APA standardi za psihološka testiranja i testiranja u obrazovanju

Izrada, evaluacija i dokumentacija testova

**1. Validnost**

**1.1**

Treba ponuditi obrazloženje za svaku preporučenu interpretaciju i upotrebu skorova testa, zajedno sa sveobuhvatnim pregledom dokaza i teorije koji se tiču predviđene upotrebe ili interpretacije.

**1.2**

Sastavljač testova treba jasno da prikaže kako bi skorovi testova trebalo da se tumače i koriste. Potrebno je precizno odrediti populaciju (populacije) za koju je test prikladan i jasno opisati konstrukt čijoj proceni je test namenjen.

**1.3**

Ako validnost nekog uobičajenog ili verovatnog tumačenja nije ispitana, ili ako tumačenje nije u skladu sa dostupnim dokazima, tu činjenicu treba jasno istaći i upozoriti potencijalne korisnike na vršenje neosnovanih tumačenja.

**1.4**

Ako se test koristi na način koji nije proveren, korisnik je dužan da opravda novu uportebu, prikupljajući nove dokaze ukoliko je to potrebno.

**1.5**

Strukturu svakog uzorka ispitanika iz kog se pribavljaju dokazi o validnosti treba opisati u što više detalja, a da ostane praktično, uključujući i glavne relevantne sociodemografske i razvojne karakteristike.

**1.6.**

Kada se validacija delimično zasniva na prikladnosti sadržaja testa, treba opisati i obrazložiti primenjene procedure pri određivanju i kreiranju sadržaja testa u odnosu na konstrukt testa koji treba izmeriti i oblasti koju treba obuhvatiti. Ako definicija uzorkovanog sadržaja testa sadrži kriterijume kao što su važnost, učestalost ili nužnost, ove kriterijume takođe treba precizno objasniti i obrazložiti.

**1.7**

Kada validacija delimično počiva na mišljenju ili odlukama stručnjaka, posmatrača ili procenjivača, procedure za odabir takvih stručnih lica i za podsticanje mišljenja ili procena treba detaljno opisati. Treba prikazati kvalifikacije i iskustvo stručnjaka. Opis procedura treba da sadrži i bilo kakvu obuku i dobijene instrukcije, da ukaže da li su učesnici samostalno došli do svojih odluka i da izvesti o stepenu saglasnosti koji je postignut. Ukoliko su učesnici bili u međusobnoj interakciji ili razmenjivali informacije, treba prikazati postupke kojima su oni možda uticali jedni na druge.

**1.8**

Ako obrazloženje za korišćenje testa ili interpretaciju skorova zavisi od pretpostavki o psihološkim procesima ili kognitivnim operacijama koje su primenjivali ispitanici, onda treba pružiti teorijske ili empirijske dokaze koji podržavaju takve pretpostavke. Treba obezbediti slične informacije kada su tvrdnje o procesima koje su primenili posmatrači i ocenjivači deo argumenta o validnosti testa.

**1.9**

Ukoliko postoji tvrdnja da na test suštinski ne utiču praksa i obuka, onda treba dokumentovati osetljivost testa na promenu usled menjanja oblika uputstava.

**1.10**

Kada se predlaže tumačenje učinka na određenim zadacima, ili malim podgrupama zadataka, trebalo bi pružiti obrazloženje i relevantne dokaze koji podržavaju takva tumačenja. Kada sastavljač smatra da je tumačenje pojedinačnih odgovora na zadatke moguće, ali ne i preporučljivo, treba upozoriti korisnika da ne vrši takva tumačenja.

**1.11**

Ukoliko obrazloženje koriščenja ili tumačenja testa zavisi od pretpostavki o odnosima između delova testa, treba pružiti dokaze koji se tiču unutrašnje strukture testa.

**1.12**

Kada se predlaže interpretacija subskorova, razlika u skorovima ili profila, treba pružiti obrazloženje i relevantne dokaze koji podržavaju takve interpretacije. Kada se dobiju kompozitni skorovi, treba ponuditi osnovu i obrazloženje za njihovo dobijanje.

**1.13**

Kada dokaz o validnosti obuhvata statističku analizu rezultata testa, bilo pojedinačnih ili zajedno sa podacima o drugim varijablama, uslove pod kojim su prikupljeni podaci treba opisati dovoljno detaljno da korisnici mogu da procene značaj statističkih nalaza za lokalne uslove. Trebalo bi skrenuti pažnju na sve odlike prikupljanja validacijskih podataka koji će verovatno odstupiti od tipičnih operativnih uslova za testiranje i koji bi možda mogli da utiču na učinak testa.

**1.14**

Kada dokaz o validnosti obuhvata empirijsku analizu odgovora na testu zajedno sa podacima o drugim varijablama, trebalo bi pružiti obrazloženje o odabiru dodatnih varijabli. Tamo gde je to pogodno i izvodljivo, treba prikazati ili citirati dokaze koji se odnose na konstrukte predstavljene drugim varijablama, kao i njihove tehničke karakteristike. Treba skrenuti pažnju na bilo koji mogući izvor zavisnosti (ili nedostatka nezavisnosti) između varijabli osim zavisnosti među konstruktima koje one predstavljaju.

**1.15**

Kada postoji tvrdnja da određeni nivo učinka testa predviđa adekvatne i neadekvatne učinke kriterijuma, treba priložiti podatke o nivoima učinka kriterijuma koji su u vezi sa datim nivoima skorova testa.

**1.16**

Kada se validacija zasniva na dokazu da su skorovi testa povezani sa jednom ili više varijabli kriterijuma, treba saopštiti podatke o prikladnosti i tehničkom kvalitetu kriterijuma.

**1.17**

Ako se skorovi testa koriste zajedno sa drugim merljivim varijablama da bi se predvideo neki ishod ili kriterijum, regresione (ili ekvivalentne) analize treba da obuhvataju te dodatne relevantne varijable zajedno sa skorovima testa.

**1.18**

Kada se vrše statističke korekcije, kao što su one za restrikciju raspona ili umanjenje, treba izneti i korigovane i nekorigovane koeficijente, kao i tačno određenu proceduru koja je primenjena i sve statističke podatke korišćene pri korekcijama.

**1.19**

Ako je preporučeno da se test koristi ili će se verovatno koristiti za određivanje ljudi za alternativne tretmane, i ako se ishodi tih tretmana mogu opravdano porediti prema zajedničkom kriterijumu, onda, kada god je to izvodljivo, treba pružiti dokaze u prilog različitim ishodima.

**1.20**

Kada se koristi meta-analiza kao dokaz o jačini veze između testa i kriterijuma, trebalo bi da varijable testa i kriterijuma na tom primeru mogu da se uporede sa onima iz studija koje su sažete. Ako relevantna istraživanja obuhvataju pouzdane dokaze da bilo koja druga karakteristika primene testiranja može da utiče na jačinu veze između testa i kriterijuma, treba opisati podudarnost tih karakteristika na tom primeru i u meta-analizi. Treba eksplicitno naglasiti svaku značajnu razliku koja bi mogla da ograniči primenljivost meta-analitičkih nalaza na tom primeru.

**1.21**

Potrebno je jasno opisati svaki meta-analitički dokaz koji se koristi da zastupa predviđenu upotrebu testa, uključujući i metodološke izbore pri identifikovanju i kodiranju tokom istraživanja, korekciji za artifakte i ispitivanju potencijalnih moderatorskih varijabli. Treba izneti pretpostavke zasnovane na korekciji za artifakte kao što su nepouzdanost kriterijuma i restrikcija raspona, kao i razjasniti posledice ovih pretpostavki.

**1.22**

Kada je jasno navedeno ili implicirano da će preporučena upotreba testa rezultirati određenim ishodom, treba navesti osnovu za očekivanje takvog ishoda, zajedno sa relevantnim dokazima.

**1.23**

Kada se korišćenje testa ili interpretacija skorova preporučuje na osnovu toga da će testiranje ili program testiranja sam po sebi rezultirati nekom indirektnom koristi pored korisnosti informacija iz samih skorova testa, treba eksplicitno dati obrazloženje o očekivanju te indirektne koristi. Za nju je potrebno pružiti logičke ili teorijske argumente i empirijske dokaze. Treba skrenuti odgovarajuću pažnju na svaki kontradiktorni nalaz koji nagoveštava važne indirektne ishode osim onih predviđenih.

**1.24**

Kada korišćenje testa prouzrokuje neplanirane posledice, treba da se ispita da li takve posledice proizilaze iz osetljivosti testa na druge karakteristike čijoj proceni test nije namenjen ili usled nemogućnosti testa da predstavi predviđeni konstrukt.

**2. Pouzdanost i greške merenja**

**2.1**

Za svaki ukupan skor, subskor ili kombinaciju skorova koji će se tumačiti, treba navesti procene relevantnih pouzdanosti i standardnih grešaka merenja ili funkcija informativnosti testa.

**2.2**

Standardne greške merenja, i opšte i uslovne (ukoliko su važne) treba navesti i u jedinicama sirovih skorova ili originalne skale i u jedinicama svakog izvedenog skora koje je preporučljivo koristiti pri tumačenju testa.

**2.3**

Kada tumačenje testa ističe razlike između dva posmatrana skora pojedinca ili dve prosečne vrednosti grupe, potrebno je obezbediti podatke o pouzdanosti za te razlike, uključujući i standardne greške.

**2.4**

Svaki metod merenja preciznosti ili doslednosti skorova treba jasno opisati i objasniti u pogledu statistike prikladne za ovaj metod. Treba navesti procedure uzorkovanja koje su korišćene za odabir ispitanika za analize pouzdanosti i deskriptivnu statistiku ovih uzoraka.

**2.5**

Koeficijent pouzdanosti ili standardne greške merenja zasnovane na jednom pristupu ne treba tumačiti kao uzajamno zamenljive nekim drugim dobijenim drugom tehnikom, osim ako njihove implicitne definicije greške merenja nisu ekvivalentne.

**2.6**

Ako su koeficijenti pouzdanosti prilagođeni restrikciji raspona ili varijabilnosti, treba predstaviti proceduru prilagođavanja, kao i korigovani i nekorigovani koeficijent. Treba predstaviti i standardne devijacije grupe koja je zaista testirana i ciljne populacije, kao i obrazloženje prilagođavanja.

**2.7**

Kada specifikacije testa određuju podgrupe zadataka u okviru testa i može se pretpostaviti da mere delimično nezavisne crte ličnosti ili sposobnosti, procedura procene pouzdanosti treba da uzme u obzir multifaktorski karakter instrumenta.

**2.8**

Treba obavestiti korisnike testa o stepenu do kog stopa rada može uticati na učinak ispitanika.

**2.9**

Kada je test osmišljen tako da odražava stopu rada, pouzdanost treba proceniti uz pomoć pristupa paralelnih formi (*alternate-form approach*) ili *test-retest* pristupa, primenjujući test u različito vreme.

**2.10**

Kada je subjektivni sud uključen u bodovanje testa, treba pružiti dokaze da su procenjivači i ispitanici konzistentni uz ponovljena merenja. Potrebno je napraviti jasnu razliku između podataka o pouzdanosti koji se zasnivaju na (a) nezavisnim grupama procenjivača koji boduju iste učinke ili produkte, (b) jednoj grupi koja boduje uzastopne učinke ili nove produkte, (c) nezavisne grupe koje boduju uzastopne učinke ili nove produkte.

**2.11**

Ukoliko postoje opšte prihvaćeni teorijski i empirijski razlozi da se očekuje da će se koeficijenti pouzdanosti, standardne greške merenja ili funkcije informativnosti testa suštinski razlikovati za različite subpopulacije, izdavači treba da obezbede podatke o pouzdanosti čim je to izvodljivo za svaku veću populaciju za koju se test preporučuje.

**2.12**

Ukoliko je predloženo da se test koristi u nekoliko razreda ili u rasponu hronološki raspoređenih uzrasnih grupa i ako su obezbeđene zasebne norme za svaki razred, treba pružiti podatke o pouzdanosti za svaki uzrast ili razred, ne samo za sve razrede ili uzraste zajedno.

**2.13**

Ako su angažovani lokalni ocenjivači kako bi primenili opšta pravila i principe bodovanja koje je utvrdio sastavljač testa, lokalne vlasti treba da prikupe i prikažu lokalne podatke o pouzdanosti kada adekvatna veličina uzorka bude dostupna.

**2.14**

Uslovne standardne greške merenja treba prikazati u nekoliko nivoa bodovanja ako ne može da se uspostavi konstantnost. Kada su određeni granični skorovi za odabir ili klasifikaciju, treba prikazati standardne greške merenja u blizini svakog graničnog skora.

**2.15**

Kada se test ili kombinacija mera koristi kako bi se donele kategoričke odluke, treba pružiti procene procenta ispitanika koji bi bili klasifikovani na isti način na osnovu dve procedure, koristeći istu formu ili paralelne forme instrumenta.

**2.16**

U nekim situacijama prilikom testiranja zadaci variraju od ispitanika do ispitanika – od naizmeničnog odabira iz baze stavki ili primene algoritama koji se zasnivaju na nivou učinka ispitanika na prethodnim zadacima do odabira na osnovu težine zadataka. Prilikom ove vrste testiranja poželjan pristup za procenu pouzdanosti jeste onaj koji se zasniva na uzastopnim primenama testa pri uslovima sličnim onim koji preovladavaju u operativnoj uporebi.

**2.17**

Kada je test dostupan i u dugoj i u kratkoj verziji, treba prikazati podatke o pouzdanosti za skorove obe verzije, po mogućstvu na osnovu njihove nezavisne primene.

**2.18**

Kada su dozvoljene značajne varijacije u procedurama za primenu testa, treba pružiti posebne analize pouzdanosti za skorove koji su dobijeni pri svakoj većoj varijaciji ukoliko su adekvatne veličine uzoraka dosupne.

**2.19**

Kada se prosečni skorovi testa grupe koriste prilikom evaluacije programa, testirane grupe generalno treba posmatrati kao uzorak neke veće populacije, čak i ako su svi ispitanici dostupni u vreme merenja testirani. U tim slučajevima treba da se prikaže standardna greška srednje vrednosti grupe, s obzirom na to da odražava varijabilnost prouzrokovanu uzorkovanjem ispitanika, kao i varijabilnost prouzrokovanu greškom merenja.

**2.20**

Kada je svrha testiranja ta da izmeri učinak grupe pre nego pojedinca, često je korišćena procedura da se dodeli mali podskup zadataka svakom od mnogih poduzoraka ispitanika. Podaci se prikupljaju preko poduzoraka i podgrupa zadataka kako bi se izmerio učinak grupe. Kada se takve procedure koriste za evaluaciju programa ili opis populacije, analize pouzdanosti moraju uzeti u obzir šemu uzorkovanja.

**3. Izrada i revizija testova**

**3.1**

Testove i programe testiranja treba razvijati na čvrstoj naučnoj osnovi. Sastavljači i izdavači testova treba da prikupe i dokumentuju adekvatne dokaze koji se tiču izrade testova.

**3.2**

Treba jasno navesti svrhu (svrhe) testa, definiciju oblasti i specifikacije testa tako da se mogu doneti sudovi o prikladnosti definisane oblasti za navedenu svrhu (svrhe) testa, kao i o odnosu zadataka i dimenzija oblasti koju oni treba da obuhvate.

**3.3**

Treba dokumentovati specifikacije testa, zajedno sa njihovim obrazloženjem i procesom tokom kojeg su one kreirane. Specifikacije testa treba da definišu sadržaj testa, preporučen broj zadataka, formate zadataka, željene psihometrijske karakteristike zadataka i raspored zadataka i odeljaka. One takođe treba da preciziraju količinu vremena predviđenog za testiranje, smernice za ispitanike, procedure koje treba slediti prilikom primene i bodovanja testa, kao i sve ostale relevantne informacije.

**3.4**

Potrebno je dokumentovati procedure koje se koriste za interpretaciju skorova i, kada je to prikladno, normativne uzorke, uzorke za standardizaciju ili korišćeni kriterijum.

**3.5**

Kada je prigodno, relevantni stručnjaci koji ne učestvuju u programu testiranja treba da pregledaju specifikacije testa. Treba dokumentovati svrhu provere, proces sprovođenja provere i njene rezultate. Kvalifikacije, relevantna iskustva i demografske karakteristike stručnjaka takođe treba dokumentovati.

**3.6**

Vrstu zadataka, formate odgovora, procedure bodovanja i primene testa treba izabrati na osnovu svrha testa, oblasti koja će se meriti i predviđenih ispitanika na testu. Koliko god da je moguće, treba izabrati sadržaj testa tako da predviđeni zaključci iz skorova testa podjednako važe za članove različitih grupa ispitanika. Proces provere testa treba da obuhvata empirijsku analizu i, kada je to pogodno, upotrebu stručnjaka za proveru zadataka i formata odgovora. Kvalifikacije, relevantna iskustva i demografske karakteristike stručnjaka takođe treba dokumentovati.

**3.7**

Treba dokumentovati postupke koji su korišćeni pri izradi, proveri, testiranju i odabiru zadataka iz baze stavki. Ukoliko su zadaci klasifikovani u različite kategorije ili podgrupe na osnovu specifikacija testa, treba dokumentovati procedure koje su korišćene pri klasifikaciji, kao i adekvatnost i preciznost klasifikacije.

**3.8**

Kada se vrše ispitivanja ili testiranja zadataka na terenu, procedure korišćene pri uzorkovanju ispitanika za testiranje zadataka, karakteristike uzor(a)ka ispitanika za testiranje zadataka i karakteristike uzor(a)ka treba da što bolje predstavljaju populaciju (populacije) kojoj je test namenjen.

**3.9**

Kada sastavljač testa vrši evaluaciju psihometrijskih karakteristika zadataka, treba dokumentovati klasični ili IRT (*item response theory*) model korišćen za procenu psihometrijskih karakteristika zadataka. Treba opisati uzorak korišćen za procenu karakteristika zadataka koji treba da bude adekvatne veličine i raznolikosti za proceduru. Proces odabira zadataka i podaci korišćeni prilikom odabira, kao što su težina zadataka, diskriminativnost i/ili informativnost zadataka takođe treba da se dokumentuju. Kada se IRT model koristi za procenu parametara zadataka prilikom izrade testa, model odgovora na zadatke, postupak procene i dokazi o tome koliko dobro se model uklapa treba da se dokumentuju.

**3.10**

Sastavljači testa treba da sprovedu unakrsne studije validacije kada se zadaci biraju prvenstveno na osnovu empirijskih odnosa, pre nego na osnovu sadržaja ili teorijskih razmatranja. Treba dokumentovati do koje mere različita istraživanja identifikuju isti postavljeni zadatak.

**3.11**

Sastavljači testa treba da dokumentuju u kojoj meri sadržaj testa predstavlja definisanu oblast i specifikacije testa.

**3.12**

Potrebno je pružiti obrazloženje i dokaze koji podržavaju kompjuterski adaptivne testove. Ova dokumentacija treba da sadrži procedure koje su korišćene pri odabiru podgrupa primenjenih zadataka, pri određivanju uslova za početak i završetak testa, pri bodovanju testa i pri kontrolisanju izlaganja zadataka.

**3.13**

Kada se skor testa određuje na osnovu različite težine zadataka, sastavljač testa treba da navede obrazloženje i proces korišćen pri izradi, proveri i određivanju težine zadataka. Kada se težina zadatka određuje na osnovu empirijskih podataka, uzorak korišćen za određivanje težine zadatka treba da bude dovoljno velik i da predstavlja populaciju kojoj je test namenjen. Kada se težina zadatka određuje na osnovu stručne procene, treba dokumentovati kvalifikacije stručnjaka.

**3.14**

Potrebno je dokumentovati primenjene kriterijume za bodovanje učinka ispitanika na pitanjima otvorenog tipa. Ova dokumentacija je posebno važna za procenu učinka, na primer za portfolije i eseje za koje se moze odrediti skor, gde kriterijum za bodovanje korisniku možda nije tako očigledan.

**3.15**

Kada se koristi standardizovani format testiranja za ispravljanje strukturisanih uzoraka ponašanja, treba dokumentovati oblast, izgled testa, specifikacije testa i materijal kao i za bilo koji drugi test. Takva dokumentacija treba da obuhvata jasnu definiciju ponašanja koja se očekuje od ispitanika, prirodu očekivanih odgovora i sav materijal ili uputstva koji su neophodni kako bi se sprovelo testiranje.

**3.16**

Ukoliko je pripremljena kratka verzija testa, na primer redukovanjem broja zadataka na originalnom testu ili uređivanjem posebne verzije od delova originalnog testa, specifikacije kratke verzije treba da budu što sličnije onima za originalni test. Treba dokumentovati procedure koje su primenjene prilikom redukcije broja zadataka.

**3.17**

Kada prethodna istraživanja pokazuju da irelevantne varijanse mogu da poremete definiciju oblasti na kojoj se test zasniva, onda, koliko je to izvodljivo, sastavljač testa treba da istraži takve izvore irelevantnih varijansi i da ih ukloni ili redukuje.

**3.18**

Za testove koji imaju vremenska ograničenja, prilikom istraživanja radi sastavljanja testa treba ispitati do kog stepena skorovi obuhvataju komponentu brzine i proceniti adekvatnost te komponente, u zavisnosti od oblasti čijem je merenju test namenjen.

**3.19**

Uputstva za primenu testa treba naglasiti i predstaviti dovoljno jasno kako bi i ostali mogli da adekvatno ponove uslove primene na osnovu kojih su pribavljeni podaci o pouzdanosti i validnosti, kao i o normama kada je to moguće.

**3.20**

Instrukcije ponuđene ispitanicima treba da sadrže dovoljno detalja da bi ispitanici mogli da odgovore na zadatke onako kako je bila namera sastavljača testa. Kada je to moguće, uzorke materijala, vežbanja i uzorke pitanja, kriterijume za bodovanje i reprezentativni zadatak koji predstavlja svaku veću oblast u klasifikaciji testa ili oblast u celini treba dostaviti ispitanicima pre sprovođenja testa ili zajedno sa materijalom za testiranje kao deo standardnih instrukcija za primenu testa.

**3.21**

Ako sastavljač testa ukaže da je dozvoljeno variranje uslova primene testa od jednog do drugog ispitanika ili grupe, treba definisati dozvoljene varijacije uslova primene i treba ponuditi obrazloženje o dozvoljavanju različitih uslova.

**3.22**

Procedure za bodovanje i, ukoliko je relevantno, kriterijume za bodovanje sastavljač testa treba da prezentuje dovoljno detaljno i jasno kako bi ocenjivanje bilo što tačnije. Uputstva za korišćenje rejting skala ili za izvođenje skorova dobijenih kodiranjem, skaliranjem ili klasifikovanjem formiranih odgovora treba da budu jasna. Ovo je posebno značajno ukoliko se testovi mogu ocenjivati na lokalnom nivou.

**3.23**

Sastavljač testa treba da dokumentuje proces odabira, obuke i kvalifikovanja ocenjivača. Materijali za obuku, kao što su uputstva za ocenjivanje i primeri odgovora ispitanika koji ilustruju nivoe na skali za bodovanje, i procedure za obuku ocenjivača treba da dovedu do usaglašenosti ocenjivača kako bi se skorovi tumačili onako kako je prventsveno bila namera sastavljača testa. Osoba (osobe) koja je odgovorna za sprovođenje obuke treba da proceni i iznese pouzdanost ocenjivača i potencijalna pomeranja njihovih standarda u ocenjivanju tokom vremena.

**3.24**

Kada se ocenjivanje vrši na lokanom nivou i potrebna je procena ocenjivača, korisnik testa je dužan da obezbedi odgovarajuću obuku i instrukcije ocenjivačima i odgovoran je za proveru usaglašenosti i preciznosti ocenjivača. Sastavljač testa treba da navede očekivani nivo usaglašenosti i preciznosti ocenjivača.

**3.25**

Test treba izmeniti i revidirati kada novi podaci istraživanja, značajne promene u predstavljenoj oblasti ili nedavno preporučeni uslovi upotrebe testa mogu smanjiti validnost interpretacije skorova. Iako nije potrebno povući ili revidirati test koji je i dalje koristan samo zato što je prošlo određeno vreme, sastavljači i izdavači testa su odgovorni za praćenje promena uslova i za menjanje, revidiranje ili povlačenje testa, kao što je ukazano.

**3.26**

Test bi trebalo da bude označen i promovisan kao “revidiran” tek onda kada je revidiran u značajnoj meri. Izraz kao što je “sa manjim izmenama” treba koristiti kada je test izmenjen u manjoj meri. Skala za bodovanje treba da se prilagodi da bi odgovarala ovim izmenama, a treba i obavestiti korisnike o korekcijama do kojih je došlo na skali.

**3.27**

Ukoliko je test ili deo testa namenjen samo za potrebe istraživanja i nije distribuiran za operativnu upotrebu, izveštaje o tome treba jasno istaći na svoj relevantnoj dokumentaciji za test i materijalu za tumačenje koji se dostavlja korisniku testa.

**4. Skaliranje, normiranje i uporedivost skorova**

**4.1**

Dokumentacija testova treba da pruži korisnicima testa jasna objašnjenja o predviđenom tumačenju značenja izvedenih skala za bodovanje, kao i njihovih ograničenja.

**4.2**

Konstrukciju skala korišćenih za objavljivanje skorova treba jasno opisati u dokumentaciji.

**4.3**

Ukoliko se opravdano veruje da su moguća određena pogrešna tumačenja skale za bodovanje, korisnike testa treba na to eksplicitno upozoriti.

**4.4**

Ukoliko je predviđeno tumačiti direktno sirove skorove, njihova značenja, predviđena tumačenja i ograničenja treba opisati i obrazložiti na isti način kao što je to slučaj sa izvedenim skalama za bodovanje.

**4.5**

Norme, ukoliko su upotrebljavane, treba da se odnose na jasno opisane populacije. Ove populacije treba da obuhvataju pojedince i grupe sa kojima bi korisnici testa najčešće želeli da uporede svoje sopstvene ispitanike.

**4.6**

Izveštaji o normativnim studijama treba da sadrže precizne specifikacije populacije iz koje je uzet uzorak, procedure uzorkovanja i stope učešća, bilo kakvo kolebanje uzorka, datume testiranja i deskriptivne statističke podatke. Pružene informacije treba da budu dovoljne kako bi korisnici mogli da procene prikladnost normi za tumačenje skorova lokalnih ispitanika. Tehnička dokumentacija treba da ukaže na preciznost samih normi.

**4.7**

Ako se lokalne grupe ispitanika suštinski razlikuju od populacija na koje se norme odnose, korisnik koji saopštava izvedene skorove zasnovane na objavljenim normama u obavezi je da opiše te razlike ukoliko se one tiču interpretacije objavljenih skorova.

**4.8**

Kada se norme koriste kako bi okarakterisale grupe ispitanika, statističke podatke korišćene da se sažme učinak svake grupe i norme na koje se ti podaci odnose treba jasno definisati, isto tako oni treba da podrže predviđenu upotrebu ili tumačenje.

**4.9**

Kada su skale sa sirovim ili izvedenim skorom osmišljene za tumačenje na osnovu kriterijuma, uključujući i klasifikaciju ispitanika u odvojene kategorije, obrazloženje za preporučene interpretacije skorova treba da bude jasno objašnjeno.

**4.10**

Potrebno je pružiti jasno obrazloženje i potkrepljujuće dokaze za svaku tvrdnju da skorovi ostvareni na različitim formama testa mogu biti uzajamno zamenljivi. U nekim slučajevima, treba obezbediti direktne dokaze o ekvivalentnosti skorova. U drugim slučajevima, dokazi mogu proizići iz demonstracije da su teorijske pretpostavke na kojima se zasnivaju procedure za utvrđivanje uporedivosti skorova zadovoljene u dovoljnoj meri. Određeno obrazloženje i potrebni dokazi će delimično zavisiti od predviđenih upotreba za koje je zahtevana ekvivalentnost skorova.

**4.11**

Kada se tvrdnje o ekvivalentnosti skorova od forme do forme zasnivaju na postupcima izjednačavanja, treba obezbediti detaljne tehničke informacije o metodu po kom su funkcije izjednačavanja (*equating functions*) i druge relacije ustanovljene, kao i o tačnosti funkcija izjednačavanja.

**4.12**

U studijama koje se oslanjaju na statističku ekvivalentnost testiranja grupa ispitanika koji dobijaju različite forme testa, treba detaljno opisati metode koji obezbeđuju tu ekvivalentnost.

**4.13**

U studijama koje koriste izgled referentog testa (*anchor test*) treba predstaviti karakteristike referentnog testa i njegovu sličnost sa formama koje se izjednačavaju, uključujući i specifikacije sadržaja i empirijski određene veze između skorova testa. Ako se koriste referentni zadaci (*anchor items*), kao u nekim klasičnim studijama ili onima koje se zasnivaju na IRT modelu, treba prikazati reprezentativnost i psihometrijske karakteristike referentnih zadataka.

**4.14**

Kada se neke procedure poređenja ili konverzije skorova koriste da povežu skorove na testovima ili na formama testa koje nisu striktno paralelne, treba jasno opisati izradu, predviđeno tumačenje i ograničenja ovih konverzija ili poređenja.

**4.15**

Kada se kreiraju dodatne forme testa uzimajući podgrupu zadataka iz postojeće forme testa ili menjanjem rasporeda njegovih zadataka i postoji opravdan razlog da se veruje da na skorove tih formi mogu uticati efekti konteksta zadataka, treba obezbediti dokaze da ne postoji neprikladno narušavanje normi za različite verzije ili relacija skorova među njima.

**4.16**

Ukoliko se specifikacije testa promene od jedne do druge verzije testa, te promene treba utvrditi u upustvu za test i ukazati na to da konvertovani skorovi za te dve verzije možda neće biti striktno ekvivalentni. Kada dođe do značajnih promena specifikacija testa, skorove treba prikazati na novoj skali ili ponuditi jasan izveštaj kako bi se upozorili korisnici da skorovi nisu direktno uporedivi sa onima od ranijih verzija testa.

**4.17**

Programi za testiranje koji pokušavaju da održe opštu skalu tokom vremena treba da sprovode periodične provere stabilnosti skale na kojoj se prikazuju skorovi.

**4.18**

Ukoliko izdavač obezbedi norme za korišćenje pri interpretaciji skorova testa, onda dokle god je test i dalje u štampi, izdavač je u obavezi da osigura da se test dovoljno često ponovo normira kako bi kontinuirana tačnost i prikladnost interpretacije skorova bile omogućene.

**4.19**

Kada preporučene interpretacije skorova obuhvataju jedan ili više granični skor, treba jasno dokumentovati obrazloženje i procedure korišćene pri utvrđivanju graničnih skorova.

**4.20**

Kada je to izvodljivo, treba utvrditi kategorije za definisanje graničnih skorova sa jasnim suštinskim interpretacijama na osnovu pouzdanih empirijskih podataka vezanih za odnos učinka testa i relevantih kriterijuma.

**4.21**

Kada se granični skorovi koji definišu uspeh - neuspeh ili nivoe znanja zasnivaju na direktnoj proceni adekvatnosti zadatka ili učinka testa ili nivoa učinka, proces procene treba da bude osmišljen tako da ocenjivači mogu da upotrebe svoje znanje i iskustvo u razumnoj meri.

**5. Primena testa, bodovanje i objavljivanje rezultata**

**5.1**

Primena testa treba da pažljivo sledi standardizovane procedure primene i bodovanja koje je sastavljač testa precizirao, osim ukoliko situacija ili nemogućnost ispitanika nalaže da treba napraviti izuzetak.

**5.2**

Treba dokumentovati izmene ili remećenje standardizovanih procedura za primenu ili bodovanje testa.

**5.3**

Kada se uspostave formalne procedure za traženje i ostvarivanje prilagođavanja, treba obavestiti ispitanike o ovim procedurama pre testiranja.

**5.4**

Okruženje predviđeno za testiranje treba da pruži određeni komfor uz minimalne distrakcije.

**5.5**

U instrukcijama treba jasno naglasiti ispitanicima kako da daju odgovore. Potrebno je takođe dati instrukcije povodom korišćenja bilo kakve opreme koja bi mogla biti nepoznata ispitanicima. Treba im pružiti priliku da vežbaju davanje odgovora ukoliko je uključena upotreba opreme, osim u slučaju da se ocenjuje upravo upotreba opreme.

**5.6**

Treba uložiti određen trud kako bi se osigurao integritet skorova testa tako što će se ispitanicima onemogućiti da ostvare skorove koristeći nedozvoljena sredstva.

**5.7**

Korisnici testa su u obavezi da zaštite bezbednost materijala testa u svakom trenutku.

**5.8**

Službe za bodovanje (*scoring services*) testova treba da dokumentuju procedure koje su primenjene kako bi se osigurala tačnost bodovanja. Potrebno je nadgledati učestalost grešaka u bodovanju i prijaviti ih korisnicima usluge na svaki smisleni zahtev. Treba ispraviti svaki sistematski izvor grešaka u bodovanju.

**5.9**

Kada ocenjivanje testa obuhvata ljudsku procenu, u uputstvima za bodovanje (*scoring rubrics*) treba precizirati kriiterijume za ocenjivanje. Potrebno je redovno nadgledati i proveravati poštovanje ustanovljenih kriterijuma za ocenjivanje. Treba dokumentovati postupke nadgledanja.

**5.10**

Kada se podaci o skoru testa objave učenicima, roditeljima, pravnim zastupnicima, nastavnicima, klijentima ili medijima, oni koji su zaduženi za programe testiranja treba da pruže adekvatna tumačenja. Tumačenja treba da opišu jednostavnim jezikom šta test obuhvata, uobičajene pogrešne interpretacije skorova testa i kako će se skorovi koristiti.

**5.11**

Kada se objave kompjuterska tumačenja protokola odgovora na testu, obrazloženje izvora i empirijska osnova za ova tumačenja treba da budu dostupni, a treba opisati i njihova ograničenja.

**5.12**

Kada se dobiju podaci o nivou grupe prikupljanjem rezultata parcijalnih testova koje su radili pojedinci, treba objaviti validnost i pouzdanost na nivou prikupljanja na kom se izveštava o rezultatima. Skorove pojedinaca ne treba objavljivati osim u slučaju da je uspostavljena validnost, uporedivost i pouzdanost tih skorova.

**5.13**

Prenošenje pojedinačno utvrđenih skorova testa neovlašćenim licima ili institucijama treba obaviti tako da se zaštiti poverljiva priroda skorova.

**5.14**

Kada se uoči materijalna greška u skorovima testa ili bilo koja važna informacija koju je objavila organizacija koja vrši testiranje ili neka druga institucija, u najkraćem mogućem roku treba dostaviti izveštaj o ispravljenom skoru svima onima koji bi u suprotnom mogli da koriste pogrešne skorove kao osnovu za donošenje odluka. Ispravljeni izveštaj treba označiti kao takav.

**5.15**

Kada su podaci sa testa o nekoj osobi sačuvani, i protokol testa i svaki pisani izveštaj takođe treba sačuvati u istoj formi. Korisnici testa treba da poštuju politiku i praksu čuvanja podataka svojih profesionalnih organizacija.

**5.16**

Organizacije koje čuvaju skorove testa pojedinaca u datotekama ili evidenciji o pojedincu treba da razviju jasan niz smernica svoje politike po pitanju trajanja čuvanja podataka pojedinca i po pitanju dostupnosti i korišćenju tih podataka tokom vremena.

**6. Propratna dokumentacija za testove**

**6.1**

Dokumentacija testova (na primer priručnici za test, tehnički priručnici, vodiči za korisnike i dodatni materijal) treba da bude dostupna potencijalnim korisnicima i drugim kvalifikovanim osobama u trenutku kada se test objavi ili pusti u upotrebu.

**6.2**

Dokumentacija testova treba da bude kompletna, precizna i jasno napisana tako da predviđeni čitalac može lako da razume sadržinu.

**6.3**

Potrebno je ponuditi obrazloženje testa, preporučenu upotrebu testa, podršku za tu upotrebu, kao i informacije koje pomažu pri interpretaciji skorova. Kada je realno moguće predvideti neke konkretne pogrešne upotrebe testa, treba upozoriti na njih.

**6.4**

Treba dokumentovati za koju populaciju je test namenjen, kao i specifikacije testa. Ako je moguće, treba opisati bazu stavki i postupke za izradu skale u odgovarajućim priručnicima za test.

**6.5**

Kada su dostupni statistički opisi i analize koje pružaju dokaze o pouzdanosti skorova i validnosti njihovih preporučenih tumačenja, te informacije treba uključiti u dokumentaciju. Kada su relevantne za tumačenje testa, dokumentacija testova redovno treba da sadrži informacije o nivou zadataka, granične skorove, normativne podatke, standardnu grešku merenja i opis postupka za izjednačavanje više formi.

**6.6**

Kada se test odnosi na kurs za obuku ili istraživanje, kurikulum, udžbenik ili paket uputstava, dokumentacija treba da obuhvata identifikaciju i opis kursa ili nastavnih materijala i treba da se naznači godina u kojoj su ti materijali pripremljeni.

**6.7**

U dokumentaciji testova treba precizirati kvalifikacije koje su potrebne za primenu testa i pravilnu interpretaciju skorova.

**6.8**

Ako je test osmišljen tako da treba da ga boduju i tumače ispitanici, izdavač i sastavljač testa treba da pruže dokaze da ispitanici mogu ispravno bodovati ili tumačiti test. Test osmišljen tako da ga boduju i tumače ispitanici treba da sadrži materijale za tumačenje koji pomažu pojedincu da razume skorove testa i koji su napisani na jeziku koji ispitanik razume.

**6.9**

U dokumentaciji treba navesti reprezentativni niz dostunih studija koje se tiču opštih i posebnih upotreba testa.

**6.10**

Materijali za tumačenje testova koji obuhvataju studije slučaja treba da pruže primere koji ilustruju raznolikost potencijalnih ispitanika.

**6.11**

Ako je test osmišljen tako da se može koristiti više od jednog metoda za primenu ili beleženje odgovora – kao što su beleženje odgovora u svesci sa pitanjima ili na posebnom listu za odgovore ili uz pomoć kompjuterske tastature – onda priručnik treba da dokumentuje u kojoj meri su uzajamno zamenljivi skorovi proistekli iz ovih metoda. Ako rezultati nisu uzajamno zamenljivi, ovu činjenicu treba objaviti i pružiti smernice za interpretaciju skorova dobijenih pri različitim uslovima ili metodama bodovanja.

**6.12**

Izdavači i službe za bodovanje koje nude kompjuterski generisana tumačenja skorova testa treba da pruže kratak pregled dokaza koji potkrepljuju ponuđena tumačenja.

**6.13**

Kada dođe do značajnih izmena na testu, dokumentaciju treba ispraviti, dopuniti ili revidirati kako bi podaci za korisnike bili aktuelni i kako bi se pružile dodatne korisne informacije ili upozorenja.

**6.14**

Na svakoj formi testa ili pratećoj dokumentaciji treba da stoji datum autorskih prava ili datum izdavanja.

**6.15**

Sastavljači testova, izdavači ili distributeri treba da pruže opšte informacije za korisnike testa i istraživače koji možda moraju da odrede prikladnost predviđene upotrebe testa u određenom kontekstu. Kada nije moguće opravdati određenu upotrebu testa, treba to jasno naglasiti u odgovoru na pitanje potencijalnom korisniku testa. Potrebno je takođe pružiti opšte informacije korisnicima testa i zakonskim starateljima koji moraju da daju saglasnost pre primene testa.

Pravednost pri testiranju

**7. Pravednost i pristrasnost**

**7.1**

Kada pouzdana istraživanja pokazuju da se skorovi testa razlikuju u značenju od jedne do druge podgrupe ispitanika u zavisnosti od vrste testa koji je u pitanju, onda, u kojoj meri je to izvodljivo, iste forme dokaza o validnosti koji su prikupljeni za populaciju ispitanika u celini takođe treba prikupiti i za svaku relevantnu podgrupu. Moguće je da se desi da se podgrupe razlikuju u pogledu adekvatnosti sadržaja testa, interne strukture odgovora na testu, odnosa skorova testa i drugih varijabli ili postupaka davanja odgovora koje je pojedinačni ispitanik primenio. Svaki takav nalaz treba dogledno razmotriti pri tumačenju i upotrebi skorova, kao i pri naknadnim revizijama testa.

**7.2**

Kada pouzdana istraživanja navode razlike u uticaju varijansi nevažnih za konstrukt u podgrupama ispitanika na učinak na nekom delu testa, test treba koristiti, ako uopšte treba, samo za one podgrupe za koje dokazi pokazuju da se mogu izvesti validni zaključci iz skorova testa.

**7.3**

Kada pouzdana istraživanja pokazuju da postoji različito funkcionisanje zadataka u zavisnosti od uzrasta, pola, rasne/etničke pripadnosti, kulturološke onesposobljenosti i/ili lingvističkih grupa u populaciji ispitanika po pitanju sadržaja oblasti koju test ispituje, sastavljači testa treba da sprovedu odgovarajuće studije kada je to izvodljivo. Takvo istraživanje treba da nastoji da otkrije i eliminiše aspekte izgleda, sadržaja i formata testa koji bi mogli da stvore pristrasnost u skorovima testa kod određenih grupa.

**7.4**

Sastavljači testa treba da teže da identifikuju i eliminišu jezik, simbole, reči, fraze i sadržaj koje generalno smatraju uvredljivim članovi rasnih, etničkih, rodnih i drugih grupa, osim ukoliko se proceni da je to neophodno za adekvatno predstavljanje oblasti.

**7.5**

Kada primena testa obuhvata pojedinačnu interpretaciju skorova testa, tj. ako nije u pitanju selekcija, skor ispitanika ne bi trebalo prihvatiti kao odraz pozicije na karakteristikama koje se ocenjuju bez razmatranja alternativnih objašnjenja za učinak ispitanika na tom testu u tom trenutku.

**7.6**

Kada se sprovode empirijske studije o različitim predviđanjima kriterijuma za članove različitih podgrupa, one treba da obuhvataju regresione jednačine (ili adekvatan ekvivalent) sračunate zasebno za svaku grupu ili tretman koji se razmatra ili analizu u kojoj su varijable grupe ili tretmana unete kao moderatorske varijable.

**7.7**

Prilikom primene testa gde nivo jezičke ili čitalačke sposobnosti nije deo odabranog konstrukta, jezički ili čitalački nivo koji test zahteva treba svesti na minimum neophodan za validnu procenu predviđenog konstrukta.

**7.8**

Kada su skorovi rastavljeni na delove i javno objavljeni za grupe određene po osobinama kao što su pol, etnička pripadnost, uzrast, nivo jezičke sposobnosti ili invaliditet, treba izneti i izveštaje upozorenja kad god verodostojna istraživanja pokazuju da skorovi testa možda nemaju uporedivo značenje u različitim grupama.

**7.9**

Kada se testovi ili procene predlažu za upotrebu kao instrumenti socijalne, obrazovne ili javne politike, sastavljači testa ili korisnici koji predlažu taj test treba u potpunosti i precizno da informišu kreatore politike o karakteristikama testova, kao i o svakoj drugoj relevantnoj i pouzdanoj informaciji koja je dostupna, a tiče se mogućih posledica upotrebe testa.

**7.10**

Kada upotreba testa rezultira ishodima koji utiču na životne ili obrazovne prilike ispitanika, treba ispitati dokaze o razlici srednjih vrednosti skorova testa između relevantnih podgrupa ispitanika, gde je to izvodljivo, za podgrupe za koje pouzdana istraživanja pokazuju razliku srednjih vrednosti na sličnim testovima. Tamo gde se pronađe razlika u srednjim vrednostima, treba sprovesti istraživanje kako bi se utvrdilo da te razlike ne proističu iz predstavljenog konstrukta ili varijansi nevažnih za konstrukt. Iako je to prvobitno bila dužnost sastavljača testa, korisnik testa snosi odgovornost za upotrebu testa sa drugim grupama osim onih koje je sastavljač naveo.

**7.11**

Kada se konstrukt može meriti na različite načine koji su približno jednaki po stepenu predstavljanja konstrukta i slobode od varijanse nevažne za konstrukt, treba uzeti u obzir dokaze o razlikama srednjih vrednosti skorova kod relevantnih podgrupa ispitanika prilikom odlučivanja koji test koristiti.

**7.12**

Proces testiranja ili ocenjivanja treba sprovesti tako da ispitanici dobiju sličan i ravnopravan tretman tokom svih faza testiranja ili ocenjivanja.

**8. Prava i obaveze ispitanika**

**9. Testiranje ispitanika različitog jezičkog porekla**

**9.1**

Postupak testiranja bi trebalo da bude osmišljen tako da smanji pretnje po pouzdanost i validnost zaključaka iz skorova testa koji mogu nastati usled jezičkih razlika.

**9.2**

Kada pouzdani dokazi istraživanja pokažu da se skorovi testa razlikuju u značenju po podgrupama jezički različitih ispitanika, onda, koliko god je to izvodljivo, sastavljači testa treba da prikupe istu formu dokaza o validnosti za svaku jezičku podgrupu koja se proučava, kao što je prikupljeno i za populaciju ispitanika u celini.

**9.3**

Prilikom testiranja ispitanika koji odlično govori dva ili više jezika na kojima je test dostupan, treba odrediti nivo jezičke sposobnosti tog ispitanika. Test bi uglavnom trebalo primeniti na jeziku koji ispitanik najbolje govori, osim ako poznavanje jezika koji ispitanik manje dobro govori nije deo procene.

**9.4**

Jezičke izmene koje preporučuju izdavači testa, kao i obrazloženje za te izmene, treba detaljno opisati u priručniku za test.

**9.5**

Kada postoje pouzdani dokazi o uporedivosti skorova kod regularnih i izmenjenih testova ili primena, ne treba staviti nikakvu naznaku pored skora. Kada nema takvih dokaza, treba pružiti konkretne informacije o prirodi izmena, ukoliko je to dozvoljeno zakonom, kako bi se omogućilo korisnicima testa da adekvatno tumače i delaju na osnovu skorova testa.

**9.6**

Kada je preporučeno da se test koristi kod jezički različitih ispitanika, sastavljači i izdavači testa treba da pruže informacije neophodne za adekvatno korišćenje i tumačenje testa.

**9.7**

Kada je test preveden sa jednog na drugi jezik, treba opisati metode za utvrđivanje adekvatnosti prevoda, a treba pružiti i empirijske i logičke dokaze o pouzdanosti skorova i validnosti zaključaka iz skorova prevedenog testa važnih za predviđene upotrebe u jezičkim grupama koje će biti testirane.

**9.8**

Prilikom testiranja za zaspošljavanje ili kredencijalizovanje, zahtevani nivo znanja iz jezika testa ne treba da prevaziđe onaj adekvatni nivo za odgovarajuće zanimanje ili profesiju.

**9.9**

Kada je predviđeno da verzije testa na više različitih jezika budu uporedive, sastavljači testa treba da prikažu dokaze o uporedivosti testa.

**9.10**

Zaključci o ispitanikovom opštem znanju jezika treba da se zasnivaju na testovima koji mere čitav niz jezičkih karakteristika, a ne samo na jednoj jezičkoj veštini.

11. Odgovornosti korisnika testa

12. Psihološko testiranje i procena

13. Testiranje i procena u obrazovanju

14. Testiranje prilikom zapošljavanja i kredencijalizovanja

15. Testiranje prilikom evaluacije programa i u javnoj politici

**9.11**

Kada se koristi usmeni prevodilac prilikom testiranja, on treba da tečno govori i jezik na kome je napisan test i ispitanikov maternji jezik, treba da poseduje stručno iskustvo u prevođenju, kao i osnovno razumevanje procesa procene.

**10. Testiranje osoba sa invaliditetom**

**10.1**

Prilikom testiranja osoba sa invaliditetom, oni koji sastavljaju, zadaju i koriste test treba da preduzmu mere da bi se osiguralo da zaključci iz skorova testa pravilno oslikavaju predviđeni konstrukt, pre nego neki invaliditet i njegove prateće karakteristike koji nisu predviđeni za merenje.

**10.2**

Ljudi koji donose odluke o prilagođavanju i izmenama na testu za osobe sa invaliditetom treba da budu upoznati sa postojećim istraživanjima o uticaju tih invaliditeta na učinak testa. Oni koji vrše izmene na testu takođe treba da poseduju odgovarajuće znanje iz oblasti psihometrije kako bi to radili.

**10.3**

Gde je to izvodljivo, test koji je izmenjen kako bi se koristio kod osoba sa invaliditetom treba probno testirati na pojedincima sa sličnim invaliditetom kako bi se ispitale adekvatnost i izvodljivost izmena.

**10.4**

Ukoliko su sastavljači testa izvršili ili preporučili izmene za ispitanike sa određenim invaliditetom, potrebno je detaljno opisati izmene i obrazloženje izmena u priručniku za test, a treba pružiti i dokaze o validnosti kada god su dostupni. Ukoliko nisu utvrđeni dokazi o validnosti za date zaključke za osobe sa određenim invaliditetom, sastavljači testa treba da ponude izveštaje upozorenja u priručnicima ili dodatnim materijalima u vezi sa pouzdanošću tumačenja zasnovanih na tim skorovima testa.

**10.5**

Tehnički materijali i priručnici koji prate izmenjeni test treba da sadrže i temeljan izveštaj o koracima pri izmeni testa kako bi upozorili korisnike na promene koje će verovatno uticati na validnost zaključaka donesenih na osnovu skorova testa.

**10.6**

Ukoliko sastavljač testa preporučuje određena vremenska ograničenja za osobe sa invaliditetom, treba koristiti empirijske procedure, kada god je to moguće, kako bi se utvrdila vremenska ograničenja za izmenjene forme takvih testova, pre nego da se jednostavno dozvoli ispitanicima da koriste više vremena nego što je to standardno predviđeno. Kada je moguće, treba ispitati zamor kao potencijalno važan faktor kada se prekorače vremenska ograničenja.

**10.7**

Kada veličina uzorka to dozvoljava, ustanova ili izdavač koji vrše izmene treba da podstaknu i prikažu validnost zaključaka donetih na osnovu skorova testa i pouzdanost skorova testa koji je primenjen na osobama sa različitim invaliditetima. Ta istraživanja treba da ispitaju efekat promena izvršenih za ljude sa različitim invaliditetima na dobijene skorove, kao i efekat primene standardnog neizmenjenog testa na njima.

**10.8**

Oni koji su odgovorni za donošenje odluka o upotrebi testa sa potencijalnim ispitanicima kojima će možda trebati ili koji će možda zatražiti određena prilagođavanja treba (a) da poseduju potrebne informacije kako bi adekvatno izabrali mere, (b) da poseduju aktuelne informacije o dostupnosti izmenjenih formi tog testa, (c) da informišu pojedince, kada je to adekvatno, o postojanju izmenjenih formi i (d) da ove forme stave na raspolaganje ispitanicima kada je to adekvatno i izvodljivo.

**10.9**

Prilikom oslanjanja na norme kao osnove za interpretaciju skorova pri ocenjivanju osoba sa invaliditetom, grupa korišćenih normi zavisi od svrhe testiranja. Regularne norme su adekvatne kada je svrha funkcionisanje ispitanika u odnosu na opštu populaciju. Ukoliko je moguće, treba koristiti normativne podatke populacije pojedinaca sa istim nivoom ili stepenom invaliditeta kada se promatra funkcionisanje ispitanika u odnosu na osobe sa sličnim invaliditetom.

**10.10**

Sve usvojene izmene na testu treba da budu adekvatne za pojedinačnog ispitanika, a da se pritom zadrže sve moguće standardizovane karakteristike. Stručnjak za testove treba da razmotri realno dostupne informacije o iskustvu, karakteristikama i sposobnostima svakog ispitanika koji bi mogli da utiču na učinak testa i da dokumentuje argumente za izmene.

**10.11**

Kada postoje pouzdani dokazi o uporedivosti skorova pri regularnoj i izmenjenoj primeni, ne treba stavljati nikakvu naznaku pored skora. Kada nema takvih dokaza, treba pružiti konkretne informacije o prirodi izmena, ukoliko je to dozvoljeno zakonom, kako bi se omogućilo korisnicima testa da adekvatno tumače i delaju na osnovu skorova testa.