

ПРОЯВЛЕНИЕ НА КУЛТУРНИТЕ ЗАВИСИМОСТИ ПРИ ТЕСТОВОТО ОЦЕНЯВАНЕ ПРЕЗ ПРИЗМАТА НА МЕЖДУНАРОДНИТЕ СРАВНИТЕЛНИ ИЗСЛЕДВАНИЯ НА ПОСТИЖЕНИЯТА НА УЧЕНИЦИТЕ

Наталия Василева

Софийски университет „Св. Климент Охридски“

Резюме. Целта на настоящата статия е да се очертаят основните аспекти на културния контекст при тестовото оценяване на постиженията на учениците и в частност – при международните програми, свързани с оценяването. Описани са няколко типични отклонения в процеса на създаването на тестови инструменти, които биха довели до ненадеждни резултати за постиженията на учениците, ако не бъде отчетено влиянието на различни културни характеристики. Споменати са някои от процедурите, които се използват в работата по тези програми с цел отчитането и минимизирането на влиянието на културния контекст върху постиженията на учениците.

Keywords: cultural context, assessment, international student assessment, bias, tests, achievements, student performance

В началото на XX и век, на фона на всеобщото развитие на психосоциалните дисциплини, е публикувана първата скала за интелигентност на Бине. По същото време У. Щерн разработва процедури за тестване на интелигентността, които публикува в книгата си „*The psychological methods of testing intelligence*“ (1914). Заедно с новите инструменти възниква и проблемът за честността, справедливостта, безпристрастността на тестовете и оценяването. Все по-често проблемите, свързани с валидността на инструментариума и равните възможности за изява на тестираните, се разискват в психологическите научни среди. До каква степен оценяването е „честно“ към всеки оценяван, е проблем, който стои много отдавна и пред изследователите, работещи в образователната сфера. И докато върху темата за субективността и обективността на измерването и оценяването има много налична литература, въпросът за културната зависимост на самия инструмент и нейното влияние

върху постиженията на учениците все още не е достатъчно изследван и е повод за все повече съвременни научни дискусии и изследвания. Този проблем е много изразен при работата по международните програми за оценяване на постиженията на учениците (като PISA, TIMSS, PIRLS, ICCS и др.), защото в тях участват много страни с различни култури.

Културният контекст, в който се осъществява човешкото общуване, има определящо влияние при взаимодействието между хората. Културата определя рамката, в която личността се учи да организира своите собствени мисли, действия, емоции и поведение, като паралелно я свързва със заобикалящата среда. Въпреки че хората се раждат в определена култура, тя не е вродена у всеки индивид.

Съществуват много и различни дефиниции за понятието **култура**. Поради постоянното разширяване на обема и съдържанието на това понятие е много трудно да се намери единствена, точна и ясна дефиниция. Все още няма консенсус между антрополози и социолози за определението на понятието култура. Още в средата на миналия век определенията са толкова много, че в книгата на Крьобер и Клакхан са събрани над 150 дефиниции (Kroeber, 1952). Все още се правят опити да се формулира една фундаментална дефиниция на понятието, която да бъде използвана в средите на науките за човека.

Културата е не само продукт на обществото, но и негова неотменна характеристика, тя не съществува откъснато, сама за себе си, а е отражение на определени социални практики и на конкретни исторически общности. В подкрепа на изброеното е приведена дефиницията на ЮНЕСКО: *„Култура е множеството от отличителни духовни, материални, интелектуални и емоционални черти на дадено общество или обществена група, тя обхваща освен изкуството и литературата начина на живот, формите на съжителство, ценностните системи, традициите и вярванията“*.

Културата е абстрактна. Тя съществува в съзнанието и навиците на членовете на определено общество, като обхваща стереотипите в начина на мислене и споделените и общите начини за „правене“ на нещата в него. Никой не може да види културата като такава, но всеки може да я види през човешкото поведение. Тя влияе върху всяка дейност, засягаща личността, включително и върху обучението, което е *„съществена, обща и вечна функция на обществото“* (Андреев, 1996).

Обучението изпълнява изключителна функция при предаването на културните ценности. Както проф. М. Андреев посочва: *„Овладеяването на езика, приобщаването към постиженията на различните видове изкуства, може да стане и извън условията на обучението, но в масов мащаб никога не може да бъде така ефективно, както при условията на обучението“* (Андреев, 1996). Базирайки се на това твърдение, може да се направи изводът, че една от целите на образованието е да предава културата от поколение на поколение,

като решаващата роля в този процес има обучението. Ако се разгледа оценяването като неразделна част от процеса на обучение, възниква въпросът до каква степен различията в културните ценности влияят върху оценяването на постиженията на учениците.

Всяко общество се интересува от качеството на продукта, резултат от определена дейност или процес. Това важи и при структурирането на образователните системи и образователните стандарти. Оценяването е важна и основна част за измерването на образователния продукт, но би трябвало да дава отговор освен за качеството му, така и да измерва адекватността на процеса на обучението, свързан с изискванията на социума.

Сравняването на резултатите от дадено измерване с предварително определени стандарти е една от най-разпространените дефиниции за оценяването. От нея може да се обобщи, че:

- измерването винаги е преди оценяването;
- може да се оценява само това, което е познато.

Основна цел на всяко измерване е да се получи информация, която да позволи да се реагира адекватно при вземането на решения (Стоянова, 1996).

Защо ни е необходима оценката? Оценката дава информация за мястото ни в реалния живот, тя предполага взимане на решение за справяне с даден проблем. Сама по себе си оценката не е цел, тя е средство за постигането ѝ. Всяка получена оценка до голяма степен оказва влияние върху личността, която е оценявана. Самото оценяване води до качествени промени в поведението, подготовката и развитието на всеки човек. От това логично следва да се запитаме до каква степен инструментът, който измерва и чрез който се съпоставят измерваните характеристики, е „честен“ към всяка група от хора, които се сравняват.

В статистическата литература **зависимост (bias)** е термин, отнасящ се до наблюдаването на систематична грешка при изчисленията на дадена стойност. В психометрията се приема, че зависим (предубеден) тест е този, който систематично надценява или подценява стойността на променливата, която трябва да се изследва (Reynolds & Ramsay, 2003). Ако тази зависимост се проявява като функция на някоя културна променлива (като етнос или пол), може да се предположи, че в този тест се наблюдава културна зависимост (или няма културна независимост) или че този тест не е „културно честен“.

Зависимост (bias) в психологическото оценяване

При интерпретирането на резултатите от психологическите и дидактическите оценявания, в частност на резултатите от различни тестове, е задължително да се обърне внимание на областите, които могат да доведат до отклонение на резултата от нормата или до **зависимост на теста (test bias)**. Според различните автори тази зависимост може да бъде причинена от отклонения на различни етапи при създаването на инструмента. Van de Vijver и Tanzer представят таксономия на три вида зависимости в социалните науки, която

обхваща и структурира различните зависимости, водещи до отклонения в измерването (Van de Vijver & Tanzer, 2004):

Зависимост при дефиниране на конструкта (*construct bias*) се наблюдава, когато има само частично припокриване на дефинициите от изследвания конструкт (променливата, която се измерва) между различните култури.

Като пример основно се посочва по-широкото разбиране на концепцията за интелигентността от „източните общества“, включващо и социални умения в допълнение на чисто схоластичната област, която покриват разработените тестове за интелигентност в „западни общества“. Този вид отклонения могат да бъдат причинени също и от недостатъчно уместно дефиниране на различните типове поведение, асоциирани в конструкта за различните култури, или от неправилния подбор на примери, илюстриращи определено поведение, изследвано с конкретния конструкт. Като цяло, причините за зависимост, идващи от отклонения в конструкта, би следвало да се търсят в следните три насоки:

- непълно припокриване на дефинициите в конструкта;
- уместност на включените типове поведение в конструкта;
- неправилен подбор на примери, илюстриращи асоциираното в конструкта поведение.

Зависимостта от избора на методи (*method bias*) най-често се свежда до три типа.

– **Зависимост при изготвянето на извадката (*sample bias*)**, или несравнимост на извадката, се проявява, когато сравняваните извадки се различават в редица допълнителни характеристики, несъвпадащи с основния конструкт. Тогава се наблюдава т.нар. „влияние на незначителните фактори“ (*nuisance factors*), т.е. на тези променливи, които не са обект на прякото изследване, но оказват влияние върху изследвания конструкт. При сравняването на култури, различаващи се много една от друга, съпоставимостта на всички основни характеристики на практика се оказва невъзможна.

Като последица от този вид зависимост всички наблюдавани межкултурни различия, както и „влиянието на незначителните фактори“, могат да бъдат отнесени към основния конструкт. В смисъла на когнитивните тестове такива „незначителни фактори“ могат да бъдат разлики в образователните системи, непознатата изпитна ситуация, различна мотивация и др.

– **Зависимост при администрирането (*administration bias*)** обхваща всички влияния, които са причинени от конкретните и специфични форми на администриране на теста. Тази зависимост може да се дължи на различия в средата на администриране, били те физически, технически или социални. Използването на технически средства вместо класическото писане на хартия, използването на записващи устройства и др. могат да доведат до значителни културни зависимости в някои от допълнителните променливи, като ниво на интерес към инструмента или желание за самостоятелност. Груповото или

индивидуалното оценяване също са форма на администриране на теста, която може да повлияе върху крайните резултати.

Друга причина за възникването на зависимост при администрирането е проблемът с комуникацията между изпитващия и изпитвания. Често недоброто владение на езика на администриране на инструмента от някоя от двете страни води до отклонения в резултата.

– **Зависимост в самия инструмент (*instrument bias*)** възниква тогава, когато е налице различна информираност относно елементите на самия инструмент. Тази зависимост се среща, когато условията на задачите или самите задачи съдържат елементи, характерни за конкретна култура или етническа група (напр. стил на обличане на западните общества).

Различната информираност, свързана с изискванията към начина за решаване на определени задачи, също може да бъде една от причините за възникване на зависимост от този тип.

Зависимост на съдържанието на задачата (*item bias*) обхваща отклоненията на ниво задача. В психометрията тази зависимост се установява, ако изследваните лица от различни културни групи са постигнали еднакви средни резултати на конкретен инструмент, но нямат еднакъв очакван резултат на конкретна задача – тя е по-лесна или по-трудна за тях.

Например причините за тази зависимост могат да се търсят в некоректен или незадоволителен превод на езика на тестирането, в нецелесъобразност на съдържанието на задачата в различните култури или в различната конотация при формулиране на условията на задачите.

В последните години все по-голяма популярност добиват международните сравнителни изследвания на постиженията на учениците. Най-мощни от тях са Програмата за международно оценяване на учениците (PISA) на Организацията за икономическо сътрудничество и развитие (OECD), Международното изследване на уменията по математика и природни науки (TIMSS), Международно изследване на уменията за четене (PIRLS) и Международно изследване на гражданското образование (ICCS) на Международната асоциация за оценяване на постиженията в образованието (IEA). Чрез тях се правят опити да се изработят и въведат единни стандарти за оценяване на постиженията на учениците. Всяко от тези изследвания има стриктно разписана рамка, в която са поместени целите, изискванията и съдържанието на оценяването. При разработването на изследователския инструментариум и при изготвянето на сравнителните анализи специално внимание се отделя на въпросите:

– Доколко постиженията на учениците могат да бъдат тълкувани и интерпретирани по един и същ начин, при положение че са получени в различни културни среди?

– До каква степен принадлежността към определена културна група оказва влияние върху резултатите?

Общ преглед на културните зависимости

Необходимостта от сравнимост на резултатите между различните държави изисква от организаторите да изготвят „културно честен“ инструмент. Но дали това е постижима цел, при положение че авторите на изпитните материали са собствено носители на съответна култура? Има ли механизъм, по който да се прецени дали една задача влиза в „културно противоречие“ с разбиранията и спецификите на определена държава или социална група?

Езикът е един от основните носители на културата. При международните изследвания на постиженията на учениците е характерна работата в различни култури и на различни езици. Отклоненията, които могат да се получат при постиженията на учениците, свързани с проблеми в превода, се отнасят към т.нар „зависимост на задачата“.

Организаторите на международните изследвания като PISA, TIMSS, PIRLS и ICCS се опитват да намалят влиянието на различните езици, на които се провеждат изследванията, върху резултатите на учениците чрез различни процедури, свързани най-вече със строги изисквания към начина на превод и адаптация на материалите. Тези изисквания са посочени в т.нар „технически доклади“ (Martin, M.O., Kelly, D.L., 1996). OECD, 2005: PISA 2003 Technical Report; OECD, 2003: PISA 2000 Technical Report) – задължителна част от документацията за работа по тези международни изследвания. Една от използваните процедури е независимият превод от два различни източника. Този метод се прилага тогава, когато основната версия на материалите е изготвена на повече от един език.

Друга процедура за избягване на културна зависимост вследствие на превода е т. нар. „обратен превод“ (*back translation*). При него първо задачата се превежда на съответния език от основната версия, а след това друг преводач превежда обратно на езика на основната версия. Така, след като носители на съответния език сравнят основната версия и обратния превод, се откриват лингвистичните нюанси, които правят езика носител на културата и в някои случаи водят до намаляване на „културната честност“ на изпитните материали.

И при двете процедури преводачите трябва да отчетат специфичните характеристики на държавата, начина на изразяване и на възприемане на информацията. С други думи, при всеки превод се налага да се адаптират тестовите задачи от езика, на който са написани, на езика, на който ще бъде проведено оценяването. Тази адаптация не трябва да влияе върху конструкта, който се оценява, или върху трудността на задачата. В международните сравнителни изследвания този процес на адаптиране е много строго контролиран. В основата си адаптацията е отговорност на екипа на учащата държава, но впоследствие тя трябва да бъде проверена и одобрена от организаторите чрез процес, наречен „верификация на превода“ (*translation verification*). Верификацията се прилага, за да се гарантира, че независимо от адаптирането на ин-

струмента той остава със същите характеристики като първоначално търсени-те при неговото създаване на основния език (не става по-труден или по-лесен за конкретната държава).

Съдържателната валидност е степента, в която съдържанието на теста е представително за конструкта, който трябва да се измерва (Anastasi, 1988). Задачи, при които се наблюдава **съдържателна зависимост (content bias)**, имат различно поведение в различни групи от ученици – по-трудни или по-лесни са за една група, отколкото за друга.

Въпросът за съдържателната валидност в тестовете за постижения е много комплексен. Важни променливи, влияещи върху тази валидност, са времето на обучението по конкретни теми, способностите на конкретната група и точността и спецификата на задачите за изследваната извадка (Reynolds, Cecil R and Ramsay, Michael, 2003). Разглеждайки съдържателната валидност като изискване при създаването на тест за оценяване на постиженията на учениците, не може да не се спомене, че един от основните принципи при оценяването в обучението е, че не трябва да се оценява нещо, което не е преподавано или изучавано.

При сравнение на две от международните изследвания (TIMSS и PISA) може да се обобщи, че TIMSS (независимо дали се администрира в IV клас, или в VIII клас) оценява постиженията на учениците, придържайки се към учебните програми на всяка участваща държава, докато PISA измерва приложението на знанията, уменията и компетентностите, придобити от учениците до определен етап от тяхното обучение.

Накратко, TIMSS се фокусира върху използването на широко дефинирана учебна програма, отчитайки образователните възможности, които се предоставят на учениците, и факторите, които влияят на това как учениците използват тези възможности. В този модел на учебна програма – съответно по математика и природни науки, се отчитат спецификите, които се очаква учениците да учат, така както са определени в учебните програми на всяка участваща образователна система. Като цяло, TIMSS се интересува от това какво и как се преподава в класните стаи, какви са характеристиките на тези, които го преподават, и накрая какво учениците са научили и какво мислят те за изучаването на тези предмети (Mullis, I.V.S. & Martin, M.O., 2013). От друга страна, PISA се фокусира върху оценяването на способността на 15-годишните да използват своите знания и умения, за да се справят с предизвикателствата на реалния живот. Тази ориентация отразява настъпващите промени в целите и задачите на учебните програми в голяма част от държавите, в които са застъпени техники и стратегии за усвояване на това, което учениците могат да правят, с това, което учат и научават в училище, а не само усвояването на специфично учебно съдържание (OECD, 2009).

При така дефинираната в TIMSS цел съдържателната валидност на тестовете в различните държави се осигурява от стриктното отчитане и спазване на

учебното съдържание, включено в различните учебни програми. Разбира се, наличието на съдържателната валидност не води автоматично до получаване на „честен“ инструмент, но в голяма степен го предполага. По-практическата насоченост и липсата на конкретен клас за изследване при PISA правят по-трудноизпълними задачите за осигуряването на съдържателна валидност. Амбициите на организаторите да представят сравнителни резултати за постиженията на учениците могат в много по-голяма степен да бъдат повлияни от културните характеристики на всяка образователна система, отколкото в TIMSS. Особено сериозен е въпросът за осигуряване на съдържателната валидност в образователни системи като българската, които са по-консервативни, по-теоретични и енциклопедични, т.е. такива, чиито учебни програми наблягат на теорията и отделят по-малко време за упражнение и практика. За да може да бъде осигурена сравнимост на резултатите между различните държави, националните екипи трябва да прегледат всички задачи и да преценят дали оценяваното учебно съдържание е включено по някакъв начин в учебните програми до VII клас включително. В тази връзка осигуряването на съдържателна валидност на задачите в PISA изисква много повече експертиза, защото в повечето случаи няма явно припокриване между измервания от PISA конструктор и конкретното учебно съдържание в учебните програми. Това налага въвличането на групи от експерти (за всяка от измерваните области – четене, математика и природни науки), които да преценят освен всички други характеристики на задачите и тези, свързани с културния контекст.

Представяйки накратко TIMSS и PISA, се вижда, че съществува вероятност резултатите да бъдат повлияни от различни характеристики, касаещи културния контекст. В този смисъл, проблемът за работата в мултикултурна среда в един клас, в едно училище или един регион се проектира в много по-голям мащаб, защото основната цел на международните изследвания е сравняването на образователни системи. По тази причина инструментите, които се използват, трябва да бъдат безпристрастни, честни и еднакви за всички участващи страни. Както през 1999 година Браун отбелязва, *„честността е морален, философски или правен проблем, по който разумните хора могат законно да спорят“*, от друга страна, *„зависимостта е емпирична характеристика на теста, когато се използва в две или повече специфични групи. Зависимостта е статистически измерима, а не се определя на принципа на дебати или становища“* (Reynolds, Cecil R and Ramsay, Michael, 2003).

Културата е неотменна характеристика на обществото. Влиянието на културния контекст върху постиженията на учениците е проблем, който все по-често се дискутира в научните среди. Трудността при определянето на степента на ефекта му върху постиженията на учениците при тестовото оценяване оказва пряко влияние върху разработването на единен и точен инструмент за неговото откриване или отчитане. И докато при културната

зависимост на задачата можем да разчитаме, че психометрията ще ни насочи къде да търсим проблем, породен от влиянието на различните култури, при проблеми със съдържателната валидност можем да разчитаме само на експертната оценка, която също може да бъде повлияна от специфичните характеристики, които всеки експерт носи, като представител на определена култура. Това наличие на предполагаема субективност на експертната оценка дава основание на тестолози, психолози, психометрици, педагози и всички, интересувачи се от влиянието на културния контекст при оценяването, да продължават да търсят методи за неговото по-точно откриване, отчитане и последващо минимизиране.

REFERENCES / ЛИТЕРАТУРА

- Andreev, M. (1996). *Protsesat na obuchenie. Didaktika*. Sofia: Sofia University Press [Андреев, М. (1996). *Процесът на обучение. Дидактика*. София: Университетско издателство].
- Andreev, M. (1998). *Образование i obshtestvo. Pedagogicheska sotsiologiya*. Sofia: Sofia University Press [Андреев, М. (1998). *Образование и общество. Педагогическа социология*. София: Университетско издателство].
- Stoyanova, F. (1996). *Testologiya za uchiteli*. Sofia: Atika [Стоянова, Ф. (1996). *Тестология за учители*. София: Атика].
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2002). *Standards for educational and psychological testing*. Washington, DC.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Domino, G. & Domino, M.L. (2006). *Psychological Testing: An Introduction*. Cambridge University Press.
- Kroeber, A. & C. Kluckhohn (1952). *Culture: A Critical Review of Concepts and Definitions*. Harvard University Press.
- Martin, M.O. & Kelly, D.L. (1996). *Third International Mathematics and Science Study Technical Report Volume I: Design and Development*, Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.
- Mullis, I.V.S. & Martin, M.O. (Eds.). (2013). *TIMSS 2015 Assessment Frameworks*. TIMSS & PIRLS International Study Center, Boston College.
- OECD (2009). *PISA 2009. Assessment Framework – Key Competencies in Reading, Mathematics and Science*.
- OECD (2005). *PISA 2003 Technical Report*. OECD Publishing.

- OECD (2003). *PISA 2000 Technical Report*. OECD Publishing.
- Reynolds, C. R. & Ramsay, M. C. (2003). Bias in psychological assessment: An empirical review and recommendations; *Handbook of psychology*; Wiley Online Library.
- Van de Vijver, F. & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 54 (2), 119 – 135.
- Weiner, I.B. and Freedheim, D.K. and Graham, J.R. & Schinka, J.A. and Velicer, W.F. (2003). *Handbook of Psychology, Assessment Psychology*, Wiley.

THE IMPACT OF CULTURAL CHARACTERISTICS IN TEST ASSESSMENT IN THE SCOPE OF THE INTERNATIONAL COMPARATIVE STUDIES OF STUDENTS' ACHIEVEMENTS

Abstract. This article aims to outline the main aspects of the cultural context in the test evaluation of students' achievement and particularly in the international assessment studies. A few typical deviations in the process of creating test instruments are described, which could lead to unreliable results of students' performance if the influence of different cultural characteristics is not recognised. The article focuses on some of the procedures used in the development and in the work on international studies such as PISA, TIMSS, PIRLS and ICCS to report on and minimise the impact of cultural influence on students' achievements.

✉ **Ms. Natalia Vassileva, PhD Student**
Faculty of Education
University of Sofia
15, Tsar Osvoboditel Blvd.
1504 Sofia, Bulgaria
E-mail: n.vassileva@mon.bg