1), 4-15.

Thomas, N. (1993). Asymptotic corrections for multivariate posterior moments with factored likelihood function. *Journal of Computational and Graphical Statistics, 2,* 309-322.

Thorndike, R. L. (1982). *Applied Psychometrics*. Houghton Mifflin Company, Boston.

TIMSS 2003 School Sampling manual (2001). IEA.

TIMSS Assessment Frameworks and Specifications 2003, IEA, http://isc.bc.edu/timss2003i/frameworks.html

TIMSS 2007 Assessment Frameworks, IEA, http://isc.bc.edu/TIMSS2007/frameworks.html

TIMSS 2011 Assessment Frameworks, IEA, http://isc.bc.edu/TIMSS2011/frameworks.html

White, K.R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin,* 91 (3), 461-481.

Willms, J. D. (1992). *Monitoring school performance: A guide for educators.* Lewes: Falmer.

Willms, J. D. (1995). The Challenge of Developing New Educational Indicators. *Educational Evaluation and Policy Analysis*, Spring 1995, Vol. 17, No. 1, pp. 113-131.

Willms, J. D. (1999). Basic Concepts of Hierarchical Linear Modeling with Applications for Policy Analysis. *Handbook on Educational Policy*, Academic Press.

Willms, J.D. (2002). *Vulnerable children: Findings from Canada's National Longitudinal Survey of Children and Youth*. Edmonton, AB: University of Alberta Press.

Willms, J. D. & Marie-Andree Somers. (2001). Family, Classroom, and School Effects on Children's Educational Outcomes in Latin America. *School Effectiveness and School Improvement*, 2001, vol. 12, No. 4, pp. 409-445.

Wilson, S. M. & Berne, J. (1999). Teacher Learning and the Acquisition of Professional Knowledge. An Examination of Research on Contemporary Professional Development. In A. Iran-Nejad & P. D. Pearson (Eds.) *Review in Research in Education* (Vol. 24). Washington D.C.: American Educational Research Association.

Yamamoto, K., Kulick, E. (2000) Scaling Methodology and Procedures for the TIMSS Mathematics and Science Scales. TIMSS Questionnaire Development, In M. O. Martin, K. D. Gregory, and S. E. Stemler, *TIMSS 1999 Technical Report*. Chestnut Hill, MA: Boston College, pp. 237-263.