

Tószegi Zoltán

A NYELVI TESZTELÉS MINT MÉRÉS

BEVEZETÉS

Magyarországon a nyelvi tesztelést (mint összetett tevékenységet) még mindig túlnyomórészt nyelvtanárok végzik. Talán ez az oka, hogy még nem alakult ki a nyelvi tesztelés precizitásával szemben támasztott olyan követelményrendszer, amely minden más mérési helyzetben teljesen természetes.

A nyelvtanároktól általában automatikusan elvárják, hogy nyelvi tesztelésben (közkeletűbben: nyelvvizsgáztatásban) is részt vegyenek, többnyire szakirányú végzettség nélkül, jó esetben legalább minimális felkészítéssel. Pedig a nyelvi tesztelés, azaz a nyelvi képességek mérése - a nyelvtanítással ellentétben - a pszichometria része. A pszichometria a pszichológiai jellemzők mérésének tudománya. A nyelvi tesztelés tehát: mérés. Ezen dolgozat célja megvizsgálni, milyen követelményeket támaszt ez a megközelítés a nyelvi tesztekkel, a nyelvi teszteléssel szemben.

A MÉRÉS

Mérésnek azt nevezzük, amikor számszerűleg kifejezünk valamely tulajdonságot. Ilyenkor egy ismert dolog ismert számszerűsített tulajdonságához viszonyítunk egy mennyiségileg ismeretlen tulajdonságot, s a mérés eredménye egy szám és egy mértékegység lesz. A méréssel szemben támasztott legfontosabb követelmény a hitelesség. Azokban a mérési helyzetekben, amelyekkel a mindennapi életben leggyakrabban találkozunk, a hitelességet az objektivitás garantálja. Ez azt jelenti, hogy az ugyanazon körülmények között elvégzett mérések ugyanazt az eredményt kell, hogy adják, függetlenül attól, hogy a mérést ki végzi. (Bonyolultabb, esetleg speciális szakértelmet igénylő mérőműszerek esetében természetesen kritérium az adott műszer kezelésére feljogosító végzettség, mint az objektivitás alapvető feltétele.)

A mérés akkor eredményes, ha eredménye minden érintett számára értelmezhető. Ennek előfeltétele, hogy a mérés eredménye mindenki számára ugyanazt jelentse. Ehhez olyan mértékegység kell, amelyet minden érintett ismer, s melynek nagyságrendjével (amennyiben az értelmezhető) is tisztában van. Ezt biztosítják a szabványok, amelyek egységessé teszik a mértékegységeket, a mérőműszerek és -eszközök, valamint a méréshatárok használatát. Amennyiben valamely területen bizonyos okból ez nem teljesíthető maradéktalanul, legalább a különböző mértékegységek összehasonlíthatóságát kell garantálni. Minél szélesebb körben fontos egy bizonyos fajta tulajdonság mérése, illetve annak eredménye, annál fontosabb az egységes szabvány bevezetése és alkalmazása.

A mindennapi életben két alapvető mérési helyzettel találkozunk. Bizonyos fizikai jellemzők viszonylag nagy pontossággal megbecsülhetők (akár pusztán érzékszerveink segítségével), s aránylag egyszerű eszközökkel igényeinknek megfelelő pontossággal mérhetők. Ilyen a hosszúság, a tömeg, a hőmérséklet, de akár az idő vagy a sebesség is. Más dolgok nem ennyire kézzelfoghatóak, s csak hatásuk alapján mérjük őket olyan mérőeszközökkel, melyek használata könnyen elsajátítható, s végzettséget sem igényel (pl. feszültség vagy keréknyomás). Akadnak olyanok is (pl. a sűrűség), amelyek több különböző mérés eredményéből következtethetők ki, vagy vezethetők le. Ezek mindegyikére jellemző, hogy a gyakorlati életben jelentőséggel bír, és azt legjobban szolgáló méréshatárok között mérjük őket.

Az átlagember mindennapi életét nem közvetlenül érintő, illetve az átlagember által nem, vagy csak nehezen mérhető dolgok mérését szakemberek végzik speciális mérőműszerekkel és mérőeszközökkel. A pszichometria tárgyát képező tulajdonságok mérése szintén ebbe a kategóriába tartozik.

A NYELVI TESZTELÉS

Az emberi tulajdonságok mérését általában nem rutinszerűen végezzük. Bár többnyire nagyon szeretnénk többet tudni a másik emberről, s a tulajdonságok számszerűsítése bizonyos helyzetekben nagyban megkönnyítené a döntéshozatalt, nyilvánvaló, hogy az ember pszichológiai jellemzői túlságosan összetettek ahhoz, hogy egyszerű, mindenki által hozzáférhető és hiteles módszerekkel, eszközökkel mérhetőek legyenek. Tovább nehezíti a dolgot, hogy a pszichológiai jellemzők sokszor egymással szorosan összefüggnek. Elég csak arra gondolnunk: a vizsgán nyújtott teljesítmény nem kizárólag a felkészültség függvénye; az úgynevezett „jó vizsgázók” könnyebben adnak valós képet megszerzett ismereteikről, mint azok, akik „rossz vizsgázók”. Az állandó jellemzők mellett időlegesen jelen lévő tényezők (pl. fáradtság, éhség, esetleg a vizsgáztató iránti szimpátia, esetleg ellenszenv) is komolyan befolyásolhatják a végeredményt.

A nyelvi készségek mérésénél már az első lépés komoly kihívást jelent, hiszen definiálni kell, mit is akarunk mérni. A válasz – „nyelvtudást” – csak látszólag egyszerű.

Nyelvtudást optimális esetben nem öncélúan mérünk; az eredmények fontos döntések meghozatalában hivatottak segíteni a mérés megrendelőit. Mind mennyiségileg, mind minőségileg más nyelvtudásra van szüksége mondjuk egy vadászpilótának, egy légiirányítónak, egy nemzetközi törzsben szolgáló főtisztnek, vagy egy külszolgálaton lévő egészségügyinek. Ugyanazon teszttel, mint mérőeszközzel végzett mérések nem szolgáltathatnak egyaránt használható információval mindezen személyek nyelvtudását illetően. A megfelelő mérőeszköz megalkotása mindegyik esetben más kihívást jelent, de mindegyikhez alapvetően szükséges kétféle tudás. Az egyik a nyelvtudásnak, mint összetett pszichológiai jellemzőnek a mélyreható ismerete, mely jóval több, mint az adott nyelv ismerete. A másik a nyelvi tesztek, mint mérőeszközök létrehozásával, működésével és használatával kapcsolatos általános elméleti és gyakorlati ismeretek összessége – mely optimális esetben tapasztalattal is párosul. Mindezekből következik, hogy a nyelvtudás mérése a méréseknek azon csoportjába tartozik, amelyeket hitelesen csak megfelelő végzettséggel lehet végezni.

Amennyiben megfelelően sikerült definiálni az adott mérési kontextusban a mérés tárgyát, a következő lépés az azt (és lehetőleg csakis azt) az igényeknek megfelelő határok között mérő eszköz, a teszt megalkotása. Ez a mérendő alapkészségektől (beszéderítés, beszéd, olvasás, írás) függően feladatokból¹, illetve azokon belül kérdésekből² áll. Több feladatra/kérdésre azért van szükség, mert a nyelvtudás az adott kontextusban alkalmazott definíciótól függetlenül egy komplex tulajdonság. Ennek minden egymástól elkülöníthetően mérhető komponenséről szükséges mérési adatokat gyűjteni. Végleges formájában a teszt az ezeket mérő feladatok és kérdések összessége lesz, mely olyan struktúrát alkot, amelyben minden feladat/kérdés azt (és csakis azt) méri, aminek mérésére tervezték.

A nyelvi tesztelés lényegesen eltér a nem pszichometriai mérésektől. Mivel az utóbbiak jól definiálhatók, a mérésekre megalkotott eszközök térben és időben csaknem univerzálisan használhatók. A technikai fejlődés (pl. a digitalizálás) pontosabbá, megbízhatóbbá teheti a méréseket, de ez a mérendő tulajdonság definícióját nem változtatja meg. Felhasználói szinten ezért általában semmi más nem érdekes, csak a hitelesített eszközzel kompetens személy által elvégzett mérés eredménye.

Ezzel szemben a nyelvtudás definíciója az adott kontextustól függ, s egyedi mérőeszköz használatát teszi szükségessé. Ebből viszont következik, hogy a különböző tesztek alkalmazásukban is eltérnek. Ezért fontos megvizsgálni mind a nyelvi teszt megalkotásának, mind pedig alkalmazásának folyamatát.

A TESZTÍRÁS

Fentebb már megemlítettük a nyelvtudás négy legáltalánosabban elfogadott összetevőjét (beszéderítés, beszéd, olvasás, írás), melyek maguk is komplex tulajdonságok. Könnyen belátható, hogy különböző módon mérhetők ezek a készségek, csakúgy, mint alkotóelemeik. Amikor ezek mérésére alkalmas

¹ Az angol nyelvű szakirodalomban: *task*

² Az angol nyelvű szakirodalomban: *item*

feladatokat, kérdéseket készítünk, két dolgot kell pontosan meghatároznunk: mit mérünk, és milyen szinten (más szóval milyen méréshatárok között) mérünk.

Amit mérünk

A tesztek tartalmi részének megalkotása viszonylag egyszerűen leírható folyamat. Első lépcsője, hogy valaki úgy gondolja: valamely döntés meghozatalához nyelvtudás mérésére van szükség. A döntéssel kapcsolatos szituációból adódik egy nyelvi célkörnyezet. Ennek elemzése után a tipikusan használt vagy szükséges nyelvi készségek összessége szolgáltatja majd a nyelvtudás adott kontextusban használható definíciójának alapját. A kimerítő definíció eredménye a nyelvtudás csak eme kontextusban értelmezhető modellje. A teszt készítése során e modell egyes elemeinek mérésére alkalmas feladatokat és kérdéseket kell alkotni. A megfelelő szakmai háttérnek itt alapvető szerepe van. Az írott illetve beszélt nyelv értésének méréséhez olyan szövegeket kell találni (autentikus teszt) vagy alkotni (nem autentikus teszt), amelyekkel a vizsgázók a nyelvi célkörnyezetben találkoznak. Az írás- és beszédkészséget mérő feladatokban olyan nyelvi produktum előállítását kell megkövetelni (illetve lehetővé tenni), amelyet a vizsgázóktól majd a célnyelvi környezetben is elvárnak. A célnyelvi környezet illetve a feladatok/kérdések ilyen egyértelmű megfeleltetése bonyolult nyelvészeti feladat.

A végzettség és a tapasztalat esetén alkalmazása azonban csak az első lépcső a hiteles teszt megalkotásának folyamatában. A teszt minden egyes elemének több lépcsőben bizonyítania kell, hogy megfelel a követelményeknek. Először a tesztelők maguk értékelik őket nyelvészeti szempontból. Ezután ugyanők, vagy kollégáik kipróbálják őket, s ekkor új, immáron mérési (de még mindig tartalmi, minőségi) szempontból értékelik őket. A következő lépcső a célcsoport, a majdani vizsgázók egy kis mintáján való kipróbálás, és az általuk történő értékelés, majd az utolsó lépcső egy nagyobb, lehetőleg reprezentatív mintán való kipróbálás. Mindezen szinteken komoly tartalmi visszajelzés érkezik a feladatokkal, kérdésekkel kapcsolatban, mely indokolt esetben azok módosítását, majd ugyanazon szinten való ismételt kipróbálását eredményezi. Ez a folyamat egyúttal a mérőeszköz alkalmazásának próbáját is jelenti. A próbák lebonyolítási rendje, szabályai, a vizsgázók és vizsgáztatók viselkedése szintén megfigyelés tárgya, s a folyamat végére összeáll az optimális vizsgarend, amely lehetővé teszi, hogy a vizsgahelyzet csak és kizárólag a mérni kívánt nyelvtudás méréséről szóljon.

Mindezek értékelése részben a szaktudás függvénye. Ez azonban még mindig meglehetősen szubjektív, a mérésekkel szembeni általános követelményeket messziről közelítő eljárás. Ezért minden ilyen szintű döntést dokumentálni kell. A tesztelők tisztában vannak a szubjektivitás szerepével, és azzal, hogy a legjobb teszt is (jó esetben) csak konvergál ahhoz, amely általában egy mérési helyzetben ideálisnak mondható. A hitelességet itt két lépcsőben biztosítják. Az első a teljes szakmai átláthatóság biztosítása folyamatos és minden részletre kiterjedő dokumentáción keresztül. Így bármilyen szakmai vizsgálat megállapíthatja, hogy a teszt kidolgozása során az adott körülmények között elvárható legjobb, szakmailag legmegalapozottabb döntések születtek-e. A második lépcsőben megjelenik az objektivitás tudományos megalapozottságának igénye. Ezért a visszajelzések és eredmények számszerűsíthető összetevőit statisztikai módszerek segítségével vizsgálják. A vizsgálatok során többek között kiderül, hogy van-e átfedés a különböző kérdések által mért összetevők között, vagy hogy a lebonyolítási rend, esetleg a kérdések tartalma diszkriminál-e valamilyen nem a nyelvtudáshoz tartozó tulajdonság alapján. A feltárt hiányosságokat megszüntetik, a hibákat kijavítják mielőtt újabb próbára kerül sor.

Mértékegység, méréshatárok

A mérés megrendelője az eredményekkel kapcsolatos döntések természetétől függően eldöntheti, milyen rendszerben vagy mértékegységben kívánja a mérés eredményét kifejezni. Ez történhet „megfelelt”/„nem felelt meg” kategóriák, „A” szint, „B” szint, „C” szint, stb., esetleg pusztán a megszerzett pontszámok szerint (esetenként valamilyen konverziót követően). Hogy a kiválasztott mértékegységgel kifejezett értékeknek milyen nyelvtudás felel meg, az a méréshatárok, illetve a mérési tartomány diszkrét elemeinek definícióján múlik. Ez *minden esetben* szubjektív döntés eredménye. A hitelesség egyik záloga itt is a dokumentáció. A másik pedig objektív bizonyíték szolgáltatása arra, hogy az egyes feladatok/kérdések valóban a definiált szinten mérnek. (A vizsgázók szempontjából köznyelven ezt hívjuk a feladat/kérdés nehézségének.) Ez szintén statisztikai módszerekkel történik, s a számítások eredménye az azok alapján meghozott döntésekkel együtt

szintén precízen dokumentálandó. Nyilvánvaló, hogy a mérés eredményességéhez még mindig szükség van egy ismert tulajdonságra, esetünkben nehézségi szintre. Egy új teszt elkészítésénél ez értelemszerűen soha nem lesz teljesen objektív, ám a szubjektivitás szintje eredményesen csökkenthető különböző megközelítésekkel.

Ha a mérni kívánt nyelvi szint definíciója köthető egy tanulói, vagy hallgatói csoport tagjainak nyelvi szintjéhez, akkor e csoportból kiválaszthatók úgynevezett mesterek³. A mesterek olyan egyének, akik nyelvtudásának szintjét megfelelő pontossággal ismerjük. A tanulói csoportokban ugyanis a folyamatos kontaktus lehetővé teszi a nyelvtudás folyamatos, ezáltal sokkal pontosabb és hitelesebb értékelését. Más szóval: a tanár, aki sokkal több és részletesebb információval rendelkezik a diákok nyelvtudásáról, mint amelyet bármely teszt egyszeri lebonyolítása szolgáltathat, jobb közelítéssel meg fogja tudni határozni bármelyikük nyelvi szintjét. Ha ebben a környezetben értelmezhető a nyelvtudásnak az adott mérési helyzethez kialakított definíciója, a (legalábbis jobban) ismert nyelvi szintű vizsgázóknak a próbán nyújtott teljesítménye megfelelő információt szolgáltathat az egyes kérdések nehézségét illetően. Ezt az eljárást kalibrálásnak nevezzük. A kalibrált (ismert nehézségű) kérdések segítségével már a többi kérdés nehézsége is meghatározható, ezek kombinációjával pedig a végleges teszt nehézsége előre meghatározott értékre állítható be. Az ezen elért eredmény a vizsgázónak az adott körülmények között hitelesen mért nyelvtudása.

A mesterek adott esetben olyan érvényes bizonyítvánnyal rendelkező egyének is lehetnek, amely a nyelvtudás elégséges mértékben hasonló definícióját alkalmazó teszthez köthető. Teljes megfeleltetésre nincs szükség, hiszen ennek fennállása azt jelentené, hogy az új teszt megírása fölösleges.

Homlokegyenest másik irányból kiindulva olyan kérdések⁴ előállítása is lehet a cél, amelyek nehézsége ismert, s ezekből már levezethető a többi kérdés nehézségi foka. Ez azonban igen költséges és időigényes folyamat lenne, mivel az ideális mérési körülményekhez való konvergencia határait kellene feszegetni hozzá, ami mind az igénybevett szaktudás, mind a statisztikai módszerek alkalmazásának maximalizálását jelenti. A nyelvtudás mérését megrendelőknak azonban aligha állnak rendelkezésre az ehhez szükséges anyagi források, illetve idő. A megoldás olyan cégek szolgáltatásainak igénybevétele, amelyek kifejezetten a pszichometria különböző területein szükséges mérések elvégzésére alkalmas feladatok/kérdések előállításával foglalkoznak. Ezúton, bár a minőségért általában ennek megfelelő magas árat kell fizetni, időt és költséget lehet megtakarítani.

Bármelyik módszert is alkalmazzák, a döntés mechanizmusát, illetve az eljárás részleteit dokumentálni kell. A mérés eredményessége és értelmezhetősége érdekében ugyanis alapvető fontosságú, hogy az adott mérési helyzet *minden egyes* érintettje számára megfelelő mélységben ismerje a mért tulajdonság tartalmával, a mértékegységgel és a méréshatárokkal kapcsolatos döntéseket, s ezek a tesztelői és vizsgázói társadalom e szegmensére számára *minden kétséget kizáróan* egységes tartalommal bírjanak. Ettől lesz a mérés mindenki számára hiteles.

A TESZTEK ALKALMAZÁSA

A nyelvi tesztek alkalmazása, használata lényegesen eltér a megfelelő eljárások alkalmazásával (esetleg tömegesen) legyártott és hitelesített mérőműszerektől. Az előbbi során a mérő és a mérendő oldalon is emberek vannak, ami még mindig nagyon sok hibalehetőséget rejt magában. A hibák bármelyike pontatlan mérést, a valós értéktől eltérő mért értéket eredményezhet, amelynek súlyos következményei lehetnek mind a vizsgázó, mind a mérés megrendelője szempontjából. Ráadásul az ilyen hibák a leghitelesebb, a legszigorúbb szakmai követelményeknek megfelelő tesztet is hiteltelenné teszik, annak elkészítőivel együtt. Ezért nem csak az a fontos, hogy a próbák során a tesztek alkalmazásának minden egyes lépését a legapróbb részletekig kidolgozzák, hanem az is, hogy a teszt elkészítői mellett mindenki, aki bármilyen szinten részt vesz a tesztek alkalmazásában, a vizsgák lebonyolításában, s nem utolsósorban az értékelésben, tisztában legyen munkája fontosságával, s rendelkezzen az ehhez szükséges ismeretekkel, előírásokkal. Ennek különböző szintjei lesznek. Természetes, hogy aki „csupán” egy írásbeli vizsgán felügyel, annak nem kell behatóan ismernie a

³ Az angol nyelvű szakirodalomban: *master*

⁴ Az angol nyelvű szakirodalomban: *anchor items*

tesztírás minden aspektusát. Tudnia kell viszont, mit *kell*, mit *szabad*, és mit *tilos* a vizsgázóknak tenniük, amíg a teremben tartózkodnak, illetve mi vezet a vizsgáról való azonnali kizáráshoz.

A tesztek még egy lényeges ponton különböznek a többi mérőeszköztől. Míg az utóbbiak csak akkor kerülnek forgalomba, ha optimális teljesítményre (precizitás, megbízhatóság) képesek, a tesztek teljesítményük maximumát nagy valószínűség szerint csak használatbavételük után érik el. Ennek magyarázata a nyelvi tesztek hitelességének egyik komponense, az úgynevezett jósló-hitelesség. A nyelvi képességeket ugyanis az esetek nagy részében azért mérjük, hogy az eredményből következtethessünk egy későbbi élethelyzetben (a nyelvi célkörnyezetben) nyújtott teljesítményre. Az azonban, hogy a teszt milyen precizitással képes megjósolni ezt, igazán csak akkor derül ki, amikor a mérési eredmények alapján kiválasztott vizsgázók már lehetőséget kaptak annak bebizonyítására, hogy nyelvi szempontból valóban alkalmasak a célnyelvi környezetben megoldandó feladatok végrehajtására. A nyelvi tesztek teljesítményét emiatt hosszú távon is figyelemmel kell kísérni, s a rendelkezésre álló információt szintén dokumentálva, a tapasztalatokat figyelembe véve tovább kell erősíteni az ideálshoz való konvergenciát. A tesztelési ciklus tehát túlnyúlik a teszt használatba vételén. Ez a magyarázata annak, hogy a nyelvi tesztek hitelességét leíró általános definíció kiterjed az eredmények alapján meghozott döntésekre is.

ÖSSZEGZÉS

E dolgozat célja az volt, hogy olyan szempontból vizsgálja a nyelvi tesztelést, amely viszonylag kevésbé van jelen a köztudatban: a nyelvi tesztelést mint mérést kívánta bemutatni. Megvizsgáltuk, hogy bizonyos a méréssel kapcsolatos alapfogalmaknak mik felelnek meg a nyelvi tesztelés kontextusában. Feltártuk azokat a tényezőket, amelyek a bennük rejlő szubjektivitás miatt veszélyeztetik a mérés pontosságát. Mindeközben számba vettük azokat az eljárásokat, amelyek a nyelvtudás mérésének hitelessége kialakítását, biztosítását, illetve megőrzését célozzák. Ezek közül alapvető fontosságú a megfelelő szakirányú végzettség és tapasztalat megszerzése, az objektív bizonyítékok szolgáltatása statisztikai módszerek segítségével, valamint a részletes és átlátható dokumentáció. Ez azt a célt szolgálja, hogy mindenki, aki a nyelvi képességek egy bizonyos kontextusban történő mérésében közreműködik, illetve abban érintett, érdekeinek megfelelő mélységű tájékoztatást kapjon. Kiemelten fontos ez annak érdekében, hogy az esemény minden résztvevője egységes szemlélettel rendelkezzen, és a mérés eredménye mindannyiuk számára ugyanazt jelentse. Mindezek alapfeltételei annak, hogy a mérés eredménye hiteles legyen, valóban a vizsgán demonstrált nyelvtudást tükrözze, és segítse a mérés megrendelőit az ezzel kapcsolatos helyes döntés meghozatalában.